INVESTIGATING THE USABILITY OF TOUCH-BASED USER INTERFACES

MD ABUL FADLE MAIDIN

School of Computing, Science and Engineering,
College of Science and Technology,
University Of Salford, Salford, UK

Submitted in Fulfilment of the Requirements of the
Degree of Doctor of Philosophy, December 2014

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my special appreciation and thanks to my supervisor Dr Pietro Murano for his unlimited support both academically and personally. His invaluable guidance and unfailing support at all stages of this work fundamentally helped me refined this study.

I thank the Brunei Darussalam Government for offering me a full studentship. Without the financial support, my academic pursuit would not have been possible.

I am indebted to Dr Adam Hj Jait for the permission to use the experiment room for the research at Seri Begawan Religious Teachers University College, Brunei Darussalam. I am deeply thankful to the undergraduate and postgraduate students for their participation in my research.

A special thanks to my family. Words cannot express how grateful I am to my mother, my father-in-law, mother-in law and family, for all of the sacrifices that you’ve made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank my beloved wife, Dk Hjh Siti Norainna Pg Hj Besar for her unwavering support. To my beloved children, Siti Nouru Badreen and Md Adam Malek, I would like to express my thanks for being such a good girl/boy.

ABSTRACT

With the emergence of pen-and-touch operated personal digital assistants (PDAs), tablet computers, and wall-size displays (e.g., Liveboard and Smartboard), touch and pen input have gained popularity. Touch-based user interfaces such as mobile phones, PDAs and tablet PCs (with touch screens) have become more attractive in consumer electronics because they enable quick learning and rapid performance whilst evoking high user satisfaction. Today, countless supermarket checkouts, restaurant tills, automated-teller machines, airport check-in kiosks, museum information-booths and voting kiosks use touchscreens.

Nevertheless, initial literature identified that the widespread use of a touch-based user interface has been limited by the high error rates shown in many studies, the lack of precision, the fatigue in arm motion, and the concern for screen smudging.

Furthermore, most research into touch-based interaction has tended to not directly investigate efficiency, effectiveness and user satisfaction. There is therefore a need to add to the body of knowledge in this area, especially as devices using touch-based interaction are becoming more pervasive.

Hence, the purpose of this research is to evaluate the usability of touch-based user interfaces in terms of efficiency, effectiveness and user satisfaction. In order to answer the question of whether a touch-based user interface is better - more effective, useful, practical and satisfying to the user -, an investigation of comparison to other, alternative interaction methods, by means of mouse, touch and stylus has been conducted.

Therefore, the research sets out to concentrate on a series of empirical experiments that will be designed and developed to evaluate the efficiency, effectiveness and user satisfaction of using touchscreen interfaces. Furthermore, in order to collect the human performance data, a series of small software prototypes involving touch-based interaction were developed and designed using Adobe Flash.

Initially a pilot experiment is carried out and followed by the abstract experiment and context experiment that were based on the guidance of The International Organization for Standardization known as ISO (ISO 9241-420, 2011).

The abstract experiment consist of four tests (Tracing test, Dragging test, One direction test and Multi directional test) which are deliberately developed as abstract tasks with the purpose of analysing the user's ability on simple tasks without a real world context.

The context experiment consist of four tests as well (Tracing test, Dragging test, One direction test and Multi directional test) which are deliberately developed as contextual tasks with the purpose of analysing the user's ability in a real world context.

Overall, the aim of both abstract and context experiments was to discover if there are differences in mouse, stylus and touch on the tracing test and dragging test with different levels of difficulty that could affect users' performance and satisfaction.

The significant contribution to knowledge that may arise from this research might provoke the gaining of evidence to show if touch-based interaction is more effective and preferred by users in real-world-type tasks and scenarios. Currently there is very little evidence to indicate whether touch-based interaction is more effective and preferred by users. It seems that the proliferation of touch-based devices is market-driven rather than usability-driven. Moreover, this is the first study that has been carried out which compare three input devices (stylus, mouse and touch) in tracing, dragging, one direction tapping and multi directional tapping test for both abstract and context tasks and therefore contributes to the up-to-date HCI literature.

The main strength of the current study is that it provides findings from well-designed experiment that is based on ISO standard (ISO 9241-420, 2011). It provided a useful guideline that can be further developed and applied to other research in this area.

CHAPTER 1: INTRODUCTION

1.1 AREA OF STUDY

Human-computer interaction is an enormous discipline that helps the researchers facilitates effectual development of latest technologies. Hundreds of thousands of technological developments have been made for the ease and comfort of humans so that they can perform their activities safely and productively; this is the concept that comes from the term human-computer interaction. It is all about how we use technology and interact with the computer. We can simply say that human-computer interaction is a study of design, implementation and performance of computing systems and devices for human use. In addition, it allows humans to communicate with computers in a more effective manner (Lazar *et al.*, 2010).

Interaction with computers is a simple thing that we do almost every moment of the day; we use our mobile phones, play games on iPads, use a microwave and watch television. These are technological developments that we, as humans, use – and so, in this way, we interact with technology. Here a question arises: are all the computing devices easy to use or do people face some difficulty while using the systems? Are some considered as better user interfaces than others? Good and Bad HCI mainly depend on system efficiency and user capabilities. For example, some websites are easy to use whereas the design of others is not appealing enough; similarly, the design of a system varies, but nevertheless, it is still the responsibility of the designer to ensure system usability and functionality, improving a pleasant user experience. The methodology of HCI in system design and analysis is to get feedback from the user experiences and to modify the design by considering the recommended changes in order to make the system acceptable and get more positive feedback.

HCI deals with two important features named ‘visibility’ and ‘affordance’. Firstly, the interface of a system should be visible, with a clear map of its design, effects and functionalities. Secondly, ‘affordance’ refers to the functional requirements of a system in reference to what it is meant to do exactly, and to how a person thinks to use it.

‘Affordance’ can also refer to characteristics; for instance, if a door affords opening and a chair affords support, it is considered perfect from a design point of view as it gives the desired outcome (Soloway *et al.*,1994).

HCI considers human disabilities, limitations and design; it uses special hardware devices and software packages that can offer happiness to the special users. Braille keyboards, software-readable webpages and screen magnifiers are designed for users with low vision, while a written edition of audio material, or headphones, are designed for those with impaired hearing. Graphical user interfaces, menus, forms and uses of natural language are beneficial aspects considered for designing a better interface for people of all ages (Rogers, 2012).

Precisely, it is difficult for humans to succeed in their lives without interacting with technology. A good user interface is mainly based on ages, cultures, education, disabilities and accessibilities. New technologies and good HCI can bring automated support to latest domains, offer new ways to interact with the computing systems and remove obstacles to democratize contribution in a large cross-section of tasks and activities (Muller, 2003).

With the emergence of pen-and-touch operated personal digital assistants (PDAs), tablet computers, and wall-size displays (e.g., Liveboard and Smartboard), touch and pen input have gained popularity (Baudisch *et al.*, 2003). Potter *et al* (1988) shared the same view - that touch-based user interfaces such as mobile phones, PDAs and tablet PCs (with touch screens) have become more attractive in consumer electronics because they enable quick learning and rapid performance while evoking high user satisfaction. Today, countless supermarket checkouts, restaurant tills, automated-teller machines, airport check-in kiosks, museum information-booths and voting kiosks use touchscreens (The Economist, 2008). Nevertheless, the widespread use of touch user interfaces has been limited by the high error rates shown in many studies - the lack of precision, fatigue in arm motion and concern for screen smudging (Ostroff & Shneiderman, 1988; Pickering, 1986).

Thus, the focus of this project is to evaluate the usability of touch-based user interfaces in terms of efficiency, effectiveness and user satisfaction. In order to answer the question of whether a touch-based user interface is more effective, useful, practical, satisfying to the user and generally better, an investigation of comparison with other, alternative interaction methods, like means of mouse, touch and stylus, were conducted.

1.2 OBJECTIVES

As there has been little research of the usability of touch-based user interfaces in terms of efficiency, effectiveness and user satisfaction, the focus of the study concentrates on a series of empirical experiments that were designed and developed to evaluate the efficiency, effectiveness and user satisfaction in the use of touchscreen interfaces. These involved comparisons of mouse, stylus and touch. Additionally, the experiments were designed to have as much ecological validity as possible, by being appropriately contextualised. Therefore, the research sets out:

1. To design and develop a series of small software prototypes involving touch-based interaction.
2. To design and develop a series of empirical experiments.
3. To evaluate the efficiency, effectiveness and user satisfaction of using touchscreen interfaces in relevant contexts in comparison with mouse, stylus and touch interaction.
4. To statistically analyse data collected from such experiments and reach appropriate conclusions regarding efficiency, effectiveness and user satisfaction.

1.3 CONTRIBUTION TO KNOWLEDGE

The significant contribution to knowledge that may arise from this research might provoke the gaining of evidence to show if touch-based interaction is more effective and preferred by users in real-world-type tasks and scenarios. Currently, there is very little evidence to indicate whether touch-based interaction is more effective and preferred by users. It seems that the proliferation of touch-based devices is market-driven rather than usability-driven.

1.4 STRATEGIES

An empirical approach was used in a series of controlled experiments with carefully designed tasks aiming to produce suitable data for analysis, so as to determine efficiency, effectiveness and user satisfaction. Willing participants had to complete a consent form as part of the process. Information/data was collected by using observational techniques and software capture and/or video recording. At the end of each experiment the findings were used to establish guidelines for possible future experiment.

1.5 RATIONALE

Research studies of human-computer interaction have failed to reveal whether one technique is superior to another. For example, studies considering functional usage of the mouse, stylus and touch focused on compatible human motion. Yet, comparisons of the three in terms of user comfort and satisfaction are seriously lacking. There is, therefore, a major need to add to the body of knowledge in this area since advancing technology is becoming more pervasive.

1.6 AIM OF THE PROTOTYPE DEVELOPMENT

The aim of developing the prototype is to investigate the usability of touch-based interfaces compared with more conventional means of interaction such as a mouse, stylus and touch. Efficiency and effectiveness in terms of accuracy, errors, speed and user satisfaction are the main factors considered.

1.7 CHAPTERS OVERVIEW

Following the introductory Chapter 1, Chapter 2 provides literature review of the usability requirement and human-system interaction in collaboration with ISO as well as reviewing related study. Chapter 3 gives a methodology review consisting of the experiment design guideline of touch based interfaces and how experiments are carried out.

Chapter 4 presents a pilot experiment that consists of 4 design version prototypes. The main task of the prototype is to enable the user to merge 2 boxes (black boxes and red boxes) by dragging and dropping. Namely prototype version 1; flick (difficulty level: easy), prototype version 2; auto merges if both touch 80% of each other (difficulty level: medium), prototype version 3; auto merges if both touch 100% of each other which means they need 100% merge accuracy (difficulty level: hard) and prototype version 4; auto merges if both touch 100% of each other which means they both need 100% merge accuracy, in which the black box must be moved along a path (difficulty level: hard).

In Chapter 5, four abstract design version tests were investigated. The first test is the tracing test which evaluates the tracing of an object. The second test is the dragging test which evaluating clicking and dragging objects to specific locations. The third test is the one-direction tapping test in which the evaluating points are movement along one axis on a horizontal rubber banding, an insert cursor at points along a character string and selecting information in columns or rows. The fourth test is the multi-direction tapping test which evaluates pointing movements in many different directions on a repositioning of the pointer at different areas on the screen, cell selection in a spreadsheet and selecting randomly located icons.

In Chapter 6, four contextual design version tests were investigated. The first test is the tracing test which is to evaluate and to measure a map route, free-hand input. The second test is the dragging test which evaluates clicking and dragging objects to specific locations. The third test is the one direction tapping test in which the evaluating point is the movement along one axis on a horizontal rubber banding, an insert cursor at points along a character string and selecting information in columns or rows. The fourth test is the multi-direction tapping test in which the evaluating point is the movements in many different directions in the repositioning of a pointer at different areas on the screen, cell selection in a spreadsheet and selecting randomly located icons.

Chapter 7 is the concluding chapter that gives a summary and discussion of the research findings.

Finally, in Chapter 7, the problems that occurred during the study are highlighted and the remarks and suggestions for further research are outlined.

CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

The literature review presented in this chapter addresses the interrelated areas of concern relevant to this study. In order to establish the list of literature to be reviewed, I used the following keywords to search the literature: TUI, ISO, Fitts Law, HCI and NKID. To include the disciplines of computer studies, I used the following databases: ACM, Google Scholar and JSTOR. I limited my findings to articles, books, book chapters, websites that provided information about one or more of the following:

1. Touchscreen Usability.
2. International Organization for Standardization Guideline
3. Fitts Law Theory
4. The principles and rules considered while testing the efficiency and effectiveness in terms of accuracy, errors, speed and user satisfaction of touchscreen technology.

2.2 TOUCH USER INTERFACE (TUI)

A touch-based user interface is an example of computer technology based on the sense of touch (haptics). Haptics is the technology which is getting more and more significant in the multimedia community as it furnishes a range of provisions where we can just touch the pages to load an internet webpage, activate the audio device or send documents (Golshani, 2007). Mainstream exposure to multi-touch technology occurred in 2007, when two products were launched - Microsoft Surface and the Apple iPhone, which started the era of touch technology that incorporated the multi-touch features. The multi-touch and gesture-based interfaces are frequently used, becoming the most desirable element of the market sector promoting mobile devices, laptops, desktop computers and large format displays. The touch technology was first used in the mid-1960s when the work at IBM started (Betts *et al.*, 1965) at the University of Illinois (Ebeling *et al.*, 1972) and in Ottawa, Canada (Hlady, 1969). In the 1970s, many new discoveries were made in the field.

In 1972, touch screens were more popular at selected grade school classrooms and computer centres as a part of the PLATO IV system. It became important when PLATO IV was integrated with the personal computer appearance and LAN networks, resulting in wide deployment when Xerox PARC began working on Alto computers. In the decade from the 70s to 80s, many different technologies evolved from touch, such as resistive, capacitive, surface acoustic wave and light interruption. Many different companies were created using these technologies for commercial purposes, such as Carroll Touch, MicroTouch Systems and Elographics (Buxton, 2010).

Dr. Buxton, in 1984, developed a multi- touch tablet, which could sense the degree of touch in multiple point systems (Lee *et al.*, 1985). Previously, the aim was to create a digital-hand drum, which is a musical instrument. It was the first multi-touch system reviewed and credited as one of the first multi-touch devices (Buxton, 2010).

A computer science expertise based on the sense of touch is recognized as a touch-based user interface. It not only allows computer-based tasks to be triggered by a sense of touch but also permits the user with ocular impairments an added rank of communication upon physical input. Although touch-based user interfaces have been in the region since the 1960s, they did not attain major awareness till 2007. In the previous five years it has become broadly recognized, and a group of mobile phone producers have commenced their own editions of touch-based user interfaces (Hansen, 2009).

2.2.1 Purpose and goals of a touch-based user interface

Touch screen gadgets are highly customizable, permitting designers to craft a wide range of user interfaces. They grant a natural and appealing understanding by permitting users to unswervingly touch and manoeuvre data.

- Should feel like an extension of the user’s body (Wigdor & Wixon, 2011).
- Offer a natural experience.
- Permit the users to operate a system.
- Permit the systems to show the outcomes of the user’s direction.

- Consider the size factor.
- System developers must make use of available interface design benchmarks and directives.

2.2.2 Evaluating touch techniques

Touchscreen interfaces incorporate both challenges and opportunities for researchers and momentous attempts have been applied to develop these interface.

Different techniques for touch-based user interfaces as stated by Nicolau *et al.* (2014, p. 4-5):

- **Tapping-it** is the most common and used technique in touch-based user interface technology and involves selecting an item by touching it. Different target sizes and screen positions are employed for this technique for the ease of users.
- **Crossing-in** crossing we do not choose the intended part as in tapping; instead an entry is chosen by just crossing it. This method offers improved performance for motor impaired users.
- **Directional gesturing-it** is the only technique that does not have a need of a target selection. In this technique the directional gestures can be made anywhere on the device’s surface.

2.2.3 Applications of touch-based user interfaces

Touchscreen interfaces have been widely accepted for many platforms and applications. They are used in kiosk displays, ATMs, home systems, mobile phones, etc. They are also very popular in older adults because of their easy to use nature (Jin *et al*., 2007).

Schöning *et al.* (2008) specified the advantages and disadvantages of user interface gadgets as below:

2.2.3.1 Advantages of touch-based user interface gadgets

- Easy and straightforward user interfaces.
- Have fewer buttons that might rupture, following a few months of function.

2.2.3.2 Disadvantages of touch-based user interface gadgets

- Many cell phones have quite small screens.
- Large screens lead to shorter battery life.
- Usually have a smaller amount of accuracy.
- Mostly the interfaces are not optimized for thumb function so a stylus is required.
- Entails substantial computing.

2.2.4 Touchscreen technologies

A touchscreen is basically an electronic image display, which is controlled through either simple or multiple touch gestures, just by touching the surface of the screen. Some of the touch screens are also capable of detecting different objects, for example a stylus or ordinary or coated gloves. Moreover, the user uses the touch screen as a response to the display as well as to control the display, such as resizing the font. The touchscreen is able to interact in a straight line with the display rather than using other devices, such as a mouse or touchpad. There are a number of touchscreen technologies that exist in the modern world (Bhalla & Bhalla, 2010).

2.2.4.1 Resistive

A resistive touchscreen comprises a number of layers, the most essential of which are a couple of thin, translucent, electrically-resistive layers separated by a thin space. These two layers face each other with just a very thin gap present between them. This touchscreen is mainly used in restaurants, factories as well as hospitals, because of its high resistance to liquids and contaminants. A major benefit of this technology is the lower cost. Subsequently, as only sufficient pressure is necessary for the signal to sense the touch, gloves, or something as flexible as a finger, can be used (Tode *et al.*, 2001).

2.2.4.2 Surface capacitance

In this fundamental technology, it is normal for only the single side of the insulator to be coated with a single conductive layer. The functionality involves provision of a very small voltage of the layer, which results in a uniformly distributed electrostatic field (Kim *et al.*, 2011). Upon getting connected with a conductor such as a finger, the capacitor forms dynamically and is controlled by a special controller, which determines the location of the finger and its movement.

2.2.4.3 Infrared sensors mounted

This technology involves infrared sensors, which are normally mounted on different displays and watched for user touchscreen input into the PLATO V terminal. The characteristic of the monochromatic plasma display containing infrared sensors and infrared technology used in touchscreen electronics contains an array of X/Y photo detector pairs incorporated with the LED infrared on the edge, which detect the disturbance in the infrared beam due to touching the screen. It is even sensitive to dust and dirt particles and used in the computer industry for highly sensitive panels as well as in medical equipment (McGookin *et al.*, 2008).

Apart from these, there are a number of other technologies which are being used for touchscreen devices. Some of them include acoustic pulse recognition; dispersive signal technology; infrared grid; infrared acrylic projection; mutual capacitance; surface capacitance; and surface acoustic wave. Input sources for these touch screens vary from fingers, nails, gorilla arms[1] and touch pens; some systems can be operated touching on

---

[1] "Gorilla arm" is a term engineers coined about 30 years ago to describe what happens when people try to use these interfaces for an extended period of time. It's the touchscreen equivalent of carpal-tunnel syndrome (Carmody, 2010). According to the New Hacker's Dictionary, "*the arm begins to feel sore, cramped and oversized - the operator looks like a gorilla while using the touchscreen and feels like one afterwards.*" (Gorilla Arm. 1996).

the screen with one's fingers while some of them require a touch pen as without it the user cannot interact with the system. However, it totally depends on the requirements of the system as to which source can be used. Cost of the touchscreen depends upon the detection sources used in it, and also varies from technology to technology (Clark & Harper, 2002).

Clark & Harper (2002) claim that, today, touchscreen technology is advancing from multipoint touchscreens, which facilitate the tracking of multiple finger movement on the screen, and therefore the architecture involved is becoming more complicated day by day. However, they stated that modern touchscreen technologies are reliable and compact, though the cost factor is also increasing and furthermore they are more user-friendly than earlier devices.

2.2.4.4 Alternative human-computer interaction

The design process of human-computer interaction is a goal-oriented problem-solving and decision-making activity in order to balance the requirements and compatibility of a product. The activity can be performed using target domain, cost, feasibility and material. The main aim is to design HCI in order to represent a development plan and set of successive elaborations and alternatives. As HCI is a vast discipline that provides ways to interact with the computers, it also deals with alternative solutions that refer to alternative human-computer interaction. *Alternative HCI deals with bad elements and determines the helpful sectors to communicate with the computing devices in an efficient manner* (Shneiderman, 2003).

Designing alternatives is important for recommending ideas to meet the specified requirements. Consideration has to be given to the conceptual design as well as the physical design. Conceptual design is about the product behaviour (what functions the product should perform); whereas physical design considers interactive details of the product including sounds, colours, menu design, images to use and icon design. However, despite the careful design planning and analysis of the system, some practical issues can arise. The customer wants the designer to resolve those issues with a variety of inspiring solutions and choices that work effectively. Keeping in mind the aspect of user satisfaction, the designer eagerly goes for a design process of alternative HCI that includes documentation, prototyping and usability engineering. Hence, a trial version of the product with detailed report helps the user to assess and evaluate the system (Holzinger & Miesenberger, 2009).

There is always room for improvement and so in this case, the HCI design alternatives enable the designer to modify his design by comparing it with the products and services which exist in the same domain. First, there is a need to list the current features of his own design then analyze the design of others and in the end, take a further step to improve his design with respect to usability and compatibility. The designer can efficiently remove the common features and add the latest ones and in this way, he can improve the basic layout of his system, which is supposed to be an alternative HCI. In addition, we can take the example of two websites following the same theme; by comparing both, the design of one website can be made better and will then be considered as an alternative. Another option is also there - the development of the website from the initial phase. Now it’s up to the designer to decide which path to adopt (Zhang *et al.*, 2005).

Reference to Zhang *et al.*, (2005) also reveals that, the designer should select an incorporative alternative for the developed system to slightly modify and manage the system. Alternative HCI is an integrated branch of HCI, which can deal not only with software models but also with socio-technical business models. It can be considered another remarkable step towards the development and revolution stage in the world of technology. No matter the system type; the only thing that matters is its compatibility and flexibility when interacting with the user, which only comes with the real practice of human-computer interaction designs and their appropriate alternatives.

2.2.4.5 Non-keyboard input device (NKID)

Input device is a piece of simple terminology that refers to all sorts of devices used for inserting data and instructions into a computing system. Input devices are available in all sizes and shapes, from inserting textual data using a keyboard to operating the system using a pointing device like a mouse. However, *the input devices used to give* *the instructions to the computer without the use of keyboard are known as non-keyboard input devices*. They allow operating the computational devices with advanced features and are equally useful for disabled people as well (Atkinson *et al.*, 2004).

As Atkinson *et al.* (2004) points out that with the increase in computer usage, the occurrence of work-related muscular disorders has relatively increased, especially in the upper limbs, including arms, neck, wrists and shoulders. They also added that, for computer users, numerous contributing factors for such muscular injuries have been determined, including awkward positions, seated work, static work, stress on connective tissue and bone, pressure on nerves and blood vessels, age, gender, inactivity, psychological stress and overuse.

Considering all these aspects, they concluded that the technology intellectuals have designed non-keyboard input devices to release the stress on muscles and bones. However, a large number of NKID have been designed for the well-being of humans, and some of them are described below:

Touchscreens: These allow the user to access the system by touching on the screen. The most common, real-life example of touchscreen evolution is in mobile phones and iPads, with which we interact in an effective manner (Atkinson *et al.*, 2004).

Game Controllers: These are used as a communication device in video games. Game controllers can be used according to different consoles. It can be a joystick, microphone or a guitar. An interesting fact about these game controllers is that they are used not only for video games but can also be used in weaponry control in order to diffuse explosive material with the help of joysticks (Atkinson *et al.*, 2004).

Biometric Devices: A biometric device is one of the most demanding non-keyboard input devices as it identifies a person seeking authorization to the computational device through voice, fingerprints, retina patterns and iris recognition. Biometric devices are used for recognition and verification. Fingerprint scanners, iris recognition systems, facial scanners, hand geometry scanners and voice recognition devices are some of the most considerable biometric devices, deployed in sensitive and confidential sectors like airports and banks (Atkinson *et al.*, 2004).

Eye-Tracking Devices: This device refers to a hands-free mouse, which moves the cursor on the screen with respect to the eye movement of the user. The user can easily control the system with the help of his eye movement. More importantly, it is a remarkable input device for disabled people (Wigdor & Wixon, 2011).

Headbands: With the use of a headband, a user can operate the system using his brain. It links the computer with the human brain, therefore allowing the user to control the mouse cursor and keyboard commands with the help of his brain in an efficient manner (Bolton, 2009).

Besides the advancements in NKID technology, there are various disadvantages of NKID such as laser beams and sensors, which are harmful to the nervous system; but still these devices have been brought into the world of technology. In addition, under proper guidance and practice of regulations, the user can make an effective use of NKID.

2.2.5 Usability

User response and reaction to technology provokes the improvements and enhancements in various features of the computing devices. User-friendly interfaces attract more users and hence increase the learn-ability and ease of use of human-made systems that can be achieved in an effective manner. For achieving this target, usability is an important discipline of human-computer interaction that can never be ignored. Usability refers to a degree of satisfaction and success that is ranked by an individual who uses the document, product or website. Usability can be defined further in terms of quality components such as ‘easy to learn’, ‘easy to remember’, ‘error free’, ‘efficient’ and ‘satisfying’ (Preece, 2000). Usability according to ISO standard (ISO 9241-11 (1998) is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

It is essential to know whether the specific product or service gives the required output or not. This can be done only if the user requirements are defined properly and hence, requirement specification is the most important phase when developing any system. The developers and designers should be aware of the user demands and needs so that they can create a user-centric system; user feedback serves as an important factor in this regard. If all the features of the system are pleasant and easy to use then it definitely increases the demand for the system or service, and hence increases productivity in this manner. People may consider it as a complete cycle that is interlinked; usability causes productivity and improved interaction of humans with the computational devices (Reiss, 2012).

Usability measures the usefulness and elegance of a computing system that differs from user experience and user satisfaction. This is the reason why efficiency of system requirements and user goals are matched by performing usability testing. It is a key that helps writers and designers catch issues in their websites and documents. They create a descriptive document including a complete plot of requirements and the respective output; users are requested to check the model of system so that essential modifications can be done before deploying the final version.

An increase in usability generally offers remarkable benefits such as increased sale and higher revenues, reduced support and development cost, increased productivity, reduced maintenance charges, and increased user satisfaction and efficiency. By improving the user demanded factors and creating ultimate levels of client satisfaction, organizations can accomplish their objectives of efficient output within their available budget resources (Lazar, 2007).

In a broader sense, usability does not only deal with the users and designers but it shows the overall performance of an organization. When the organizations promise the high quality services and products, it becomes their responsibility to show how they increase the efficiency rate, meeting the user requirements and expectations. For this purpose, most of the software development companies control the product quality through inspection and inquiry. Usability practitioners and industrial engineers perform such tasks to ensure effective human-computer interaction (Fuglerud & Rossvoll, 2010).

In particular, usability is considered to be one of the most important software quality traits, acquiring its position among traditional attributes like robustness and performance. Moreover, various academic programs regarding usability and emergence of usability consultancy organizations depict the real importance of usability in the field of computing and technology (Gregor *et al.*, 2002).

2.3 HUMAN-SYSTEM INTERACTION IN COLLABORATION WITH ISO

ISO is an abbreviation for the International Organization for Standardization and it is a highly reliable body for setting standards for services as well as for products. Different member bodies are included in this organization and they all work in accordance with their personal areas and fields of expertise. According to BS EN ISO 9241-420 (2011), usability of human-computer interaction can be defined as *"the extent to which a device can be used by the specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"*. This is an all-inclusive definition and defines how ISO visualizes the interaction between humans and their computers just as it is the definition for other products and services as well. Different bodies regulate the development of different aspects of this HCI and thus there are different ISOs in this regard. For instance, ISO 13407 (1999) focuses solely on the human-centred design processes that are operational in all such interactions.

A complete and thorough guideline has been given in the BS EN ISO 9241-420 (2011) for understanding the basics of standardization for computer interaction and the guidelines have been divided into annexes so that everyone can understand it. ISO aims at providing ease to practitioners and designers so that they can meet the expectations of the users. It is now clear that user experience is one of the most important factors that can help building the system as required or demanded; focusing on this aspect, ISO illustrates the fact that user experience is influenced by the context of use, system and user. It is known that input devices work in accordance with the input given by the user; in such a case that the design is not good enough then it creates a problem for both the producer and the consumer. This is why it is said that user experiences matter a lot in helping to build the design by considering the factors and issues which may resist the successful human interaction with the system in an effective manner. Although usability and user experiences present two different perspectives, at the same time both help achieving the specified goals and objectives with satisfaction, efficiency and effectiveness. Usability deals with getting a task done while user experience refers to the person's responses and perceptions that derive from the used product, service or application.

Putting together a huge variety of authoritative guidelines and prerequisites for evaluating, designing and developing the usable services, BS EN ISO 9241-420 (2011) imposes safety, compatibility and consistency. Highlighting the great contribution of ISO in the world of human system interaction, Bevan (2009) explains the purpose of these international standards: a standard is to make sure that a mobile phone will be able to accept any SIM card and can produce the transmissions which are attuned with all cell networks by not creating the dangerous radiations. This is called the implementation of defined standards while designing the cell phone; hence the usability that may be affected if it is quite complicated to use the cell phone. The users may get irritated with the complex design; so ensuring the usability depends upon the constraints on resources, design environment and context of use. The BS EN ISO 9241-420 (2011) standards have solved these issues in various ways in numerous areas.

2.3.1 Fitts Law Theory

There have been many theories and practices of human-computer interaction developed for studying human-centred performance. One famous theory is Fitts' Law. Fitts law is a representation commonly employed for human-computer dealings. It is utilized to model the performance of aiming by physically feeling an entity with a hand or finger or virtually directing to an object with the aid of a pointing device. It is a model of speed accuracy tradeoff. It describes generally of what ensues when user attempts to choose several information on the display.

Fitts law states that $T = a + b.\log\left(\frac{A}{W}\right) + 1$

- T is the time needed to carry out a choice.

- A is the distance from where the user begins to the objective to be chosen.
- W is the width of the objective.
- The coefficients "a" and "b" are empirically decided for some exacting grouping of input tool and display.

Fitts law effectively divides selection theory into two elements which can be measured independently. The expressions A and W are supposed to be unchanging for a given option and then "a" and "b" can be derived by changing the input devices (Olsen, 2009).

2.3.1.1 Assumptions of Fitts Law

- For input mechanisms such as screen and mouse, and b are disregarded which distinguish the input mechanisms and attention is made on which is consigned to as the index of difficulty.
- Any course from the begin location to the target location is agreeable and the user is liberated to optimize that course.
- Time to visually situate the goal is zero (Olsen, 2009).

2.3.1.2 Speed-Accuracy Tradeoff

- The width variable is considered as the accuracy element.
- He supposed that the motor structure has rigid information. If the cyclic progress of rigid amplitude is pacing up than the progress changeability will decrease.

It is principally a form of speed-accuracy exchanges. It predicts time involved to obtain records on display as a function of the distance of the aim to the objective and size of the aim. Fitt initially conceptualized the human motor structure as a communication path, progress amplitude as the indicator and aim width as the noise (Seow, 2005).

2.3.1.3 Advantages of Applying Fitts Law

- The model offers conventional power further than the task parameters.
- Fitts' law converts the experimental measurements to an index of performance that is sovereign of the particular task constraints (size and distance) (Accot & Zhai, 1999).

2.3.1.4 Summary of Fitts Law

- The time to spot an object with the help of a device by means of a device is a property of the distance from the marked object and the object's size.
- The more the time is necessary to situate and point an object, the object will be smaller and at a great distance.
- It is constructive for estimating structures for which the time to find an object is vital, e.g. a cell phone, a handheld device (Fitts, 1954).

2.4 REVIEWING RELATED STUDY

In a study, Bevan (2001) mentions the relationship between the international standards and usability; he discussed one of the failures of a graphical user interface standard titled drivability; it has been said that standards should define the elements of user interface where the designer and developer can judge the applicability of each of the guidelines. He also described that usability tests are important before finalizing the product, system or service. Usability must be incorporated for the system and software quality; the design should be human centered and this can be done by using the standards that explain the activities connected with the good usability and good user friendly interface design. Putting the emphasis on the importance of usability, user experiences and user friendly interface, the next section is designed; tracing test, dragging test and tapping test are the important aspects that should be considered in order to evaluate and assess the system, in this case tests have been discussed in context of touch screen technology. The functional properties of the system are tested to ensure that the design is not influenced when the user track the screen using his finger, when he tends to drag and drop the item or when he wants to tap the button. Hence, the main intention is to define the principles and rules that should be considered while testing the efficiency and effectiveness of touchscreen technology.

2.4.1 Tracing Test:

Looking at the ecosystem of touchscreen from a broader perspective, Padre (2014) presents a basic overview of the components of touchscreen; this study helps understanding of how the components of touchscreen work and allows the user to interact with the input device in an efficient and effective way. The hardware and software components are involved to make the device work; touch panel hardware, host touch driver, event management framework, graphics framework and display driver are the components combined together to form a touch based input device. The components of the touchscreen system work together to translate the user input to the graphical feedback; the touch responsiveness is based on the components of the touchscreen system and also on the input lag when using the touch system.

After explaining the components of a touchscreen system, Padre (2014) explains the important concepts regarding touch responsiveness. Touch responsiveness refers to the time the device takes after receiving the user input; the device acknowledges the users input and take some moments to respond to the user or the user waits for the feedback from the device; both conditions can be taken as touch responsiveness. In simple terms, it can be said that the time of translating the touch events from frame updates and the touch hardware on display is termed as touch responsiveness. The user presses the basic component of the touchscreen i.e. touch panel and waits for the device to respond; in the meanwhile, the system performs some internal functions and responds to the user while updating the user interface that would be displayed on the screen as the requested output. It’s the system latency which is experienced by the user to interact with the input device; the less is the system latency the more is the device performance and responsiveness. Move latency, initial move latency and tap latency are important to improve the touch experience; these latencies are measured during the taping and swiping gesture. Based on the above discussion, it is not wrong to state here that it is important to test the responsiveness of a touchscreen device; it is necessary to test the device from various perspectives and for this purpose tracing test has been used in order to check what exactly is needed. Testing will help building the features, eliminating the bugs and cleaning up and modifying the code in a better way. Tracing tests help the user and developer check the execution of the application; the user requests the input device for the particular action and the device then responds to the user; system latency and responsiveness are the other things but the main point is how fast the system gets the user input; it is important to know whether the system gets the command request without any delay or not; this is what one can know using different testing procedures including the tracing test.

Tracing or software tracing is the advanced form of logging that is used to note down the system progress in relation to the execution of a program (Davis, 1995). The information recorded during tracing is used for the debugging purposes depending upon the information detail included in the trace log. It should be noted that tracing can be done either by a technical support team or system administrators using the monitoring tools and policies in order to diagnose the common problems with the particular system. There are various forms of tracing, but the concept of tracing discussed here primarily deals with those having diagnostic and debugging purposes. Helander *et al.* (1997) in their study illustrated that tracing and dragging tests are not only limited to check the accuracy of trackballs and touchscreens but they are also used to trace the accuracy when using the mouse. The results have not shown much difference; tracing tests performed using stylus and fingers show by comparison that the task completion times when using the mouse are faster than those of stylus or fingers. The results were not differing much with reference to the errors but it has been said that touchscreens and mouse are the useful interaction techniques with fewer errors.

Illustrating a touch system in the simplest way, Padre (2014) carries out the relevant threads from the touchscreen system trace session. The researcher has described the system tracing test through an example; he has explained what will happen when the android framework gets the input from the user graphics framework and set to send to the display driver; the touchscreen controller scans sensor and then finalize the provided frame by display itself and the display driver. The effective tracing session contains a set of steps including the touch system first reads the input and sends off to the event management framework that takes 2-3ms; after receiving the signal this input event is dispatched to the destination in case of testing the application; after couple of milliseconds the event processing is started and the system responds the user accordingly.

Various studies explain the different methods used to perform the tracing test for input devices and the touchscreen systems. According to BS EN ISO 9241-420 (2011), moving the input device or pointer over the lines of an image or shape, one can trace the object on screen. For completion of this task, an application with freeform paths is required as tracing is based on free-hand input; though, freeform paths are not required if the input is operated in a magnetic environment or grids system. According to the given ISO study, the object should be traced both in a clockwise and anticlockwise direction on different devices; by calculating the standard deviations and means for each device, performance and time accuracy can be measured in an efficient manner.

Highlighting the need and importance of touch accuracy, Ezor (2010) presents a study in which he has explained the touch accuracy tests conducted by the MOTO and CNET development groups. The touchscreen tests allow the user to check the screens for 24-bit compatibility and numerous errors. These tests enable the user to check black level, contrast ratio and brightness displayed on the pattern screens by tooling around with photos and games in order to check the compatibility and reliability of the system. The development groups consider several different aspects of screen testing including the touch accuracy; the results may or may not be in accordance with the expectations of the user. The MOTO development group made a robot with counterfeit capacitive fingers, performing the tracing test with the same patterns on different phones that ended up with varying findings. The findings show under the lower light level the things fell apart when performing the test on Palm Pre and Blackberry Storm 2 whereas Apple iPhone and HTC Droid Evis show the better results. In this regard, it is concluded that the more finger contact there is, the better will be the tracking accuracy.

Figure 2.1: Tracing Tests on Different Devices
Source: Ezor, J. (2010).

Besides numerous testing techniques, different tracing applications have also been designed to check the accuracy of touchscreen systems. For instance, Touch Test is one of the applications used to perform the touchscreen test on the android Smartphone; it allows the user to trace the fingers on a grid to check whether the system supports multi-touch in an accurate way or not (Moin, 2010).

It also significant to highlight the study by Zabramski *et al*. (2013) on comparative evaluation of input devices in tracing tests. Participants in their study show faster movement time using touch for the reason that the tasks are learnable forwarding to the next level and participants could predict their next move which explains the fastest movement time, however leading to carelessness. It is likely that the task formulation forcing participants to perform “as fast and as accurate as possible” is responsible for creating different operational bias (Zhou & Ren, 2010).

2.4.2 Dragging Test:

It is no secret that various interaction techniques are designed for the users of touch and pen based display systems; thus, drag and pick or drag and pop are the techniques designed for the users to provide them with an access play around with the screen content that might be impossible otherwise. Drag and pop is an alternative title used for drag and drop which indicates that the user drags an icon and drops it onto the target icon; the process actually contains three simple steps, namely picking an object icon, dragging it towards the destination and dropping it onto the target icon. These techniques allow the users to interact with the icons displayed on screen, using comparatively small hand movements. Extending the drag and pop technique, drag and pick interaction style allows the activating icons to launch applications or open the folders.

Motti *et al.* (2013) in their study, discuss the importance of the dragging technique when it is about older adults using the touchscreens. It has been said that dragging is one of the most effective interaction techniques that enhances the user experiences in a positive way. The research shows that it is no doubt familiarity with the mouse which results in higher performance, but particularly for the experienced users; more recently, it has been reported that the performance gap can be reduced by touchscreen technology using various manipulation tasks like steering, crossing, dragging and clicking. Kobayashi *et al.* (2011) recommend that dragging is a much better interaction technique than tapping for the older adults, when it is about small touchscreen devices.

According to BS EN ISO 9241-420 (2011), one can evaluate the touch user interface by performing the dragging test. When dragging the test object to the target, three results can be observed: dropping the icon exactly at the destination point or near to it; dropping the icon a few millimetres away from the target or dropping the icon beyond the target icon. Based on these measurements, the accuracy of touchscreen system can be evaluated through the dragging test. Hitting the object in the centre of the target scores 3 points; the exact hit on the point scores 3 points; hitting few milliseconds away scores 2 points and hitting beyond the target outline scores 1 point. This test is conducted on different devices and thus the standard deviation and overall score is calculated for each device in order to measure the performance, time and accuracy of the touchscreen system.

Finding more about the touchscreen dragging and dropping technique, Astala *et al.* (2003) conducted a research in which he explained an apparatus and method for dragging and dropping the items exhibited on touchscreen. The dragging technique is practised in a way that an item on touch screen is touched slightly first and then with a greater pressure than the first attempt; with the same pressure the item is dragged to the targeted location; first, second locations and first, second predetermined pressures for a specific period of time should be considered to perform the dragging test in an efficient manner.

Confirming the findings, Collomb *et al.* (2005) state that drag-and-drop technique works well as long as the source and destination location is located in the same unit but it fails when there exist long distances or the source and target icons are in the different display units. It has been concluded that dragging technique is efficient in terms of efficiency, time and speed when in the same display unit.

2.4.3 One-direction Tapping Test:

ISO has defined the extent and standard following which a particular user can use the product in order to achieve the defined goals with efficiency, effectiveness and satisfaction; efficiency is achieved when the user tasks are completed in time, effectiveness deals with the tasks completion procedures by the users and satisfaction refers to the response by the user with respect to the experience. The usage and performance of touchscreen system is measured by the specified context of tasks, environments, equipment and users. As discussed earlier various tests have been performed to evaluate the ergonomics of touch based system as defined in BS EN ISO 9241-420 (2011); thus, one-direction tapping is one of the evaluation measures used to evaluate the pointing movement along either an x-axis or y-axis. It should be noted that there are two types of tapping; one-direction tapping and multi-directional tapping; but the emphasis in the given section is on one-direction tapping; multi-directional tapping will be explained in the next section.

Jota *et al.* (2014) in their research highlighted the aspects of advanced computing systems; in the given study they stated that the complex gestures and user interfaces can be made simple by abstracting a well-defined set of specified tasks while interacting with the contemporary computing systems. The study explains that according to ISO 9241-9 (2000), one-direction tapping test is one of the major tasks of computing, comprised of pointing and clicking. In short, tapping is done when an icon is targeted and clicked to perform any particular task; the selected content is first tapped and then dragged to the destined location. Hence, pointing and touching on the screen on only one axis is termed as one-direction tapping; for instance, with vertical or horizontal rubber banding, inserting cursor along the characters and selecting the information in columns or rows. Once a pointer object is selected, it is clicked and dragged to the particular location with the specified distance; the accuracy is measured in terms of time and throughput on different devices in order to evaluate which device gives the better results as expected or required.

Testing the pointing devices in accordance with the ISO 9241-9 (2000), Douglas *et al.* (1999) described that for the purpose of evaluation and comfort, the throughput was measured for both one and multi-directional selecting and pointing. The researchers conducted an experiment in which the participants rated the devices for comfort, usability, fatigue and operation. It has been concluded that results for the multi-directional testing was more significant than the results calculated for the one-direction tapping; the tests were performed using the joystick and throughout was measured accordingly for each type of tapping.

A testing company, OptoFidelity, performed an automated test to check the accuracy of touchscreens; different devices including Samsung Galaxy S3, iPhone 5c and iPhone 5s were selected to conduct the test. Artificial robot fingers were used to make several different taps across the screen of each of the phones. The researchers compared the tap locations where the tap was registered by touch panel. If the robot fingers actually hit the location where the tap was registered, the point was shown a green dot; more the numbers of green dots, greater is the tapping efficiency of the touchscreen system. The results show that Galaxy S3 performs very well in the test; losing the efficiency and accuracy at the edge of display the iPhone shows a huge amount of inefficiency and inaccuracy with 75% of inaccurate results. Comparatively, iPhone 5c is better than iPhone 5s and Samsung Galaxy S3 performs best in this tapping test conducted to check the accuracy of touchscreens (Arnott, 2013).

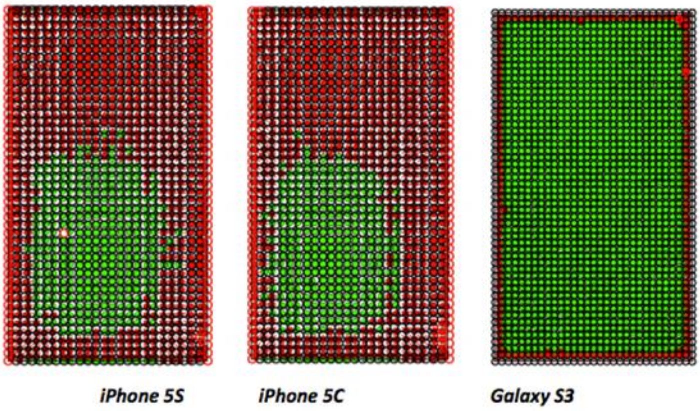


Figure 2.2: One-direction Tapping Test
Source: Arnott, N. (2013).

2.4.4 Multi-directional Tapping Test:

With the enhancements and advancements in technology, the companies emphasize that offering the friendly user interface to older adults improves their access to streams of internet services by using the touchscreen systems including tablet computers. Attempting to acquire the approval for the recently developed touch interface, Burkhard and Koch (2012) appraised the touch input performance and size of the multi-touch tablets using the multi-directional task as recommended by ISO; for the evaluation, the authors considered the multi-directional tapping assignment as a conventional application for the touchscreen systems including Smartphones, Android and tablet computers.

Unlike one-direction tapping task, multi-directional tapping task is about assessing the pointer movements in various directions; for instance, selecting a cell in spreadsheet, making a selection of random icons located at different locations and relocating the pointing movements at several different locales on the device screen are some of the common examples used to explain multi-directional tapping. Using the multi-directional tapping testing technique, each device is measured on the basis of pointing movements made in different directions along both axes. Multi-directional tapping tests can be conducted in various ways; one of the ways is under discussion. Different objects are placed in a circle on the display screen; some objects are at a distance and some of the objects are located closer comparatively; selecting, clicking and dragging the icons from one target object to the other is multi-directional tapping. The evaluation is made on the basis of distance of the source icons and target objects; in addition, the assessment is made in accordance with the speed, time and accuracy measured during the given input and the received output (ISO 9241-420, 2011).

Comparing one-direction and multi-directional tapping tasks, Horsley *et al.* (2014) explain the concept with reference to the eye tracking system. According to their research, for one-direction tapping test, the participant is required to move from source area to destination target and then back to the source whereas the multi-directional tapping test requires 24 boxes around the edge of the circle and the participant is expected to progress from the centre of the circle to the target area. The participant then is supposed to shift to and click in the box opposite to the selected box and then continue the process in a clockwise direction; in this way, the participant is required to click all the objects, placed around the circumference of the circle, and then back to the first chosen box. The given study shows that in order to test the eye tracking system as the effective input device, Zhang and MacKenzie (2007) used multi-directional tapping tests; four conditions were made with varying time and look and respond methods; the participants are required to tap the spacebar in order to make the target active and the cursor on the screen moved in all directions with respect to the eye movement. This method was considered as one of the amazing eye tracking techniques with the best throughout when compared to mouse movements.

2.5 SUMMARY

There is no doubt that ergonomics in the context of human system interaction is a broader term and presents quite a challenging task to encapsulate in a few words; reviewing different studies which were conducted to consider the touchscreen systems, a base and efficient user interface is a core requirement; the given research results in an effective understanding and need of fast and accurate graphical user touch interface. In order to check the efficiency, effectiveness and satisfaction related to a system, different tests have been performed such as tracing tests, dragging tests, one-direction tapping tests and multi-directional tapping tests. These testing tasks help the user evaluate the performance of the input device that ultimately assists the user selecting the input device for interacting with the system in an efficient and effective manner.

Nevertheless, review of the literature reveals that the research in the field of user satisfaction and accuracy using the touch screen systems is minimal at best. Primarily, the research available does not compare three input devices (i.e. mouse, stylus and touch) and does not consider user satisfaction or preference across the broader field. In current research, there is a lack of any definitive user satisfaction consideration of the comparisons. Further study of user compatibility comparing satisfaction levels of various devices has yet to be considered as an entirety.

CHAPTER 3: METHODOLOGY

3.1 INTRODUCTION

Chapter 3 describes the methodology that underpins this investigation. It begins by highlighting necessary procedures that should be considered by looking into Experiment Design Guideline. This is followed by a discussion of experiment measures of data collection used, and the analytical procedure employed.

3.2 EXPERIMENT DESIGN GUIDELINE

Planning an experiment properly is very important in order to ensure that the right type of data and a sufficient user size and power are available to answer the research questions of interest as clearly and efficiently as possible. Hence, necessary procedures should be undertaken by looking into research texts on the design guideline.

3.2.1 Usability requirements

Maguire (2001b) specified that according to the ISO 13407 (1997) standard on human-centred design there are five essential processes which should be undertaken in order to incorporate usability requirements into the software development process. The processes are carried out in an iterative fashion as depicted in *Figure 3.1,* with the cycle being repeated until the particular usability objectives have been attained.

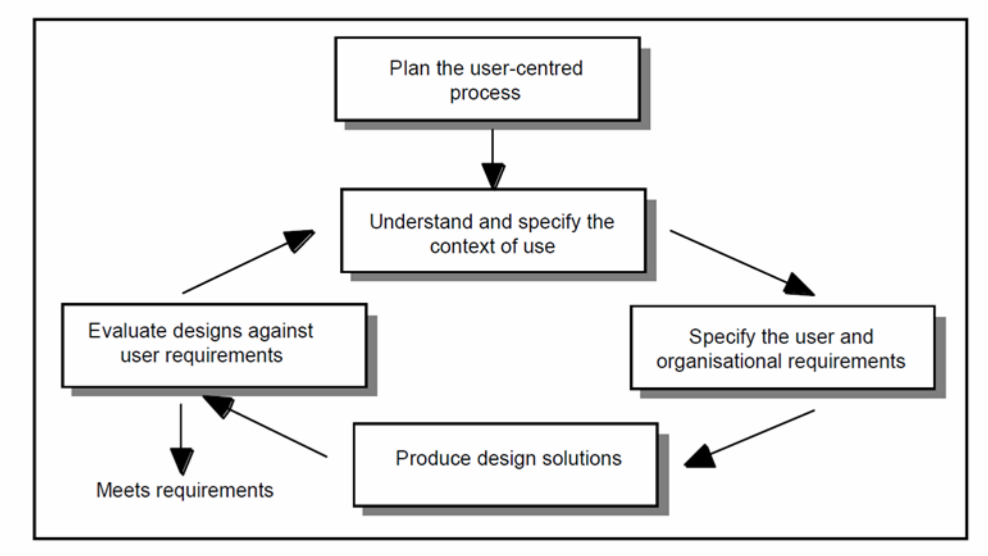


Figure 3.1: The human-centred design cycle
Source: Bevan & Curson (1999)

3.2.2 Context-of-use factors

The context of use becomes one of the main stages within the user-centred design process. The main elements of a context analysis are shown in *Table 3.1*. The international standards community has also recognised the role of 'Context of use' within usability. The ISO 9241 standard Part 11 - Guidance on Usability (ISO, 1997) refers to the context of measurement in its definition of usability: "Usability is the extent to which a product can be used with efficiency and satisfaction by specific users to achieve specific goals in specific environments."

This definition emphasises that the usability of a product is affected not only by the features of the product itself, but also by the specific circumstances in which a product is used (Maguire, 2001a). As defined by the standard (ISO, 1997): *"The Context of Use consists of the users, tasks and equipment (hardware, software and materials), and the physical and social environments in which a product is used."*

Once the context of use has been characterised, with the participation of key stakeholders, the next steps are to identify which characteristics of users, tasks and environment may be relevant to usability, and then to select representative users, tasks and environment for the evaluation (Macleod *et al.*, 1997)

User group	Tasks	Technical environment	Physical environment	Organizational environment
System skills and experience Task knowledge Training Qualifications Language skills Age and gender Physical and cognitive capabilities Attitudes and motivations	Task list Goal Output Steps Frequency Importance Duration Dependencies	Hardware Software Network Reference materials Other equipment	Auditory environment Thermal environment Visual environment Vibration Space and furniture User posture Health hazards Protective clothing and equipment	Work practices Assistance Interruptions Management and communications Structure Computer use policy Organizational aims Industrial relations Job characteristics

Table 3.1: Context-of-use factors
Source: Maguire (2001b)

3.2.2.1 Methods for evaluating usability

According to Macleod and Rengger (1993), methods giving reasonably rich data about usability fall into three principal categories:

1. Expert methods, based on expert judgement about a system or design (and hence dependent on available expertise).
2. Theoretical methods, based on models of the user and system, and how they interact.
3. User-based methods, where usability data is gained as a result of people using systems or prototypes. User-based methods divide broadly into survey methods, which give a picture of users' subjective views, and observational methods.

In any user-based evaluation it is required to ensure that the circumstances in which a prototype or system is evaluated match as accurately as possible the (intended) circumstances of system use. This includes characteristics of the users, the tasks they perform, and the organisational, physical and technical environments (Macleod and Rengger, 1993).

3.2.2.2 Choosing usability measures

Bevan and Macleod (1994) stated that a description of the quality of use should consist of appropriate measures of user performance (effectiveness and efficiency), and of user satisfaction. They defined usability measures as follows:

1. Effectiveness: measures of effectiveness relate the goals or sub-goals of using the system to the accuracy and completeness with which these goals can be achieved.
2. Efficiency: measures of efficiency relate the level of effectiveness achieved to the expenditure of resources. The resources may be mental or physical effort, which can be used to give measures of human efficiency, or time, which can be used to give a measure of temporal efficiency, or financial cost, which in turn can be used to give a measure of economic efficiency.
3. Satisfaction: measures of satisfaction describe the perceived usability of the overall system by its users and the acceptability of the system to the people who use it and to other people affected by its use. Measures of satisfaction may relate to specific aspects of the system or may be measures of satisfaction with the overall system.

Bevan and Macleod (1994) specified that it is normally necessary to provide at least one measure for each of the components of quality of use, and that it will often be necessary to repeat measures in several different contexts.

3.2.2.3 Recording and analyzing data

As stated by Macleod and Rengger (1993), observational evaluation can be carried out most conveniently - from the evaluator's viewpoint - in a usability laboratory, where users can be provided with a simulation of the workplace, and sheltered from the observers by a two-way mirror (one way window). They added that alternatively, video and audio data and observational notes can be collected in the actual workplace, and the subsequent analysis conducted in a usability laboratory, or a suitably equipped office.

3.2.2.4 Prototyping and Software Development

Prototypes are ‘instruments’ used within the software development process. Different kinds of prototypes are employed to achieve different goals. Naumann and Jenkins (1982) consider an information systems prototype to be: *“... a system that captures the essential features of a later system, is the most appropriate definition of a prototype. A prototype system, intentionally incomplete, is to be modified, supplemented, or supplanted”*

3.2.2.5 Prototyping Process Models

Floyd (1984) describes the prototyping process as consisting of functional selection, construction, evaluation and further use. Those functions that are to be prototyped are selected and a prototype is constructed. This prototype is evaluated and the prototype is further used for outlining specification or as a part of the new system.

Naumann and Jenkins (1982) characterise prototyping as a four step, iterative procedure involving users and developers:

1. User’s basic needs are identified;
2. A working prototype is developed;
3. The working prototype is then implemented and used;
4. The prototyping system is revised and enhanced.

This process undergoes several iterations and steps three and four are repeated until the user accepts the system (Carr and Verner, 1997).

3.3 EXPERIMENT PROCEDURES

There are three stages of experiment were carried out. The first stage was a pilot experiment that consisted of 4 design version prototypes. The second stage of prototypes that comprised abstract task type, and final stage was contextualised task type. The second and third stages experiments were based on the guidance of The International Organization for Standardization known as ISO (ISO 9241-420, 2011).

3.3.1 Pilot Study

The first prototype consists of dragging a black box to a red box with auto-merge. That is, even if the black box only touches a bit of the red box, it will automatically merge (flick).

The second prototype is designed to fix the first prototype design. Whenever the users dragged a black box to a red box there was an auto merge if the black box touched 80% of the red box.

To revise the first and second prototype design, the third prototype gives the user a new task where they must merge the black box and red box with 100% accuracy.

A follow-up fourth prototype design complemented the third one by requiring the user to merge the two objects by following a line/path.

3.3.2 Abstract Experiment

The second stage experiment consist of four tests (Tracing test, Dragging test, One direction test and Multi directional test) which are deliberately developed as abstract tasks with the purpose of analysing the user's ability on simple tasks without a real world context. Overall, the aim was to discover if there are differences in mouse, stylus and touch on the tracing test and dragging test with different levels of difficulty that could affect users' performance and satisfaction.

3.3.3 Context Experiment

The third stage consist of four tests as well (Tracing test, Dragging test, One direction test and Multi directional test) which are deliberately developed as contextual tasks with the purpose of analysing the user's ability in a real world context. Overall, the aim was to discover if there are differences in mouse, stylus and touch on the tracing test and dragging test with different levels of difficulty that could affect users' performance and satisfaction.

3.3.4 Apparatus

For the pilot study, a multi-touch PC tablet (Acer Iconia Tab W500) with a 10.1 inch diagonal LED display and a resolution of 1280 x 800 pixels, equipped with capacitive stylus and finger sensitive display, as well as a Wireless Optical PC five button mouse, manufactured by Acer was used.

A series of small software prototypes involving touch-based interaction were developed and designed using Adobe Flash. For instance, 2 square boxes – a primary box and target box (tap plus drag the primary box and place onto the target box then release) - and multiple sizes of boxes: small, medium and large (placed in random section).

Meanwhile for the abstract and context experiments, a convertible Samsung computer tablet with keyboard (Samsung ATIV Smart PC XE700T1C Tablet with Keyboard) Tablet PC with an 11.6 inch diagonal LED display and a resolution of pixels 1920 x 1080 pixel Full HD Resolution, equipped with S-Pen[2] stylus and finger sensitive display (Full HD Touch Screen), as well as a Wireless Optical PC five button mouse with 1000 dpi, manufactured by Samsung together with mousepad was used.

There were eight prototypes developed and designed using the Flash-based: Abstract – Tracing, Dragging, One Direction, Multi Direction, and Contextual - Tracing, Dragging, One Direction, Multi Direction.

The PC was used in "laptop mode" while using all the three devices input. Moreover, all three input methods used their default settings and their standard Windows 7/8

---

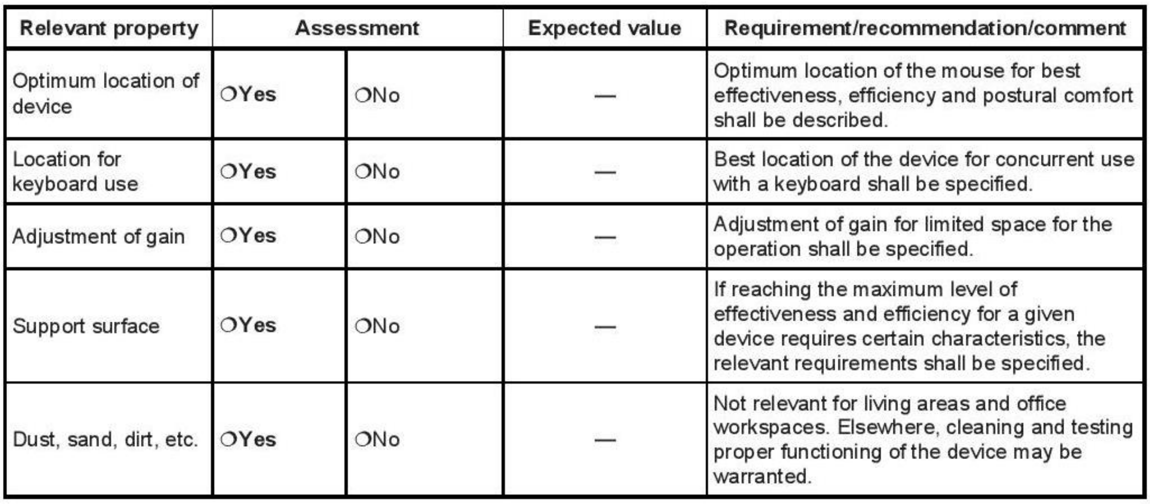
[2] The tip of the Samsung's S-Pen simulates a finger touch, tap or swipe to provide a full touch-screen environment without leaving a single fingerprint behind. With its attractive design, the S-Pen allows the user to make quick notes, draw or edit in addition to performing typical touch-screen commands. The user can also use the Replacement Stylus' multifunction button in conjunction with shortcut gestures for faster input.

system cursors visible while interacting. Flash-based applications were developed for all experiment prototypes. All movement time and error rate data were recorded during the interaction in every task.

In order to investigate the efficiency and effectiveness in terms of accuracy, errors, speed and user satisfaction following requirements provided by ISO 9241-420 (2011) is initially considered.

3.3.4.1 *Mouse*

In using mouse in the experiments, recommendation of its used is referred as stated in *Table 3.2.*

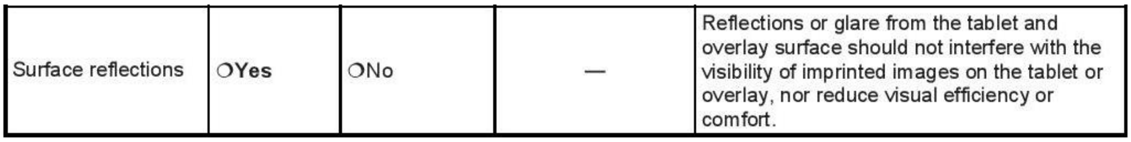


Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Optimum location of device	○Yes	○No	—	Optimum location of the mouse for best effectiveness, efficiency and postural comfort shall be described.
Location for keyboard use	○Yes	○No	—	Best location of the device for concurrent use with a keyboard shall be specified.
Adjustment of gain	○Yes	○No	—	Adjustment of gain for limited space for the operation shall be specified.
Support surface	○Yes	○No	—	If reaching the maximum level of effectiveness and efficiency for a given device requires certain characteristics, the relevant requirements shall be specified.
Dust, sand, dirt, etc.	○Yes	○No	—	Not relevant for living areas and office workspaces. Elsewhere, cleaning and testing proper functioning of the device may be warranted.

Table 3.2: Interdependencies and documentation of mice - Documentation
Source: BS EN ISO 9241-420 (2011)

3.3.4.2 *Stylus*

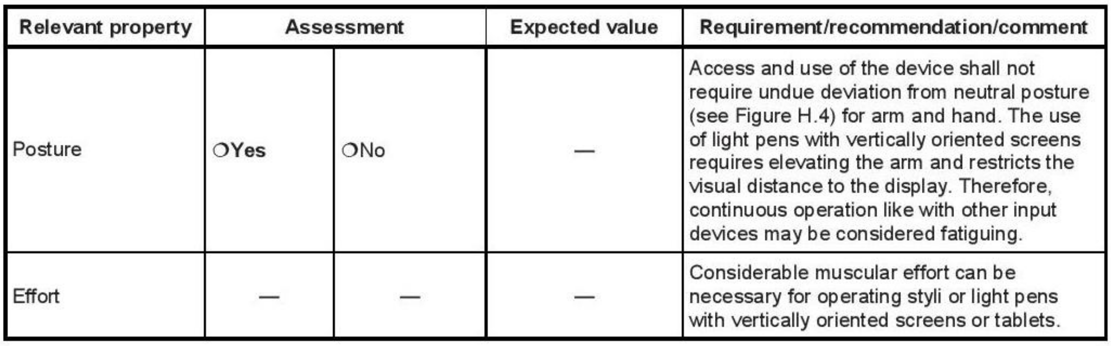
The ISO Standard has outmoded description of how a stylus is used (typically Wacom style with stylus and pad), although it does reference light pen and tablets/overlays. Light pen attributes validly apply in some cases. Tablet requirements are met in this experiment, such as the surface reflection as recommended by ISO standard.



Surface reflections	○Yes	○No	—	Reflections or glare from the tablet and overlay surface should not interfere with the visibility of imprinted images on the tablet or overlay, nor reduce visual efficiency or comfort.

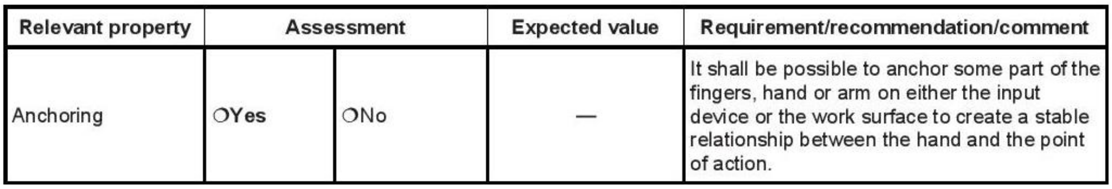
Table 3.3: Other properties of tablets/overlays - Legibility and visibility of legends and graphical symbols
Source: BS EN ISO 9241-420 (2011)

Biomechanical load and Functional properties is also considered in this study as recommended by ISO standard (*Table 3.4 & Table 3.5*)



Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Posture	○Yes	○No	—	Access and use of the device shall not require undue deviation from neutral posture (see Figure H.4) for arm and hand. The use of light pens with vertically oriented screens requires elevating the arm and restricts the visual distance to the display. Therefore, continuous operation like with other input devices may be considered fatiguing.
Effort	—	—	—	Considerable muscular effort can be necessary for operating styli or light pens with vertically oriented screens or tablets.

Table 3.4: Correspondence with generic requirements on styli and light pens - Biomechanical load
Source: BS EN ISO 9241-420 (2011)



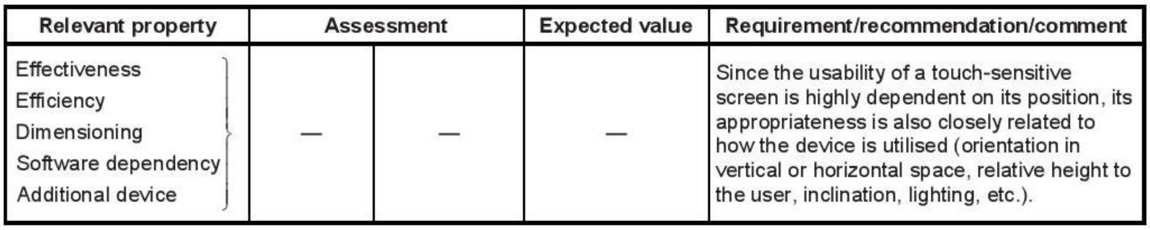
Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Anchoring	○Yes	○No	—	It shall be possible to anchor some part of the fingers, hand or arm on either the input device or the work surface to create a stable relationship between the hand and the point of action.

Table 3.5: Functional properties of styli and light pens - Functional properties
Source: BS EN ISO 9241-420 (2011)

3.3.4.1 Touch

As been said previously the PC was used in “laptop mode” while using all the three devices input. Therefore, touch experiment used 60-degree laptop mode for touchscreen and style for the reason that it aimed to investigate subjects with the same style degree to avoid bias result. Natural/typical uses of tablets are trickier because subjects are likely to hold them so many different ways. The tablet alone measures 0.5 by 11.6 by

7.2 inches (HWD) and weighs a reasonable 1.9 pounds. According to Clark (2012), subjects tend to grab, tilt, lean, cradle, and clench in a whole variety of embraces, many of which depend upon stance. Moreover, in this experiment subjects will be sitting at a table, where they are likely to prop a tablet with one hand at the lower third and tap with the other. Therefore, taking appropriateness of the experiment held in a lab into account, requirement by ISO standard is considered.



Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Effectiveness Efficiency Dimensioning Software dependency Additional device	—	—	—	Since the usability of a touch-sensitive screen is highly dependent on its position, its appropriateness is also closely related to how the device is utilised (orientation in vertical or horizontal space, relative height to the user, inclination, lighting, etc.).

Table 3.6: Correspondence with generic requirements on touch-sensitive screens - Appropriateness
Source: BS EN ISO 9241-420 (2011)

In this experiment, the touch screen is put on the docking keyboard so that subjects in the experiments can go freeform any fingers or hands, jabbing any fingers or hands at the screen. According to Clark (2012), users frequently adopt a bottom-corner grip, resting their arms alongside the keyboard and this can avoid gorilla arms. This showed that in this experiment, controllability and biomechanical load is adhered to as recommended by ISO 9241-420 (2011) standard (*see Table 3.7* & *Table 3.8*).

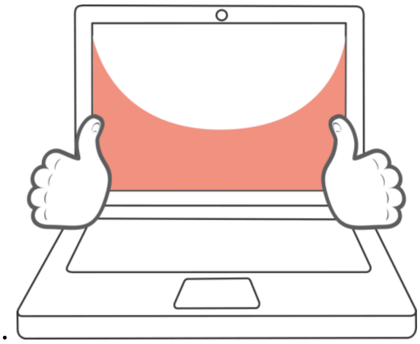
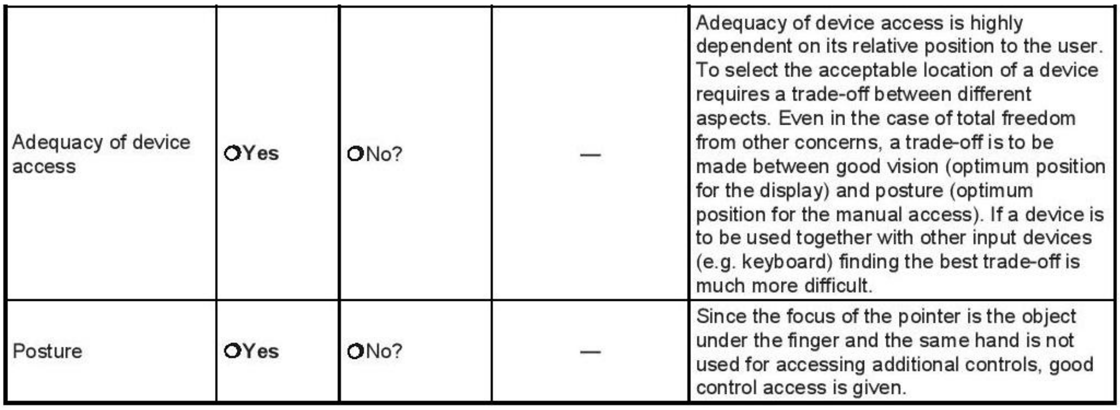
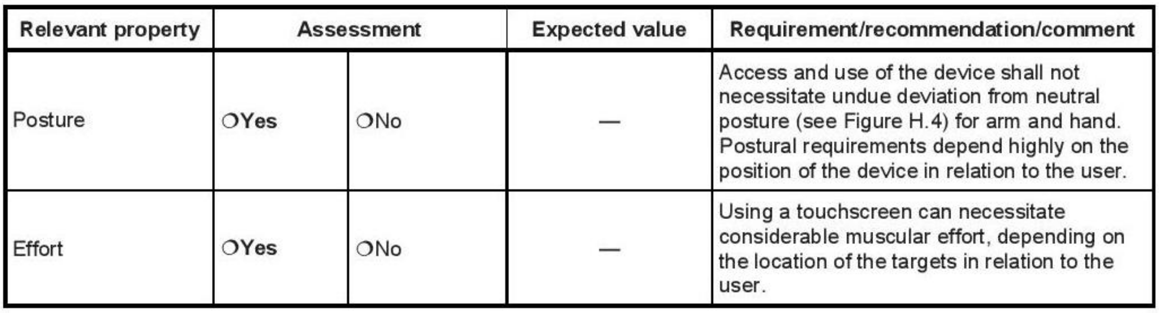


Figure 3.2: Bottom-corner grip, resting their arms alongside the keyboard
Source: Clark (2012)



Adequacy of device access	○Yes	○No?	—	Adequacy of device access is highly dependent on its relative position to the user. To select the acceptable location of a device requires a trade-off between different aspects. Even in the case of total freedom from other concerns, a trade-off is to be made between good vision (optimum position for the display) and posture (optimum position for the manual access). If a device is to be used together with other input devices (e.g. keyboard) finding the best trade-off is much more difficult.
Posture	○Yes	○No?	—	Since the focus of the pointer is the object under the finger and the same hand is not used for accessing additional controls, good control access is given.

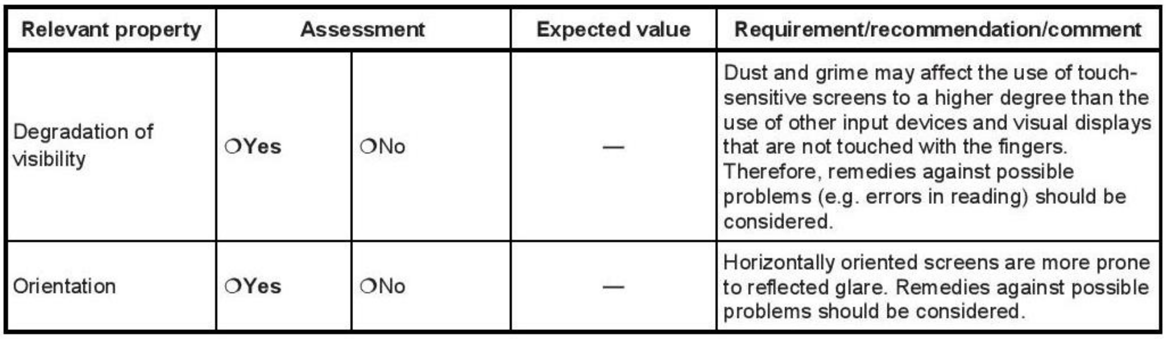
Table 3.7: Correspondence with generic requirements on touch-sensitive screens - Controllability
Source: BS EN ISO 9241-420 (2011)



Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Posture	○Yes	○No	—	Access and use of the device shall not necessitate undue deviation from neutral posture (see Figure H.4) for arm and hand. Postural requirements depend highly on the position of the device in relation to the user.
Effort	○Yes	○No	—	Using a touchscreen can necessitate considerable muscular effort, depending on the location of the targets in relation to the user.

Table 3.8: Correspondence with generic requirements on touch-sensitive screens - Biomechanical load
Source: BS EN ISO 9241-420 (2011)

The interdependency with use environment is also considered as recommended by ISO 9241-420 (2011), for example the degradation of visibility and orientation (*see Table 3.9*). The justification placing the laptop 60-degree laptop mode is intended for the visual attention naturally focuses on the whole thing at a glance and also taking into account the lighting at the lab which may prone to produce reflection.



Relevant property	Assessment		Expected value	Requirement/recommendation/comment
Degradation of visibility	○Yes	○No	—	Dust and grime may affect the use of touch-sensitive screens to a higher degree than the use of other input devices and visual displays that are not touched with the fingers. Therefore, remedies against possible problems (e.g. errors in reading) should be considered.
Orientation	○Yes	○No	—	Horizontally oriented screens are more prone to reflected glare. Remedies against possible problems should be considered.

Table 3.9: Other properties of touch-sensitive screens - Interdependency with use environment
Source: BS EN ISO 9241-420 (2011)

3.3.5 Users

Earlier literature had asserted that an appropriate number depends on the size project, with 7 users being optimal in as small project and 15 users being optimal in a medium-to-large project (Nielsen and Landauer, 1993). On the contrary, Dix *et al*. (2004) viewed that the user size must be large enough to be considered representative of the population and, if the intention is to run a controlled experiment and perform statistical analysis on the results, at least twice the number is recommended.

Since the research was a medium-to-large project that was run in controlled experiment conditions and implemented statistical analysis on the results, 20 to 30 users have been identified as participants. When selecting participants, people were mainly recruited from a higher education student population who have related background areas because expertise is always an important consideration (Lazar *et al.*, 2010).

The selected samples for the experiment represent the target population of this study as it aim high education student population who have related background areas in terms of computer experience. Intended user population is also referred to throughout the ISO standard. In page 12 of ISO 9241-420 (2011) stated that *"Only few technical products are likely to be usable for all human beings to the same extent and function in the same manner under all circumstances. In general, a product is designed to fully satisfy the needs of a certain user population, the target user or the intended user population"*

Therefore, the target user must be able to use the prototypes in this experiment. This is significant for the reason that this study intent to generalize from the sample to target population. This study should be as representative as possible of the target population. The more representative the sample, the more confident the researcher, that the findings can be generalized to the target population.

3.3.6 Measurement

The data was measured using quantitative measurements. The three most common quantitative measurements are Movement Time (*MT*) Error Rate (*ER*) and User Satisfaction using the five-point 'Likert-type' scale (Likert, 1932) questionnaire.

As been stated in literature review, usability according to ISO standard (ISO 9241-11 (1998) is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. The indicator of effectiveness, efficiency and satisfaction as indicated by ISO standard (ISO 9241-11 (1998) are:

1. Effectiveness Indicators: quality of solution and error rates.
2. Efficiency Indicators: task completion time and learning time.
3. Satisfaction Indicator: Preference that can be measured by attitude rating scales.

Hence, in this study:

1. Effectiveness is measured as Error Rates (*ER*) which is extracted from the interaction log or the log sheets.
2. Efficiency is measured as movement time (*MT*), which is also extracted from the interaction log or the log sheets.
3. Satisfaction is measured as preference of the user expressed by a grade on a five-point 'Likert-type' scale.

3.3.7 Statistical Analysis

Numerical quantitative data was analysed using the Statistical Package for Social Sciences (SPSS) to report what had been found. It is used to describe and present data.

According to ISO (ISO 9241-420, 2011), *"Given that the proper underlying assumptions are met, standard analysis of variance statistical techniques can be used to analyse this data. However, in instances where the necessary assumptions are not met (i.e. with small sample sizes or non-normal distributions) non parametric techniques of hypothesis testing should be used and tend to be computationally less complex"*.

Therefore, in this study there are some assumptions that need to be considered before a test is conducted. Since the participants in this study are to test the same individuals for differences in the participants' Movement Time *(MT),* Error Rate *(ER)* and User Satisfaction using Likert scales, between three input devices (stylus, touch and mouse), hence the appropriate test to be applied is ANOVA with repeated measures and also referred to as a within-subject ANOVA. Lund & Lund (2013) specify the five assumptions that need to be considered before applying this test:

> Assumption #1: One dependent variable that is measured at the continuous level (i.e., it is measured at the interval or ratio level).
> Assumption #2: One within-subjects factor that consists of three or more categorical levels.
> Assumption #3 No significant outliers in any level of the within-subjects factor
> Assumption #4 Dependent variable should be approximately normally distributed for each level of the within-subjects factor
> Assumption #5 Known as sphericity, the variances of the differences between all combinations of levels of the within-subjects factor must be equal

On the other hand, a Friedman test is used as a nonparametric alternative to a one-way repeated measure ANOVA. Therefore, this test is used if the data "passes" the following four assumptions (Lund & Lund, 2013):

Assumption #1: One group that is measured on three different occasions.
Assumption #2: Group is a random sample from the population.
Assumption #3: The dependent variable is measured at the ordinal level. For example of ordinal variables is Likert scales
Assumption #4: Samples is not normally distributed.

If one-way repeated measures ANOVA test is statistically significant (p < .05), a post hoc test investigate further to determine where the differences between levels of the within-subjects factor lie. If one-way repeated measures ANOVA is not statistically significant (p > .05), this result will not be followed up with any post hoc test or contrast, but just report the result of the one-way repeated measures ANOVA.

If the Friedman test is statistically significant (*p* < .05), post hoc tests is investigated further to determine where exactly the differences between groups lie. If your Friedman test is not statistically significant (*p* > .05), this result will not be followed up with any post hoc tests, but just report the result of the Friedman test.

3.3.8 Ethical considerations

Two main areas of ethical considerations are ‘informed consent’ and ‘confidentiality’. Each volunteer received an explanation about the study being carried out and read and signed a consent form (*see appendix A4*). Participants were allowed to ask questions at any time and could withdraw from the study at any time, should they wish to do so. In the case of a participant withdrawing their consent, the data collected from them would not be used. All data collected would be anonymised and used in this manner.

CHAPTER 4: PILOT STUDY

4.1 INTRODUCTION

Chapter 3 presents a pilot experiment that consists of 4 design version prototypes. In this pilot experiment, it was treated very seriously and conducted in exactly the same way as was planned for the actual experiment. As pointed out by Preece *et al.* (1994), a pilot study is the only chance to fix mistakes before running the main study. Whatever flaws or problems are discovered during the pilot study allow modifications to be made before the main study begins (Lazar *et al.*, 2010). The users are also encouraged to suggest alternative solutions and to criticize solutions they found unsatisfactory. In addition, some users' behaviour is also being observed in terms of expressions of emotion, without commenting, in order to indicate if the task is designed poorly. It may arise that in some situations users might think they are using a control correctly, but in fact they are not. This would entail further refinement of the task.

4 design version prototypes will be investigated in this chapter,. The main task of the prototype is to enable the user to merge 2 boxes (black boxes and red boxes) by dragging and dropping. Namely ***prototype version 1*****;** flick (difficulty level: easy), ***prototype version 2*****;** auto merges if both touch 80% of each other (difficulty level: medium), ***prototype version 3*****;** auto merges if both touch 100% of each other which means they need 100% merge accuracy (difficulty level: hard) and ***prototype version 4*****;** auto merges if both touch 100% of each other which means they both need 100% merge accuracy, in which the black box must be moved along a path (difficulty level: hard). In this experiment the prototype is deliberately developed as abstract tasks on the purpose of analysing the user's ability on simple tasks without a real world context. Overall, the above 4 prototypes are tested with 3 different input types (mouse, stylus and touch) in its aim to discover if these differences in three input devices and the box sizes with different levels of difficulty affected user performance and satisfaction.

4.2 HYPOTHESIS

Several hypotheses were devised around the area of efficiency of use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

H1: There is a difference in the amount of speed using touch, mouse or stylus.
H2: There is a difference in the amount of error using touch, mouse or stylus.
H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus interaction.

4.3 USERS

The initial recruitment of the participants took place by means of a recruitment questionnaire. Since the experiment involved testing aspects of using touch-based technology, it was deemed important to have participants with a certain amount of experience in using touchscreen devices and computers in general. This is because if there happened to be a number of beginners to such activities, the times and outcomes could potentially be compromised. Therefore:

- Twenty-two students in the Faculty of Education, Seri Begawan Religious Teachers University College (KUPUSB) took part in the experiment.
- Participants consisted of 14 male and 8 females.
- All participants were in the 18-39 age range.
- All participants had computer experience.
- All participants had experience with a touch-based device.
- All the participants' uncorrected visual problems or physical limitations that would inhibit their use of the mouse, stylus and touch as an input device were accounted for.

4.4 EXPERIMENT DESIGN

A within users design was used. The 22 participants were assigned to do all 4 version design prototypes, making 228 tasks altogether.

4.5 VARIABLES

4.5.1 Independent Variables

The independent variables were:

1. Prototype Version 1(Flick) – Difficulty level : Easy
2. Prototype Version 2 (Auto merges if both touch 80% of each other) – Difficulty level : Medium
3. Prototype Version 3 (Auto merges if both touch 100% of each other with 100% accuracy) – Difficulty level: Hard
4. The Prototype Version 4 (Auto merges if both touch 100% each other 100% accuracy – the black box must be moved along a path) – Difficulty level: Hard

4.5.2 Dependent variables

The dependent variables consisted of the objective human performance and the subjects' feelings about the device design.

4.5.3 Dependent measures

The dependent measures were that the performance was measured by examining the movement time (*MT*) and Error Rates (*ER*). The movement time was measured from the start point and the time when the cursor/pointer enters the target. An error was recorded (for version 4 design prototype only) if a participant deviated from the path to achieve a task by moving the box away from the path, which produced the appearance of a yellow circle.

Dependent measures	Description
Error Rate (*ER*)	~~E~~rror attempt is recorded
Movement Time (*MT*)	The time length between the start point and the time the cursor/pointer enters the target is measured

Table 4.1: Objective measures of the human performance (Pilot Experiment)

In regards to the objective human performance, these objective measures were collected during the experiment with mouse, stylus and touch (summarized in *Table 4.1*. As for the subject feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire (*see Appendix A7*) shown in *Table 4.2*.

Independent Variables	Scale	
Design	Easy	Difficult
Feelings	Stressful	Comfortable
	Very Frustrating	Not very Frustrating

Table 4.2: Subject attributes of the Five-point Likert scale (Pilot Experiment)

The subjective opinions were measured by means of a post-experiment online questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on all 4 versions of the design prototype, considering the task as well as thoughts for routine/future tasks, feeling about running the task, ease of remembering ordering of the last version to the next version and feelings of satisfaction in using the mouse, stylus and touch for the task.

4.6 APPARATUS AND MATERIALS

4.6.1 Testing apparatus

The room used for the experiment was a small office room on the Second Floor in the Multimedia and Technology Centre, KUPUSB. The max capability of the room allows one participant to be assessed in a single shot, shown in *Figure 4.1*:

Figure 4.1: Experiment Area (Pilot Experiment)

This experiment was conducted based on the following equipment/tools:

- A Multi touch-based PC tablet (Acer Iconia Tab W500 - capacitive touchscreen[3] supports four-point multitouch)
- Windows 7 32-bit (AMD 1.0 GHz, 2GB RAM)
- Resolution of 1280 x 800
- Screen size 10.1”
- Capacitive Stylus
- Standard three-button optic mouse with 800 dpi, manufactured by Acer®;
- A Five-point 'Likert-type' scale (Likert, 1932) questionnaire (*see Appendix A*7), used to collect the user profile (i.e. age, gender, etc.) and subjective feeling about the device design and the discomfort in the particular body region;
- The data analysis is performed using SPSS version 13.

4.7 TASK

As shown in *Table 4.3* the boxes measurements in cm were 4 x 4, 3 x 3, 2 x 2 and 1 x 1. The boxes position for prototypes Version 1, 2, 3 and 4 are all similar (Black D,3: Red

---

[3] Capacitive touchscreens are those that respond to the electrical properties of the human body. This means that they can be controlled by a light touch, and don't require the user to exert heavy pressure on the screen.

D,7) . The merge accuracy between prototypes Version 1, 2, 3 and 4 shows that as the user moves to the next tasks the boxes become smaller and the level of difficulty changes from Easy to Hard.

Factors/Parameters	**Levels**
Box Size (cm)	4 x 4
	3 x 3
	2 x 2
	1 x 1
Merge Accuracy	Prototype Version 1 – Difficulty level: Easy *(Flick)*
	Prototype Version 2 – Difficulty level: Medium *(Auto merges if both touch 80% of each other)*
	Prototype Version 3 – Difficulty level: Hard *(Auto merges if both touch 100% of each other with 100% accuracy)*
	The Prototype Version 4 – Difficulty level: Hard *(Auto merges if both touch 100% of each other with 100% accuracy – the black box must be moved along a path*)

Table 4.3: Target Condition (Pilot Experiment)

4.7.1 Prototypes

There are 4 prototypes.

Prototype 1, 2 and 3 box positions are all similar, the difference are the difficulty level of each prototype.

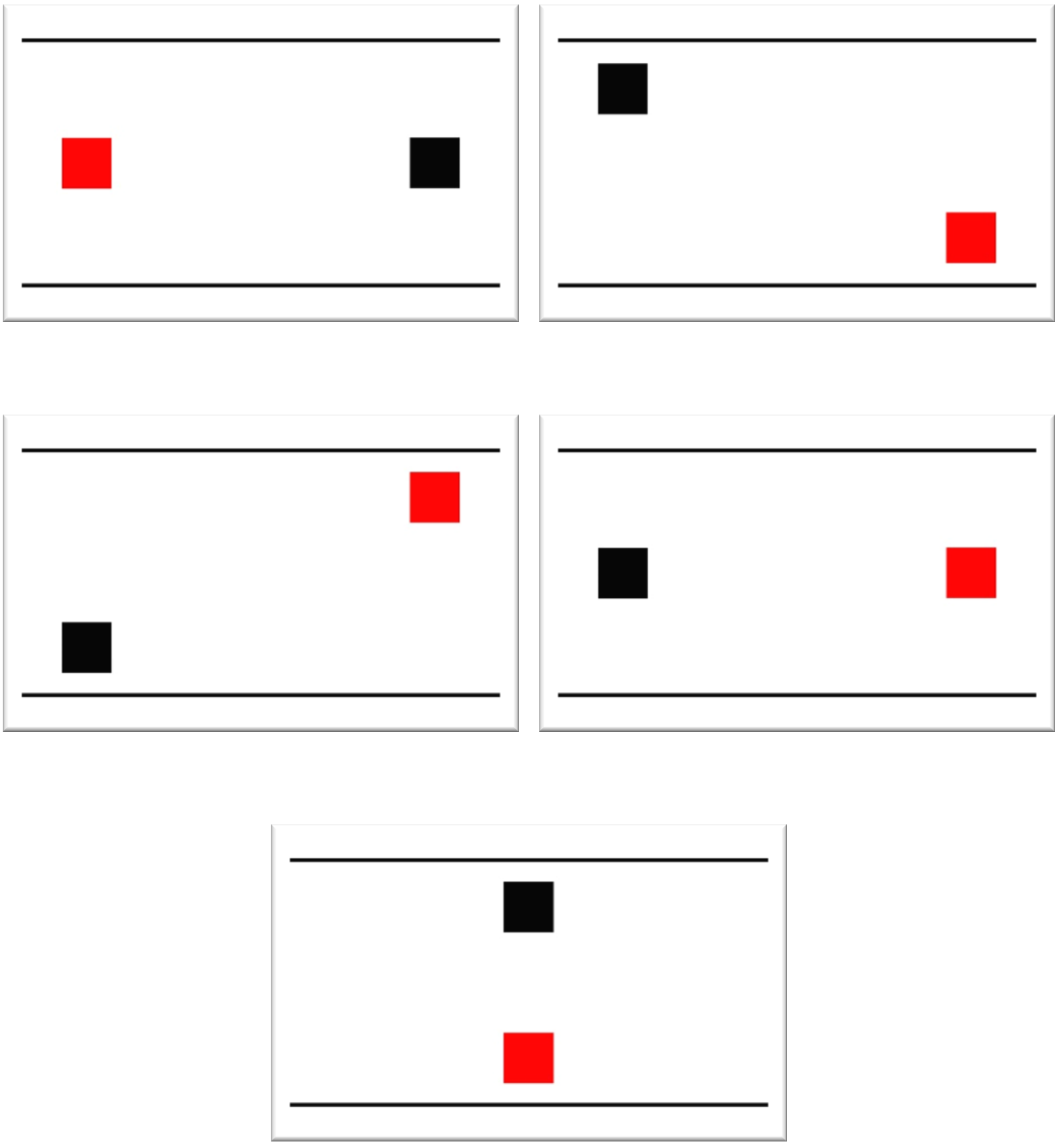


Figure 4.2: Experiment prototype layout, Difficulty: Easy (Pilot Study)

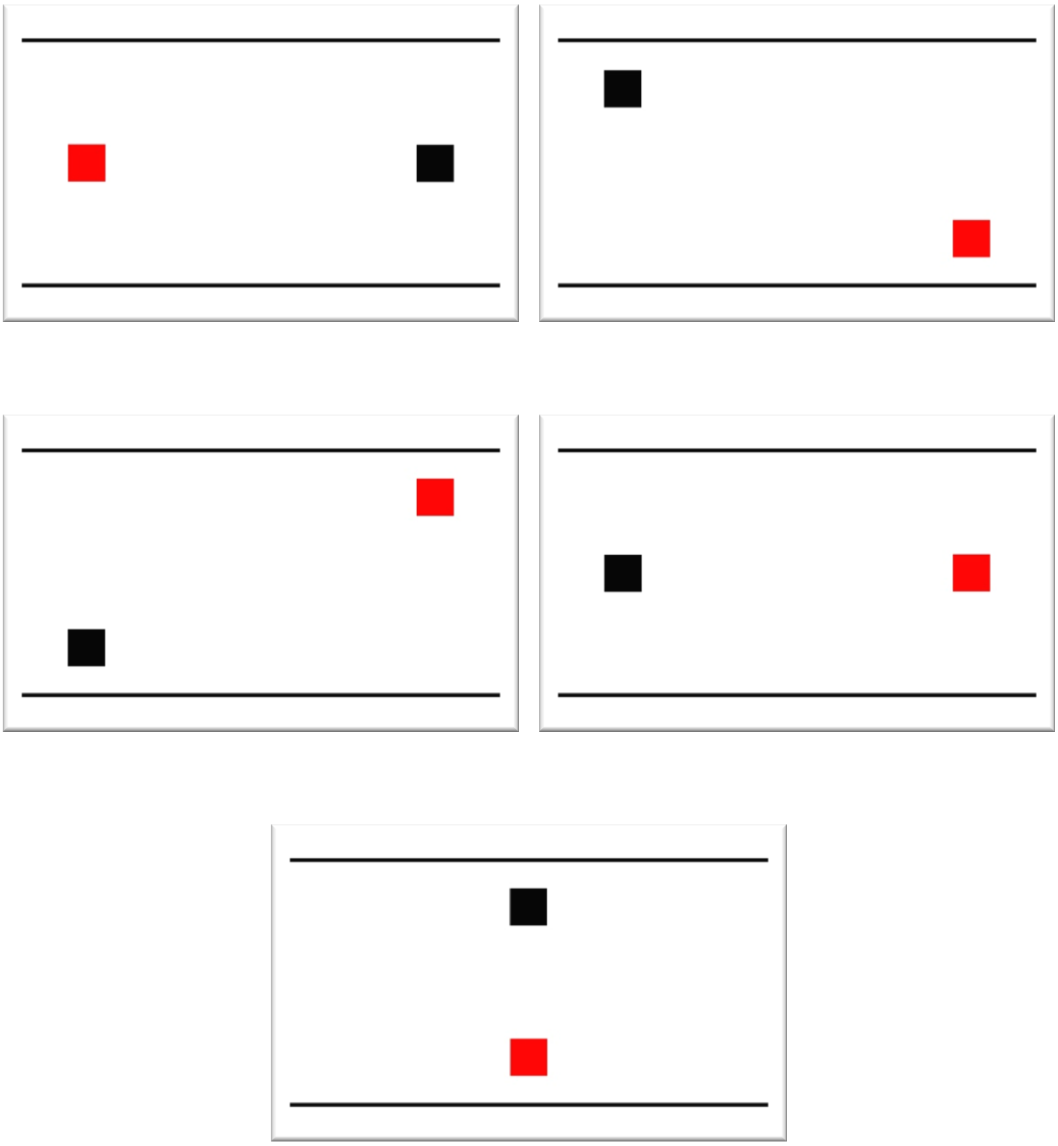


Figure 4.3: Experiment prototype layout, Difficulty: Medium (Pilot Study)

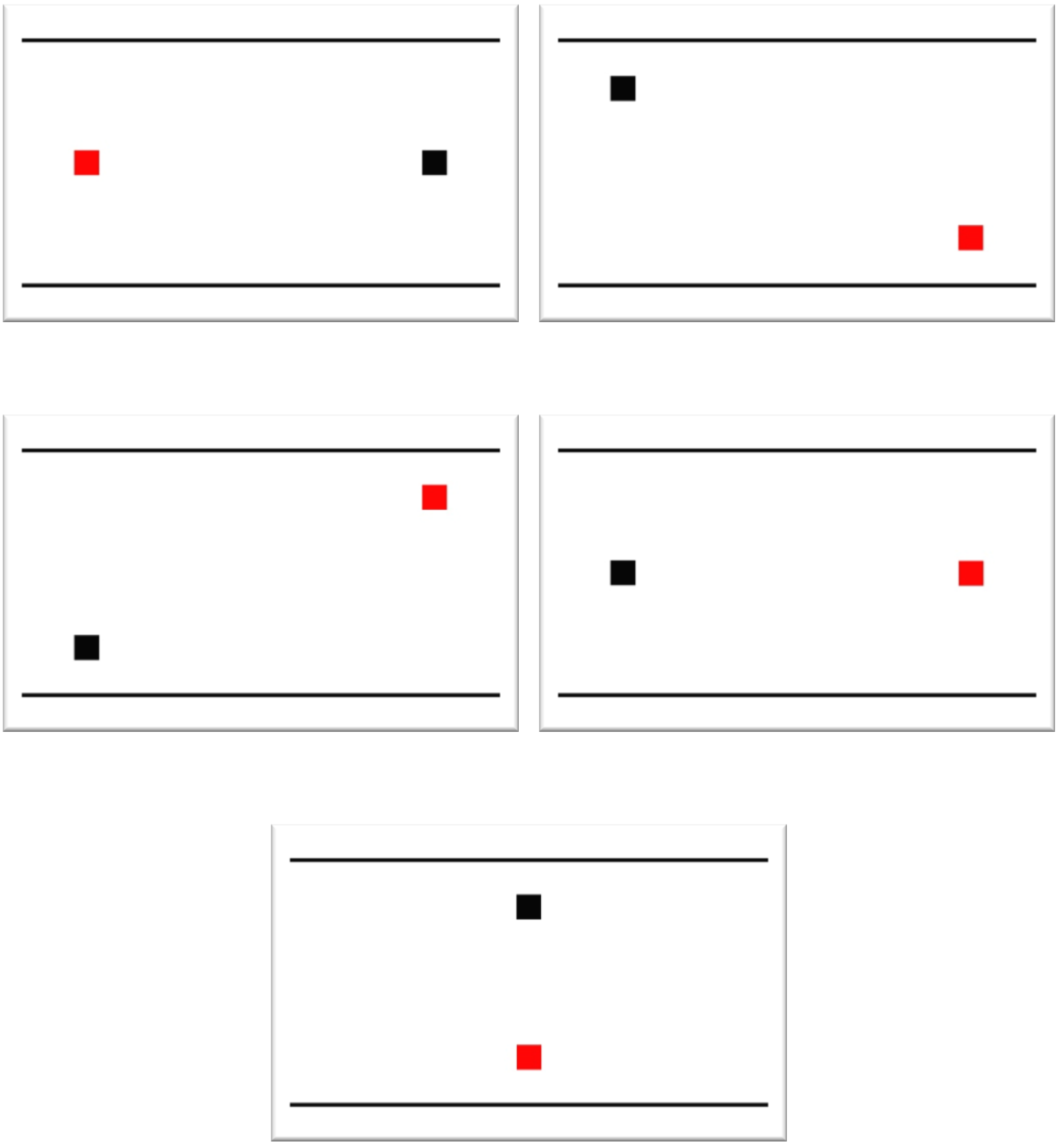


Figure 4.4: Experiment prototype layout, Difficulty: Hard (Pilot Study)

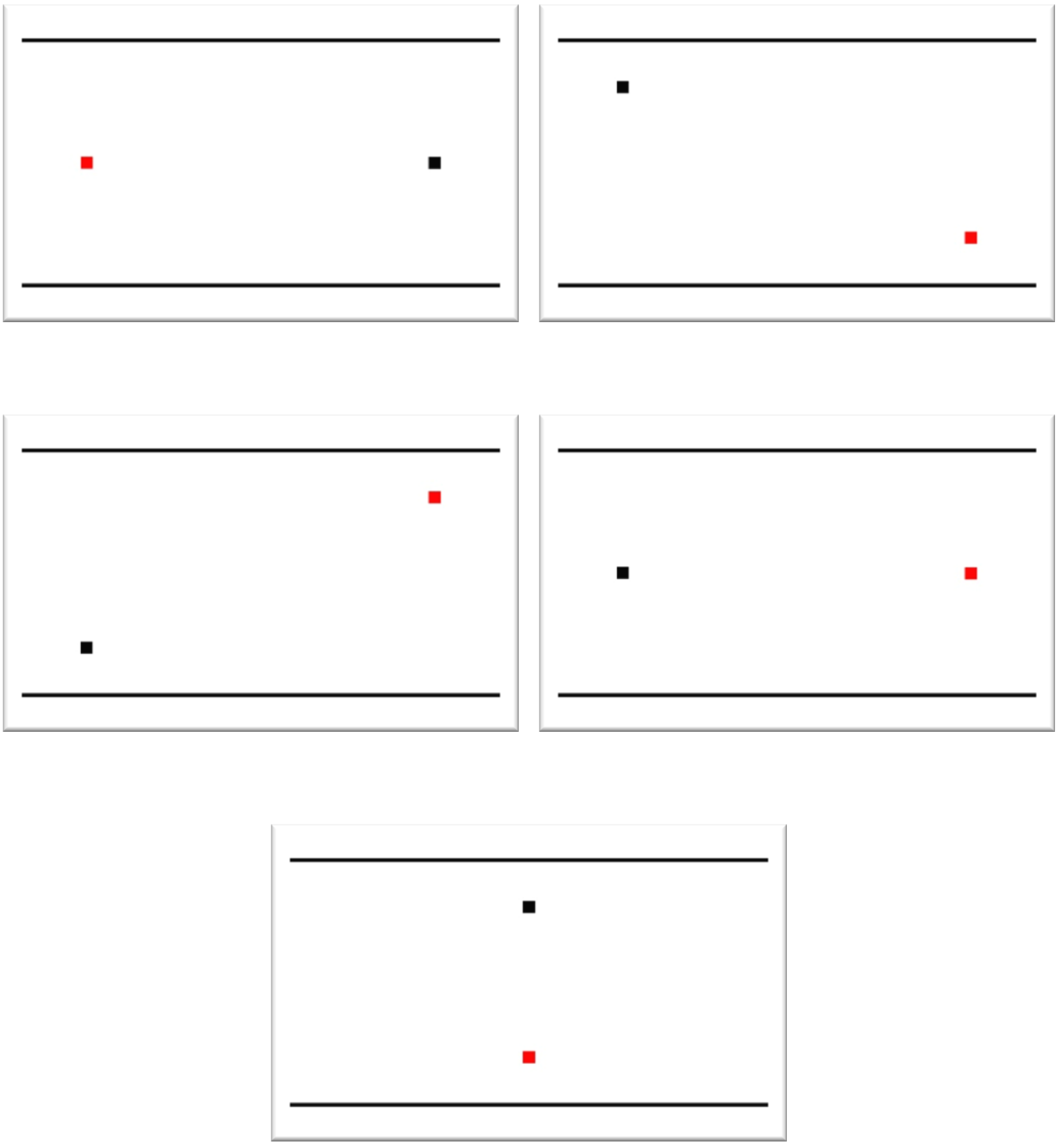


Figure 4.5: Experiment prototype layout, Difficulty: Difficult (Pilot Study)

4.7.2 Prototype Version 1

Using *Touch, Stylus and Mouse,* participants were asked to move the Black Box to the Red Box. Altogether this version comprises 20 tasks. This is the easiest difficulty level of the test as the merge accuracy for this version is using ‘Flick’. To complete the task, both boxes need to touch each other before it will automatic merge.

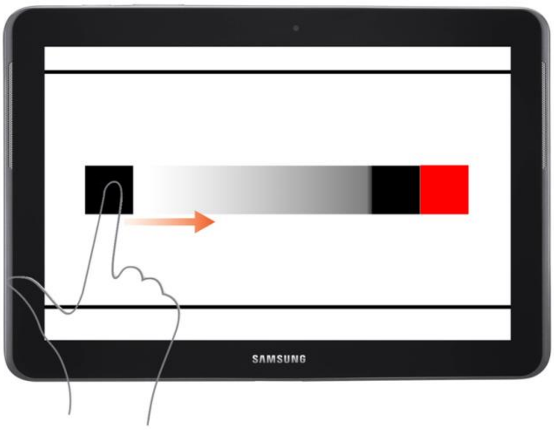


Figure 4.6: Experiment prototype layout version 1 (Pilot Study)

4.7.3 Prototype Version 2

Using *Touch, Stylus and Mouse,* participants were again asked to move the Black Box to the Red Box. In total this version likewise consists of 20 tasks. The level of difficulty of the test is medium as the merge accuracy for this version is using 80% auto merge. In other words, if both boxes touch 80% of each other, they will automatically merge. To complete the task, both boxes need to touch 80% each other before it will automatic merge.

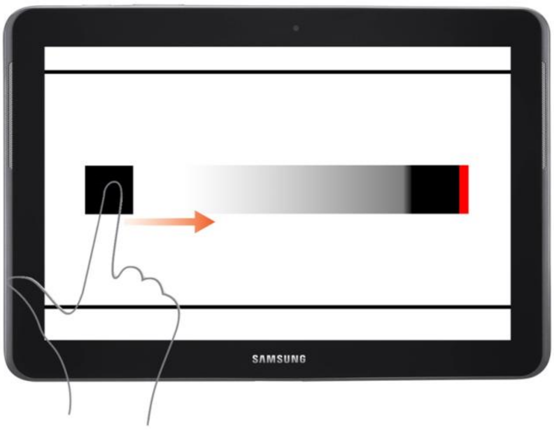


Figure 4.7: Experiment prototype layout version 2 (Pilot Study)

4.7.4 Prototype Version 3

Using *Touch, Stylus and Mouse*, participants were yet again tested to move the Black Box to the Red Box. At this point, the level of difficulty increased from medium to hard since the merge accuracy for this version is 100% auto merge. That is, if both boxes 100% touch each other, they will automatically merge. As the previous version, this version similarly involves 20 tasks. To complete the task, both boxes need completely touch each other before it will automatic merge.

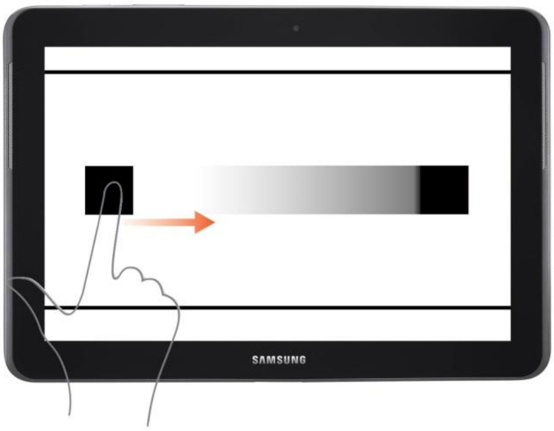
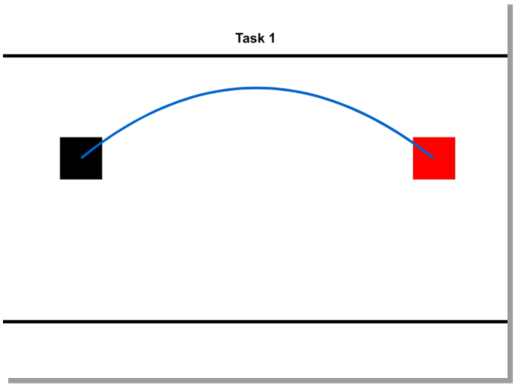
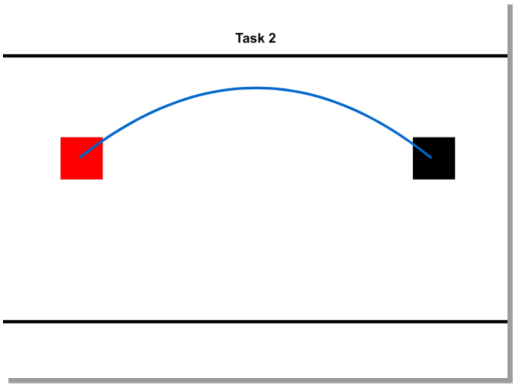


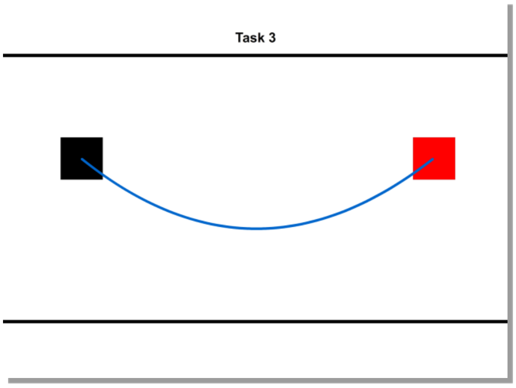
Figure 4.8: Experiment prototype layout version 3 (Pilot Study)

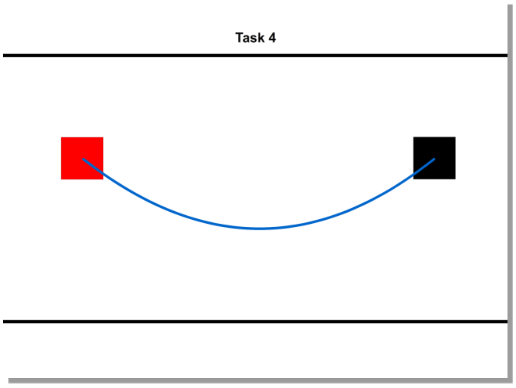
4.7.5 Prototype Version 4

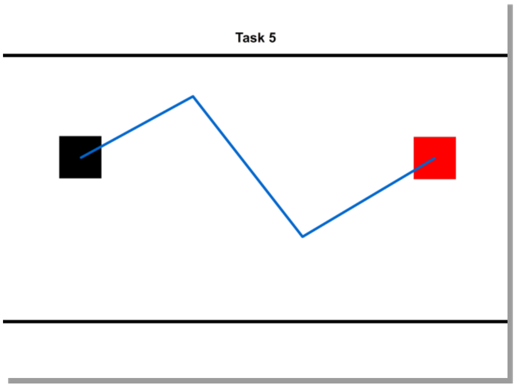
The last version is correspondingly testing the movement of the Black Box to the Red Box using *Touch, Stylus and Mouse.* This is the hardest level of all the versions. Despite the participants' abilities to move the Black Box to the Red Box accurately with 100%, merge, the black box must now be moved along a path before it can touch the Red Box. This version only consists of 16 tasks. In what follow are the screenshots of the experiment prototype layout for version 4 (Pilot Study) (*see Figure 4.10*).

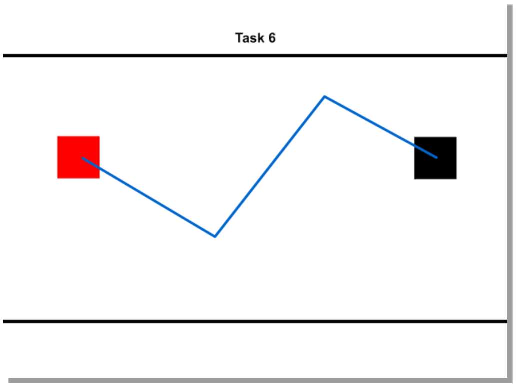

Task 1

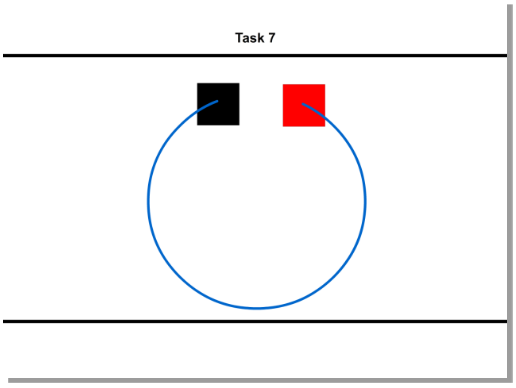

Task 2

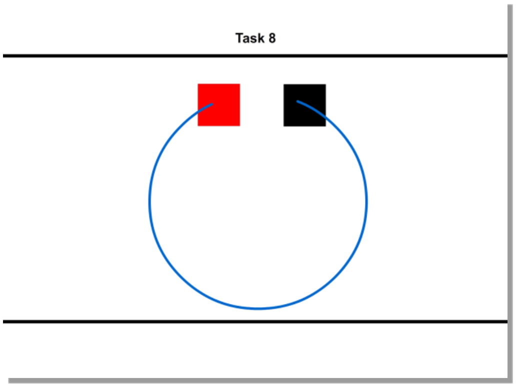

Task 3


Task 4


Task 5


Task 6


Task 7


Task 8

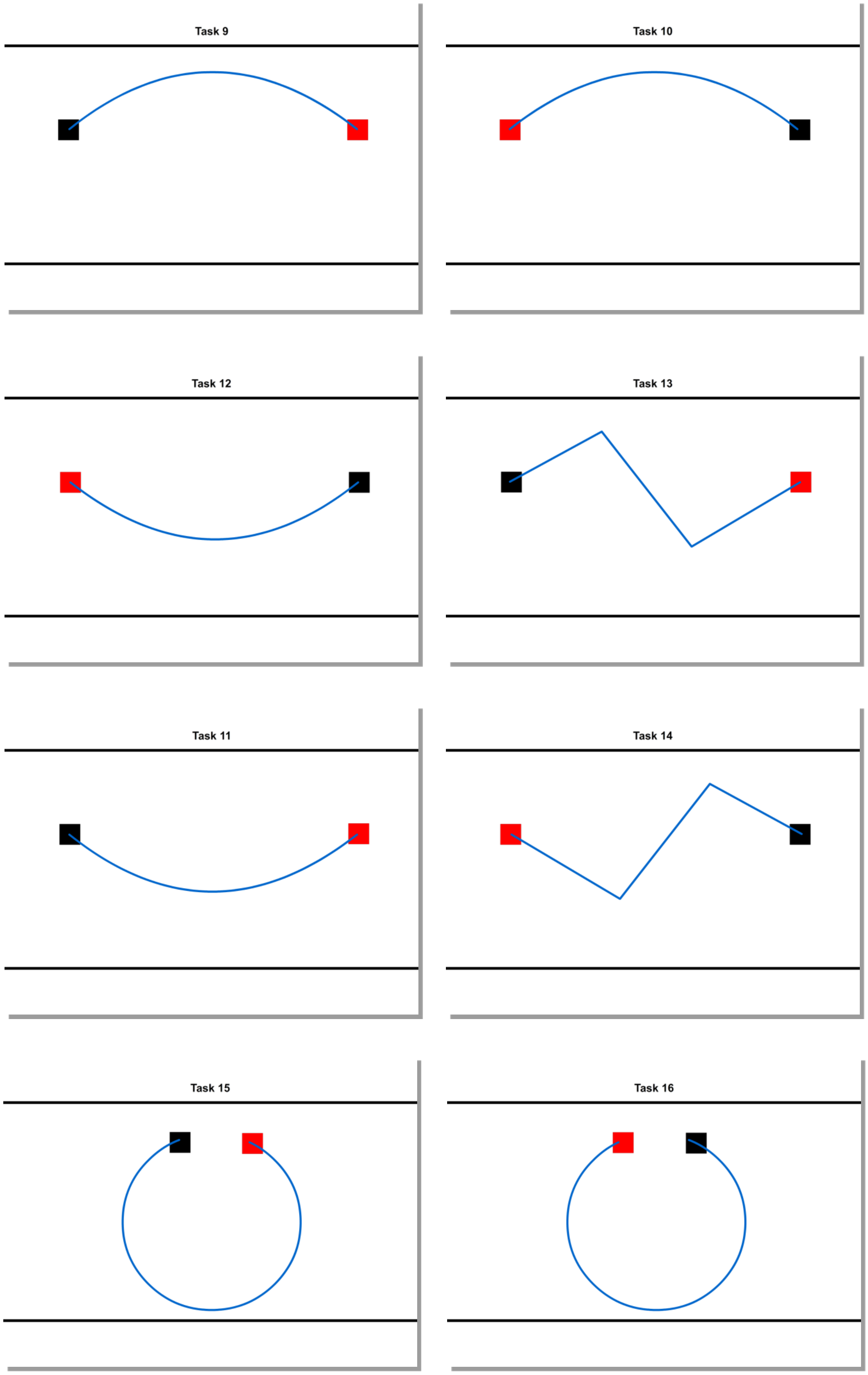


Figure 4.9: Experiment prototype layout version 4 (Pilot Study)

4.8 PROCEDURE

Each participant was asked to present themselves to a specific room in the institution set aside for the experiment. During the experiment each participant was seated at the desk in the room with the laptop facing them and the researcher sat opposite the participant with the second monitoring laptop facing the researcher.

Each participant was briefed about the touch-based prototype and it was stated that the study was evaluating the touch-based prototype rather than the participants.

A Standard Operation Procedure (SOP)[4], shown in *Figure 4.10*, is developed using a checklist to allow each participant to follow the same procedure during the experiment, which could help in reducing process bias during the experiment and to ensure reliability of the study.

There were two sections to the experiment: in section 1 of the SOP, the experimenter introduced the SOP to participants and demonstrated each task to familiarize the participants with the task and the laboratory environment. After that, participants were asked to sign a Consent Form to ensure that ethics requirements were met to the experiment. Participants were then asked to fill out ‘personal information’ to gather demographic data, i.e. age, gender, preferred hand, and experiential data such as computer experience.

In section 2 of the SOP, participants were instructed to perform each task *“as accurately as possible and as fast as possible”* (Zhai *et al*., 2004). The task was designed using a simple dragging task to determine the speed and accuracy of the movement object. There were 2 boxes with 3 different sizes (small, medium, large) and the participant was instructed to ‘drag’ the primary box and place it onto the target box, then release. Each size of the boxes was tested on 6 different target box locations,

---

[4] *‘A Standard Operating Procedure (SOP) is a set of written instructions that document a routine or repetitive activity followed by an organization. The development and use of SOPs are an integral part of a successful quality system as it provides individuals with the information to perform a job properly, and facilitates consistency in the quality and integrity of a product or end-result’.* (U.S. Environmental Protection Agency (2007, p.1).

making a total of 18 tests (3 different sized boxes x 6 different target box locations). In order to perform the test, the participant was asked to use touch, stylus and mouse in the same order. Firstly they were asked to use mouse, then touch and followed by stylus.

Throughout the tasks, the system recorded the movement time (*MT*) and error rates (*ER*) for each participant. After completing the tasks, the participants were prompted to fill out an electronic post-experiment questionnaire online via SurveyMonkey concerning user satisfaction.

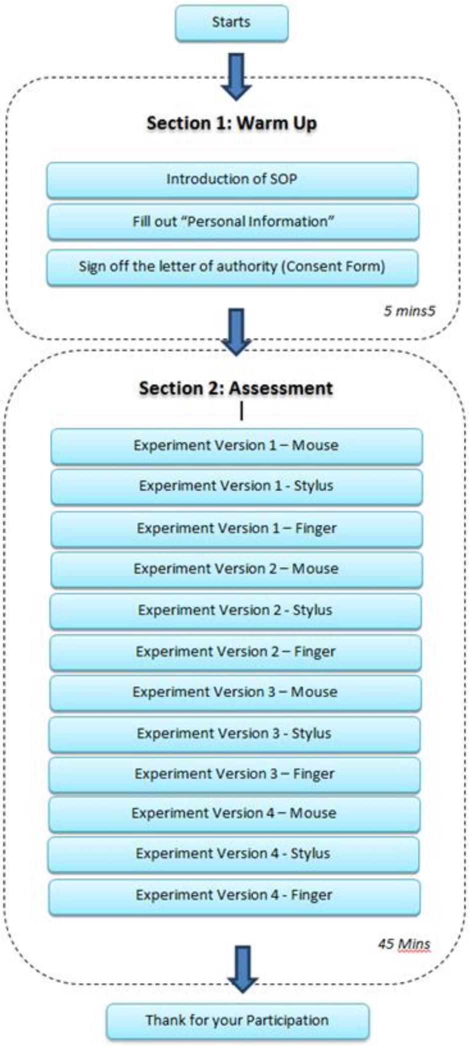


Figure 4.10: Standard Operation Procedure (SOP) (Pilot Experiment)

4.9 RESULT& DISCUSSION

4.9.1 Movement Time for a task

In the first stage, participants were asked to complete four different versions of the tasks using three different input devices (mouse, touch and stylus) and movement time was recorded for each participant.

Version	N	Device			P value
		Mouse	Stylus	Touch	
1	22	24.38 (17.84)	27.78 (25.63)	21.88 (17.34)	.114
2	22	20.98 (141.69)	25.74 (23.31)	25.39 (23.55)	.727
3	22	36 (27.60)	45.50 (43.81)	48.82 (47.48)	.036
4	22	58.29 (43.43	62.83 (64.9)	57.03 (55.90)	.956

Table 4.4: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Friedman test (Pilot Experiment)

** Version 1 - Accuracy in relation to speed: Easy (Flick)*

** Version 2 - Accuracy in relation to speed: Medium (Auto merges if both touch 80% each other)*

** Version 3 - Accuracy in relation to speed: Hard (Auto merges if both touch 100% of each other with 100% accuracy)*

** Version 4 - Accuracy in relation to speed: Hard(Auto merges if both touch 100% of each other with 100% accuracy – the black box must be moved along a path)*

Table 4.4 indicates that touch has the fastest overall movement time (*MT)* on versions 1 (21.88 sec ±17.34) and 4 (57.03 sec ±22.90), whereas the mouse shows the fastest movement time (*MT*) on versions 2 (20.98 sec ±141.69) and 3 (36 sec ±27.60).

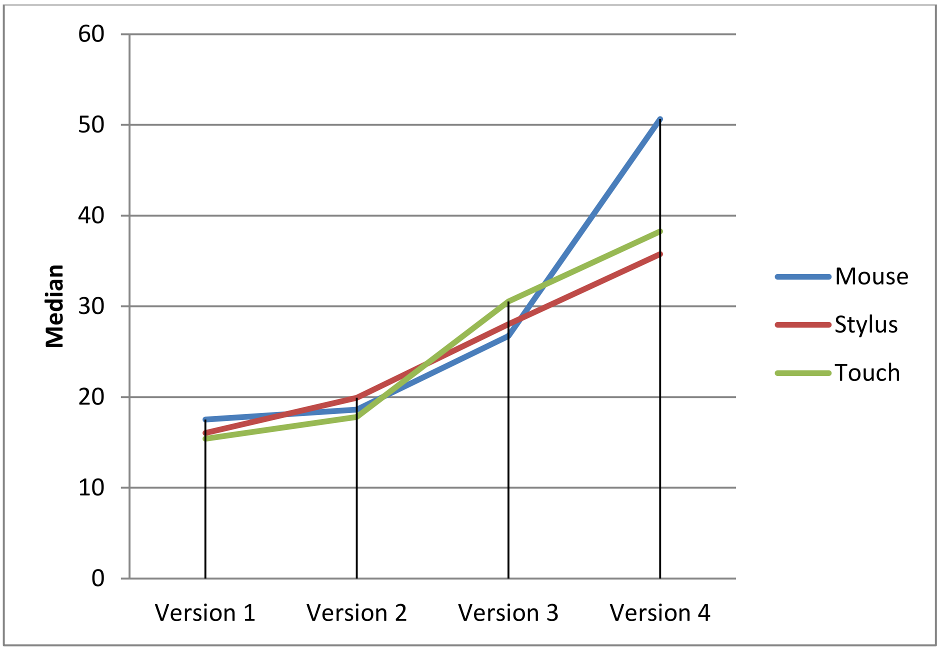


Figure 4.11: Movement time for all versions (Pilot Experiment)

A Friedman test was run to determine if there were differences in the movement time between three input devices. The test shows that there is no statistically significance difference among all versions of the test.

The movement time for **version 1** indicate that the three input devices became slower from mouse (Mdn = 17.51), to stylus (Mdn = 16.05), to touch (Mdn = 15.39), but the differences were not statistically significant, $\chi^2(2) = 4.345$, $p < .0005$

The movement time for **version 2** indicate that the three input devices became slower from stylus (Mdn = 19.93), to mouse (Mdn = 18.6), to touch (Mdn = 17.79), but the differences were not statistically significant, $\chi^2(2) = .636$, $p < .0005$

The movement time for **version 3** indicate that the three input devices became slower from touch (Mdn = 30.55), to stylus (Mdn = 28.04), to mouse (Mdn = 26.71), but the differences were not statistically significant, $\chi^2(2) = 6.636$, $p < .0005$

The movement time for **version 4** indicate that the three input devices became slower from mouse (Mdn = 50.64), to touch (Mdn = 38.26), to stylus (Mdn = 35.76), but the differences were not statistically significant, $\chi^2(2) = .091$, $p < .0005$

4.9.2 User satisfaction level

In the pilot study a brief survey was conducted online. Participants were asked to answer questions related to the functionality of the three input devices and were asked to rate each input devices based on their experience of using it.

Subjective Feeling	N	Devices		
		Mouse	Stylus	Touch
Overall consider the task of using	15	4.71 (.61)	2.43 (.94)	3.21 (1.12)
Overall consider using different input devices for routine/future task	15	4.53 (.64)	2.27 (.80)	3.60 (1.45)

Table 4.5: Result analysis of the Five-point Likert scale subjective assessment with the touch, mouse and stylus (Pilot Experiment)

A one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences in the user consideration of using different input devices and the user consideration of using different input devices for routine/future task. There were no outliers and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively.

In regards to the users' views on using the three input devices, the assumption of sphericity was not violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 2.790$, $p = .248$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.828$). Moreover it was statistically significantly different in the users' views on using the three input devices, $F(2, 26) = 21.432$, $p < .0005$, partial $\eta^2 = .622$, with the participant using mouse (4.71 ±0.611) followed by touch (3.21 ±1.122) then stylus (2.43 ±0.938. Post hoc analysis with a Bonferroni adjustment revealed that the users' views on using the three input devices was statistically significantly high from mouse to stylus (2.29 (95% CI, 1.56 to 3.02), $p < .005$), and from mouse to touch (1.50 (95% CI, 0.47 to 2.53), $p = .004$), but not from touch to stylus (0.79 (95% CI, -0.34 to 1.91), $p = .230$).

In regards to the users' views on using the three input devices for routine/future tasks, the assumption of sphericity was not violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 2.962$, $p = .227$. Therefore, a Greenhouse-Geisser correction was applied (ε = 0.831). Furthermore it was statistically significantly different in these views for future work, $F(2, 28) = 15.845$, $p < .0005$, partial $\eta^2 = .531$, with the increase in the order of the task of using **mouse** (4.53 ±0.640) followed by **touch** (3.60 ±1.454) and then **stylus** (2.27 ±0.799).

Post hoc analysis with a Bonferroni adjustment revealed that the difference between the users' views when considering their use of the three input devices for routine/future task was statistically significantly high from mouse to stylus (2.27 (95% CI, 1.45 to 3.08), $p < .005$), but not from mouse to touch (0.93 (95% CI, -0.27 to 2.13), $p = .004$), and from touch to stylus (1.33 (95% CI, 0.1 to 2.57), $p = .230$).

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
Version 1	Consider the task of using	15	4.79 (.426)	3.43 (1.28)	3.86 (1.41)	.008*
	Feel running the task - Stress	15	4.86 (.363)	3.36 (1.28)	3.57 (1.40)	.000
	Feel running the task - Frustration	15	4.79 (.579)	3.29 (1.33)	3.57 (1.56)	.002
Version 2	Consider the task of using	15	4.50 (1.09)	2.71 (.99)	2.86 (1.35)	.000
	Feel running the task - Stress	15	4.50 (1.09)	2.79 (1.12)	2.64 (1.39)	.000
	Feel running the task - Frustration	15	4.43 (1.09)	2.64 (1.15)	2.57 (1.28)	.000
Version 3	Consider the task of using	15	4.14 (1.09)	2.43 (1.02)	2.64 (1.22)	.000
	Feel running the task - stress	15	4.29 (1.14)	2.43 (1.02)	2.21 (.975)	.000
	Feel running the task - Frustration	15	4.21 (1.12)	2.43 (.94)	2.21 (.975)	.000
Version 4	Consider the task of using	15	4.43 (.852)	2.57 (1.28)	2.71 (1.44)	.000
	Feel running the task - Stress	15	4.50 (.855)	2.57 (1.16)	2.50 (1.29)	.000
	Feel running the task - Frustration	15	4.21 (1.250	2.50 (1.16)	2.43 (1.22)	.000

Table 4.6: Result analysis of the Five-point Likert scale subjective assessment with the touch, mouse and stylus (Pilot Experiment)

Participants indicated their agreement with a series of statements about each device using a 5-point 'Likert-type' scale (Likert, 1932); Consider the task of using (1=Difficult, 5=Easy), Feeling in running the task (1=Stressful, 5=Comfortable) and

Feeling in running the task (1=Very Frustrating 5= Not Very Frustrating). The statements used, along with their mean values, are listed in *Table 3.6*.

The Repeated Measure ANOVA test indicates that there is no statistical significance in regards to the subjective feeling except in version 1 – consider the task of using. The participants overall considered it much easier using the mouse on the task given for all versions, rather than touch and stylus.

In terms of the overall feeling among the participants, the data shows that the mouse had significant approval of its usability, whereas the stylus recorded the worst results and as the mean shows, that majority of the users did not consider using the stylus in the tasks.

4.10 SUMMARY AND DISCUSSION

After reviewing the time interval study and user satisfaction survey it was concluded that the mouse is by far the most convenient and efficient input method. Average time duration for the mouse (*M = 34.6543, SD = 30.82233*) was the lowest followed by Touch (*M = 38.6126, SD = 41.42465*) and stylus (*M = 40.4619, SD = 30.82233*) and the result was statistically significant *($p<0.05$)*. Similarly the mouse was the easiest input device among others $\chi^2(8) = 28.379$, $p < 0.05$ and participants thought they would always prefer the mouse for their future/routine tasks as well $\chi^2(8) = 28.317$, $p < 0.05$.

It was suggested that there was a possibility that this prototype was easier for a mouse user rather than a stylus or touch. However, findings from the user feedback indicated that the problem with the stylus pointer was that it stuck to the tablet screen and therefore most likely led to bias result. Moreover the tablet used in the pilot experiment has a capacitive touch screen which needs a capacitive stylus. A capacitive stylus pen is typically made of soft rubber and works effectively on a smooth screen surface. The likelihood is that during the experiment dust and debris fell on the screen, which could explain the nonresponsive screen. Additionally, the capacitive stylus that was used was not built in with the tablet and had to be purchased separately whereby its suitability with the touch screen is uncertain. Therefore, four participants failed to complete the experimental tasks because of technical error where the tablet did not respond, which meant that the movement time and error rate could not be recorded, hence their data was excluded from the analyses. All reported analyses are only based on data obtained from the 22 participants who successfully completed the experiment. Thus, it is recommended to revise the use of apparatus and materials in future experiments to avoid this issue and show clearer results.

Furthermore it can be reflected from the pilot experiment questionnaire that the participants' feedback was not as productive as expected. The use of SurveyMonkey as a tool to conduct the questionnaire online that targeted 22 participants in the study only succeeded in collecting 15 feedbacks. It is concluded that in future experiments the questionnaire will be distributed immediately after the experiment to acquire all of participants' feedback. Additionally, the benefit of immediately distributing the questionnaire is that participants could still remember the tasks they had undertaken and could ask questions of the researcher if they failed to understand the requirements of the questionnaire.

To sum up, there were some issues that needed attention in the future experiment, such as the design of the questionnaire. It was thought that instead of just recording the participants' feeling about using the stylus, mouse and touch in the experiment tasks, it would also be useful to identify the fatigue effect of using the three input devices during the experiment. It was decided that in future an existing survey tool that had already been tested and validated would be used. As recommended by Lazar *et al*, (2010) *'if a survey tool has already been developed, there is no need to create one from scratch'.*

CHAPTER 5: ABSTRACT EXPERIMENTS

5.1 INTRODUCTION

In this chapter, experiments are deliberately developed as abstract tasks with the purpose of analysing the users' ability on simple tasks without a real world context. Chapter 4 consists of tracing, dragging, one direction tapping test and a multi directional tapping test. Overall, the aim was to discover if there are differences in mouse, stylus and touch on the tracing test and dragging test with different levels of difficulty that could affect users' performance and satisfaction.

Findings from pilot studies were taken into account in the experiment. The use of apparatus and materials was revised. The design of the prototype and questionnaire is based on the guidance of The International Organization for Standardization known as ISO (ISO 9241-420, 2011)

5.2 USERS

Participants were paid volunteers who were recruited through posters and personal contact. They were rewarded £10 each for completing both tests.

The initial recruitment of the participants took place by means of a recruitment questionnaire. Since the test involved testing aspects of using touch-based technology, it was deemed important to have participants with a certain amount of experience in using touchscreen devices and computers in general. This is because if there happened to be a number of beginners to such activities, these could lead to bias result. Therefore:

- All participants had computer experience.
- All participants had experience with a touch-based device.

- All the participants' uncorrected visual problems or physical limitations that would inhibit their use of the mouse, stylus and touch as an input device were accounted for.
- *Table 5.1* is the demographic of the sample group

		ABSTRACT			
		Tracing	**Dragging**	**One Direction Tapping**	**Multi Direction Tapping**
TOTAL USER		23	23	22	22
Gender	Male	19	19	7	7
	Female	4	4	15	15
Age	18 – 25	9	9	16	16
	26 – 39	12	12	4	4
	40 and above	2	2	2	2
Hand Use	Right	21	21	19	19
	Left	2	2	3	3
Academic Level	Undergraduate	3	3	15	15
	Postgraduate	20	20	1	1
	Others	0	0	6	6

Table 5.1: Demographics of sample group (Abstract)

They all signed an informed consent document informing them of the goals and activities of the study, their rights to terminate, and the confidentiality of their performance.

5.2.1 APPARATUS AND MATERIALS

5.2.1.1 Testing apparatus

The room used for the experiment was a small room on the first floor at the University of Salford library. The max capability of the room allows one participants to be assessed in a single shot, shown in *Figure 5.1*:

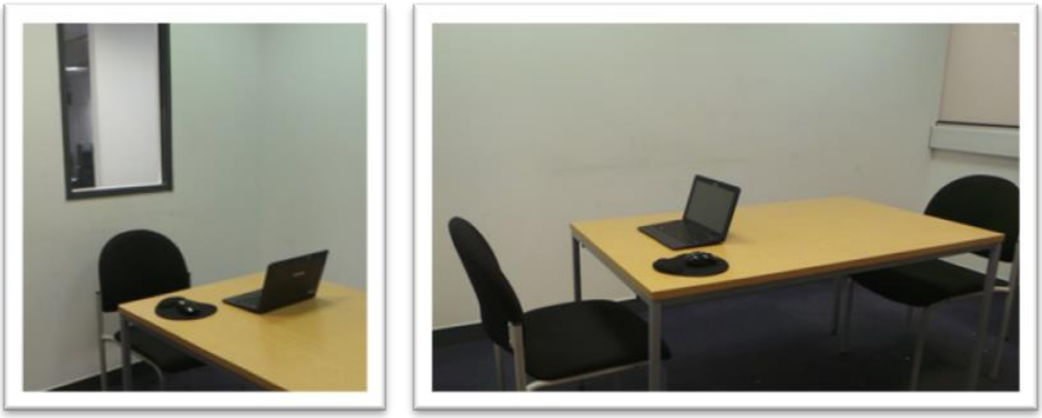


Figure 5.1: Experiment Area (Tracing Test and Dragging Test - Abstract)

This experiment was conducted based on the following equipment/tools:

- A convertible Samsung computer tablet with keyboard (Samsung ATIV Smart PC XE700T1C Tablet with Keyboard)
- Intel Core i5-3317U Dual Core Processor,
- Microsoft Windows 8 64bit, 64GB Storage, 4GB DDR3 RAM
- 11.6" Full HD Touch Screen
- 1920x1080 Full HD Resolution
- Wireless Optical PC five button mouse with 1000 dpi, manufactured by Samsung
- S-Pen stylus
- A Five-point 'Likert-type' scale (Likert, 1932) questionnaire (*see Appendix A7*), used to collect the participant profile (i.e. age, gender, etc.) and subjective feeling about the device design (general indices) and the discomfort (fatigue indices) in the particular body region.

5.2.1.2 Tasks

Experimental abstract tasks were presented by four different tests: Test 1 is a Tracing Test, Test 2 is a Dragging Test, Test 3 is One direction tapping and test 4 is Multi-directional tapping. All tests were written in adobe flash that runs under Windows 8.

Test 1 (Tracing Test) consists of four circles, each with a diameter of 100 mm. The participants were instructed to draw a free-hand line using *Touch, Stylus and Mouse* around each of the circles (*see Figure 5.2*).

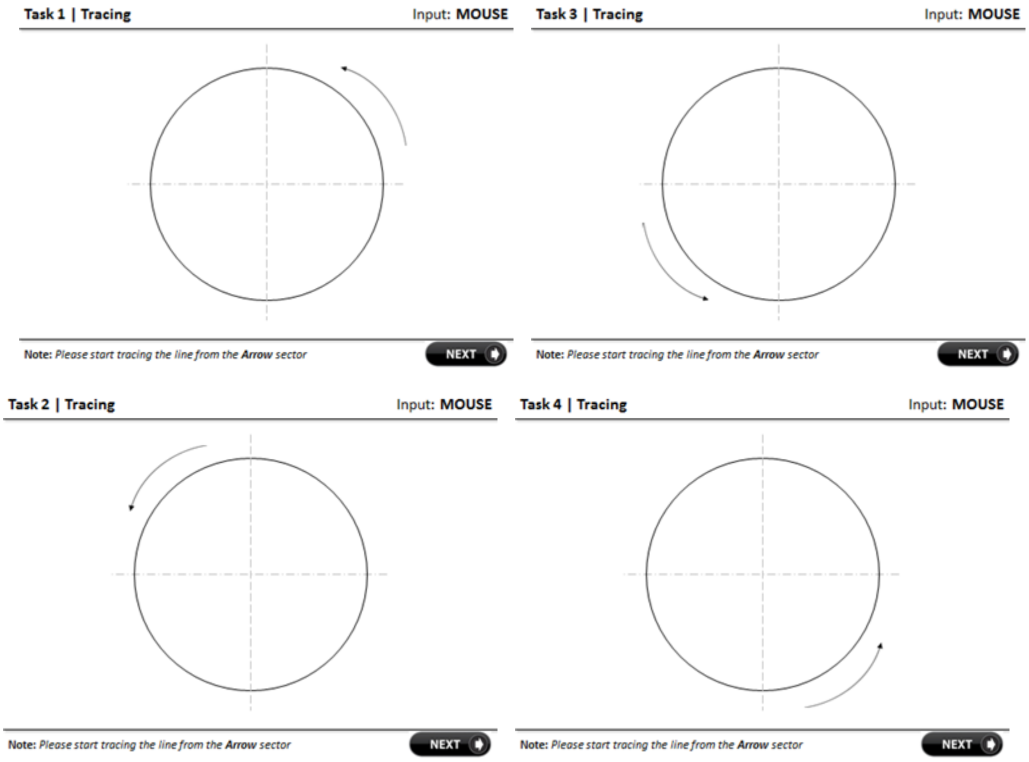
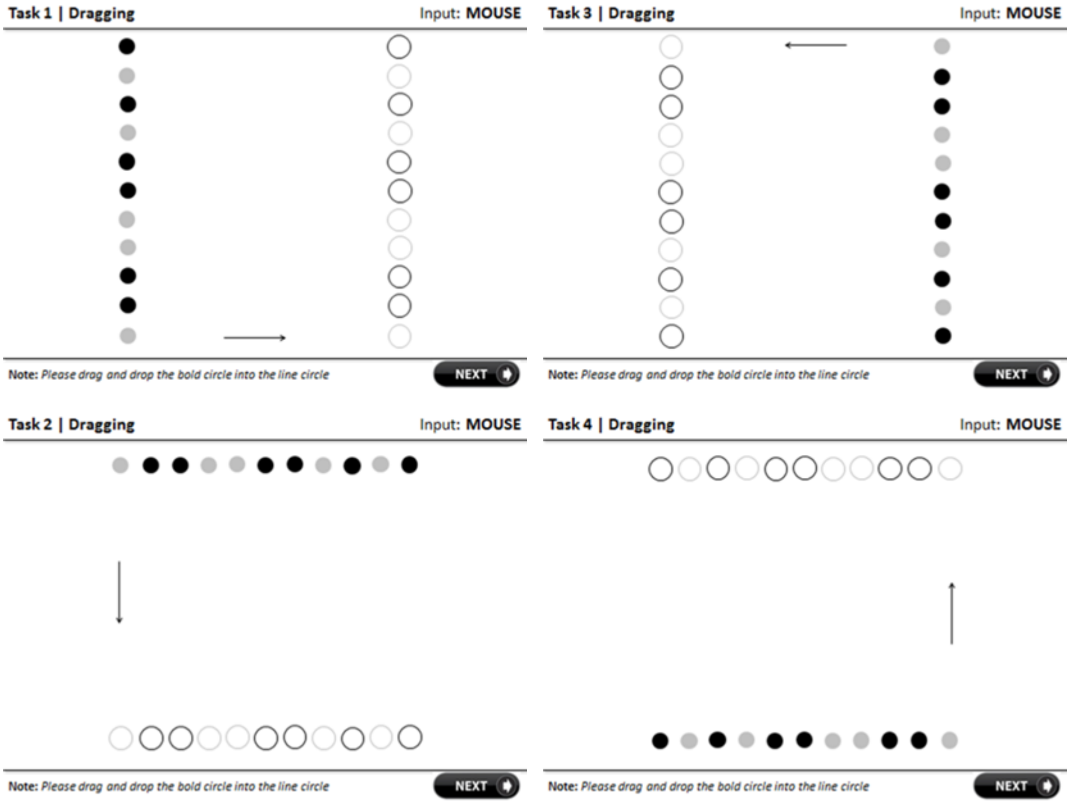


Figure 5.2: Test object and direction of movement (Tracing Test - Abstract)

Figure 5.3 shows the test 2 (Dragging Test) task in which the test object consist of circles with a diameter of 8 mm and over a distance of 100 mm. Participants were asked to place them in circles with a diameter of 10 mm and perform the task in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction. The movement time (*MT)* for each direction were then measured



Note: *Please drag and drop the bold circle into the line circle*

Note: *Please drag and drop the bold circle into the line circle*

Note: *Please drag and drop the bold circle into the line circle*

Note: *Please drag and drop the bold circle into the line circle*

Figure 5.3: Arrangement of the objects for the dragging task for the cardinal directions (Dragging Test - Abstract)

Test 3 (One Direction Tapping Test) consists of two rectangles with a defined width in the direction perpendicular to the direction of the movement (*see Figure 5.4*). The task consists of alternately tapping between the two rectangles. The participants were instructed to point and click, along one axis, within each rectangle 25 times using Touch, Stylus and Mouse. Each test session starts when the user first moves the pointer into a rectangle and actuates a button. This allows the participant to move quickly back and forth between the two rectangles.

This experiment consists of four tasks in which with the increasing difficulty of the task, the targets become smaller and the distance greater.

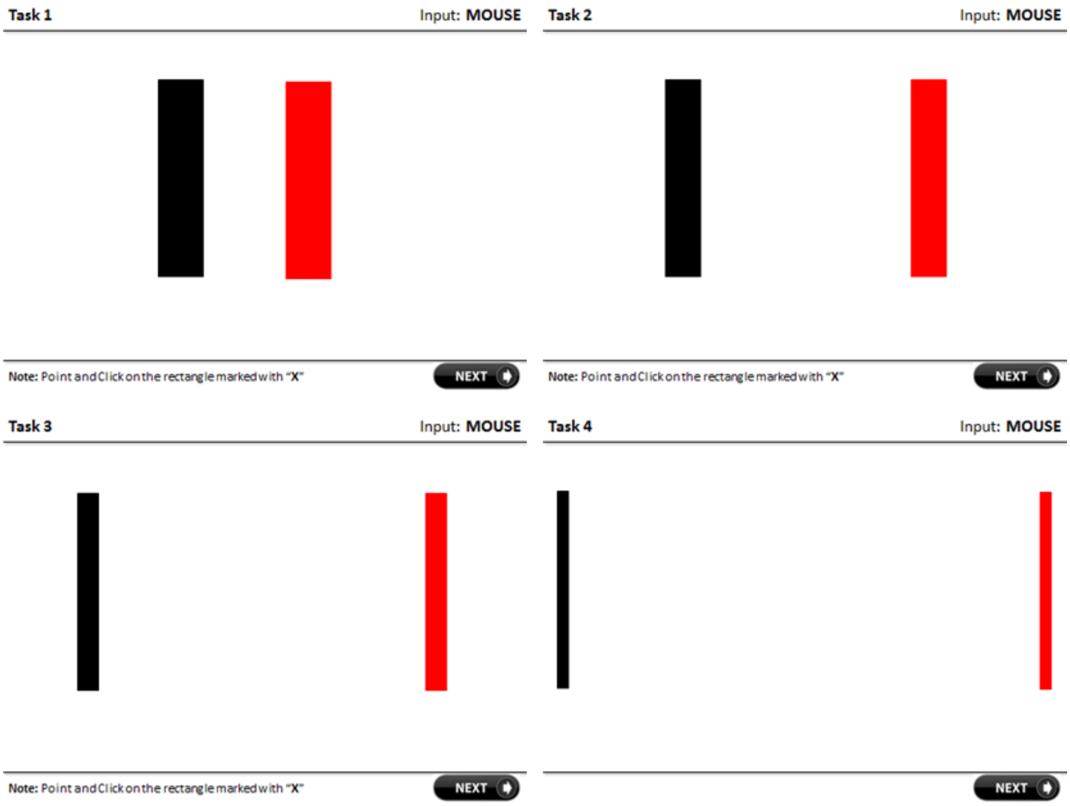
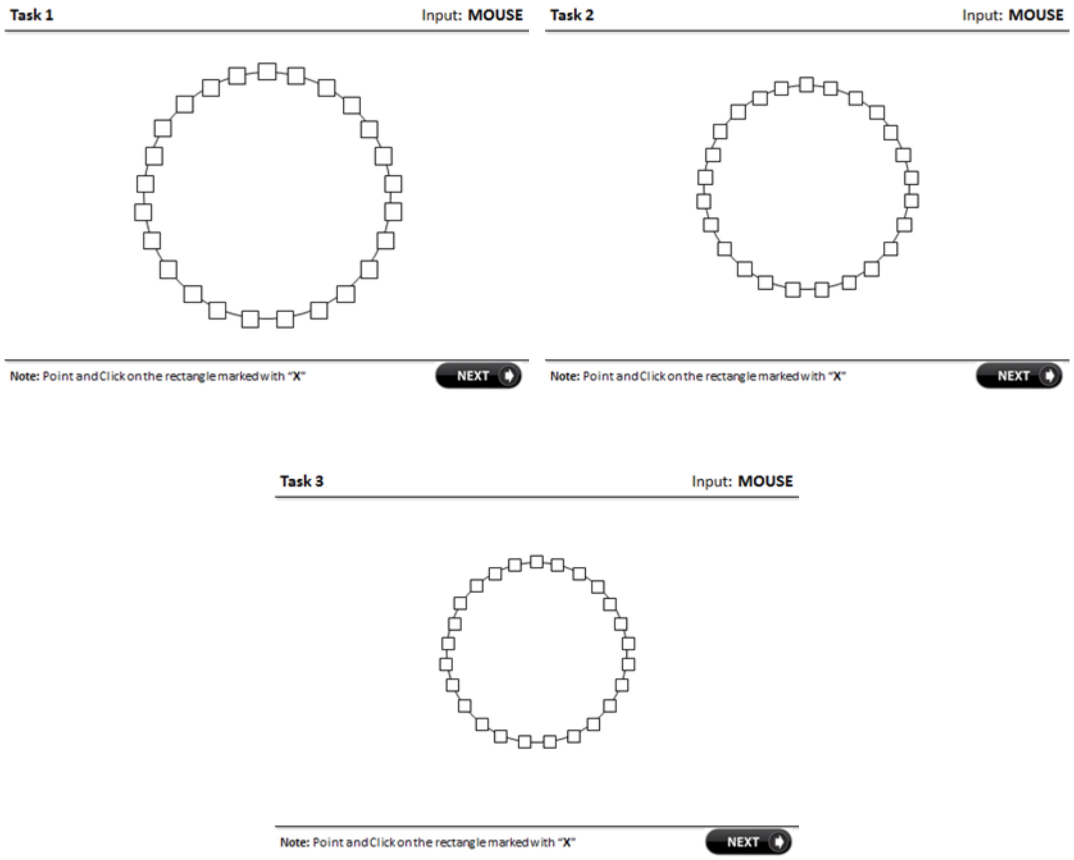


Figure 5.4: One-direction tapping task (Abstract)

The test 4 (multi-directional tapping test) consists of targets positioned around the circumference of a circle. The task consists of alternately tapping around the circumference of a circle made up of twenty-five small squares. The participants were instructed to point and click, along the circumference of a circle tapping each of the squares using Touch, Stylus and Mouse. Each test session starts when the user first moves the pointer into a square and actuates a button. This allows the participant to move quickly back and forth between squares. The target to which the participant should advance was marked with X. Each test session starts after the participant points to the topmost target and ends when the sequence is completed (at the topmost target).

This experiment consists of three tasks in which the target becomes smaller with the increase in difficulty of the task.



Note: Point and Click on the rectangle marked with "X"

Note: Point and Click on the rectangle marked with "X"

Note: Point and Click on the rectangle marked with "X"

Figure 5.5: Multi-directional tapping test (Abstract)

5.2.2 PROCEDURE

Participants were asked to present themselves to a specific room in the institution set aside for the experiment. During the experiment each participant was seated at the desk in the room with the laptop tablet facing them and the researcher sat opposite the participant.

There were two sections in the experiment: in section 1 of the SOP (*see* ***Error! Reference source not found.***), the experimenter introduced the SOP to participants and demonstrated each task to familiarize the participants with the task and the laboratory environment. After that, participants were asked to sign off a Consent Form to ensure commitment to the experiment. Participants were then interviewed and filled out ‘personal information’ to gather demographic data, i.e. age, gender, preferred hand, and experiential data such as computer experience.

In section 2 of the SOP (*see Figure 5.7*), participants were instructed to perform each task "*as accurately as possible and as fast as possible*" (Zhai *et al*., 2004). There were two different tests: test 1 the tracing test and test 2 the dragging test. The participants were instructed to use touch, mouse and a stylus in order to perform the test.

Participants were given the tracing test first. The task was explained and demonstrated to the participant. They were instructed to work as fast as possible while still maintaining high accuracy. Participants were also instructed to continue without trying to correct errors. Moreover, the prototype recorded variables such as movement time (*MT)* and the error rates (*ER*).

After completion of the tracing task, participants rested for a few minutes before receiving instruction on the dragging task.

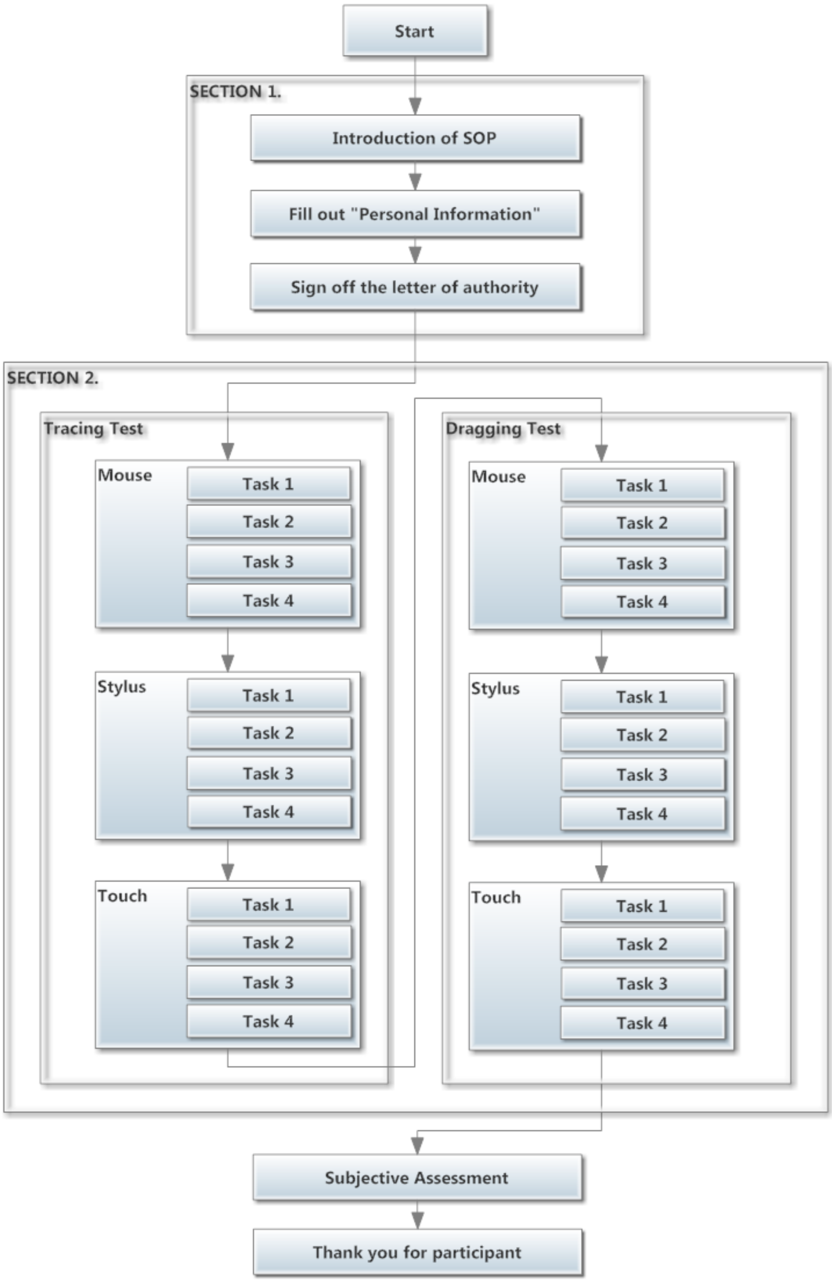


Figure 5.6: Standard Operation Procedure (SOP) (Tracing Test and Dragging Test - Abstract)

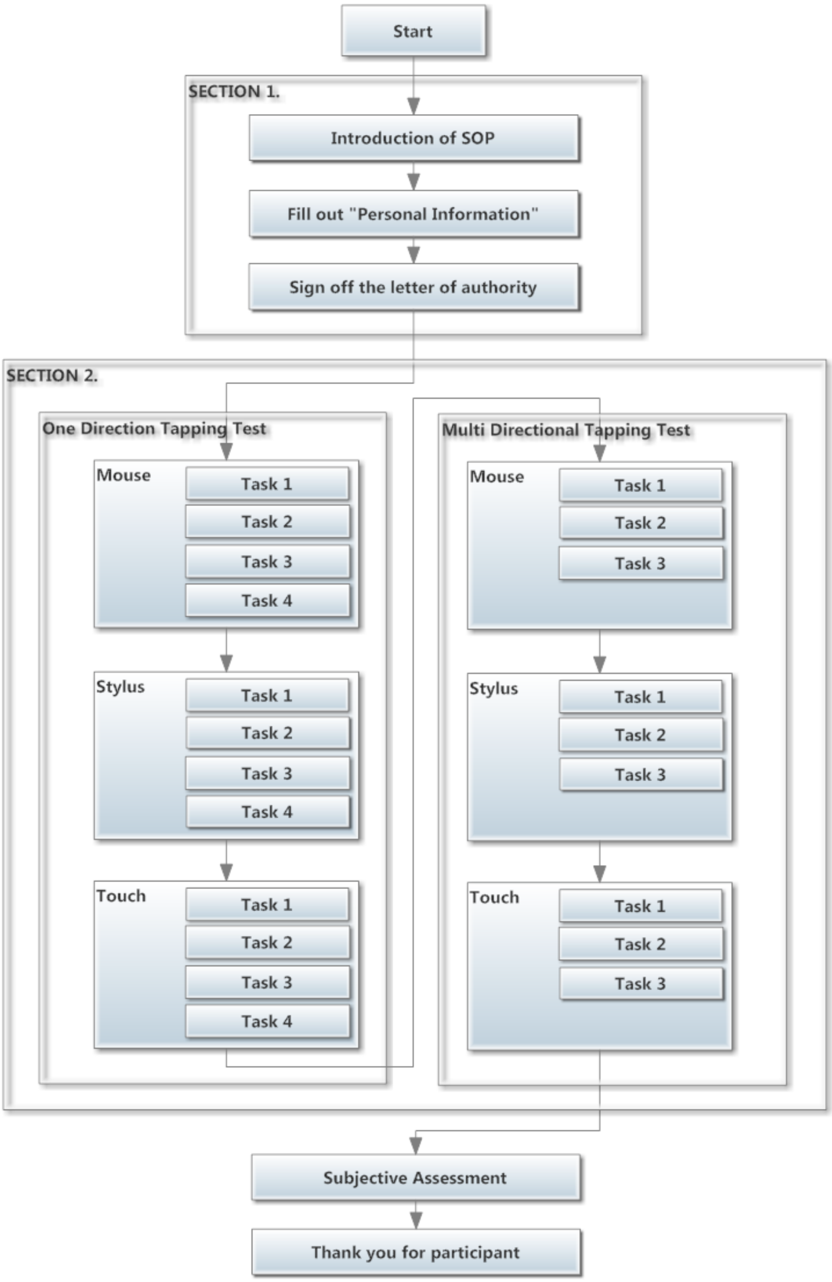


Figure 5.7: Standard Operation Procedure (SOP) (One Direction Tapping Test and Multi Direction Tapping Test - Abstract)

At the conclusion of the performance portion of the experiment, participants were asked to respond to a written questionnaire asking them to rate their experience in using the device. The questionnaire consisted of forty-seven questions covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

In this questionnaire, the first 5 questions asked the subjects about general questions. The next 21 questions were asking about the first prototype tests and followed by another 21 questions asking the second prototype. In each experiments there are two prototype to be test by the subjects i.e. Test 1 (Tracing and Dragging), Test 2 (One Direction Tapping and Multi-Direction Tapping). Researcher then categorised the responses into fourteen data points as been listed in *Table 5.2*: Device Assessment Questionnaire (Tracing Test and Dragging Test - Abstract) that illustrates this device assessment questionnaire.

Participants were also explained the meaning of the various surveys terms. Researcher also in some occasions translated the incomprehensible and misunderstood terms to Malay language on the way to ensure that participants fully understood each term.

General indices

(Please (√) Tick Appropriate Box)

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

Fatigue indices

(Please (√) Tick Appropriate Box)

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

Table 5.2: Device Assessment Questionnaire (Tracing Test and Dragging Test - Abstract)

Source: BS EN ISO 9241-420 (2011)

The total time spent by each participant ranged from slightly less than an hour to one hour and 30 minutes. The performance section took between 45 minutes to one hour to complete.

5.2.2.1 Statistical Analysis

The data for movement time (*MT)* and Error Rate (*ER*) was collected directly by the prototype which presented the experimental tasks. The data was then prepared for further statistical analysis by computing values for movement time (*MT*) and Error Rate (*ER*). Descriptive Statistics and Inferential Statistics were performed using SPSS.

5.2.2.2 Device Assessment Questionnaire

The mean and standard deviation of the ratings for each of the forty-seven questions was computed. Given the ordinal nature of the data, the Friedman test non-parametric statistic was computed to test for significant differences between participants in the three device groups.

5.3 EXPERIMENT 1: TRACING TEST

The first test is the tracing test which evaluates the tracing of an object. The test object consists of four circles, each with a diameter of 100 mm. The participants are instructed to draw a free-hand line around each of the circles in the clockwise direction.

5.3.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

5.3.1.1 Tracing test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

H2: There is a difference in the amount of error using touch, mouse or stylus.

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus interaction.

5.3.2 EXPERIMENT DESIGN - TRACING TEST

A within users design was deployed because the tracing test aimed to carry out an investigation of the difference of the three input devices and the same participants engage in all the conditions. Twenty-three participants were assigned to do the test which included 12 tasks (3 inputs x 4 task) altogether. This design was chosen firstly, as stated by Langston (2014), to equate the groups in the experiment as every participant is in every group and there cannot be any differences because they are all the same people. Secondly, it is also chosen to promote efficiency as it greatly reduces the number of participants that are needed since it only requires a smaller sample size (Bannan-Ritland, 2003) and enhances the quality of research carried out (Langston, 2014) Moreover, Lazar *et al*. (2010) argue that for researchers who have difficulty in finding and recruiting qualified participants which is the frequent problem faced by many HCI researchers, within-in group design is more appropriate.

5.3.3 VARIABLES - TRACING TEST

5.3.3.1 Independent Variables

The independent variables were:

Factors/Parameters	Level
Circle diameter	100 mm
Task	4

The test object consists of four circles, each with a diameter of 100 mm. The participant were instructed to draw a free-hand line tracing on each of the circles (*see Figure 5.8*)

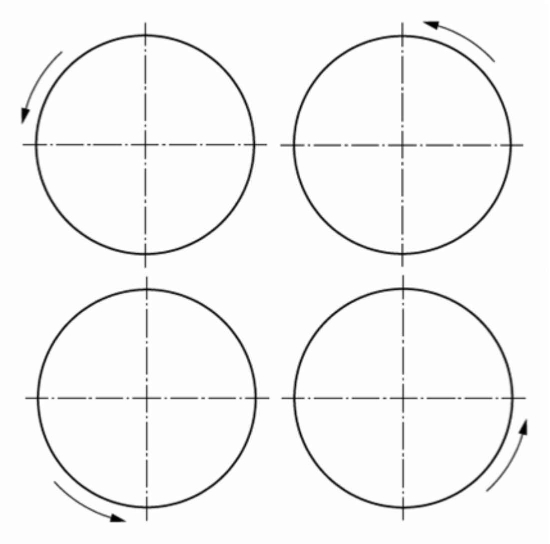


Figure 5.8: Test object and direction of movement (Tracing Test - Abstract)
Source: BS EN ISO 9241-420 (2011)

5.3.3.2 Dependent variables

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (i.e. **design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* for each task, and error rate (*ER*). Error rate is the millimetre of targets selected when the pointer trace outside the target.

5.3.3.3 Dependent measures

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch (summarized in *Table 5.3*).

Dependent measures	Description
Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 5.3: Objective measures of the human performance (Tracing Test - Abstract)

Figure 5.9 show the distance measurement of the test object (each of the circles) and the free-hand-drawn line, in full millimeters, at 36 locations.

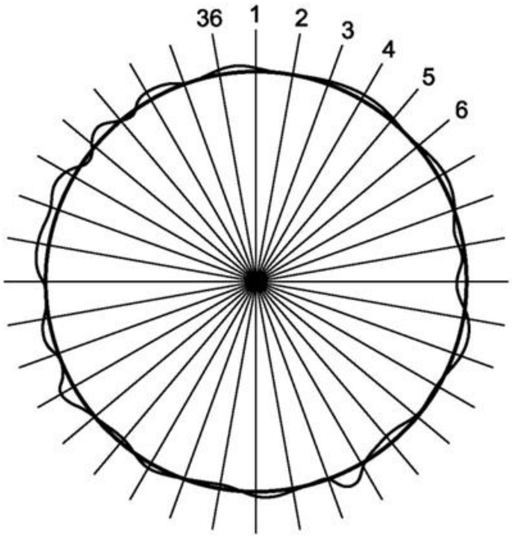


Figure 5.9: Test object and points of measurement for deviations (Tracing Test - Abstract)
Source: BS EN ISO 9241-420 (2011)

As for the participants' feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire shown in *Table 5.4*.

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both test, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force Operation smoothness Operation effort Accuracy Operation speed General comfort Overall operation	
Fatigue Indices	Finger fatigue Wrist fatigue Arm fatigue Shoulder fatigue Finger fatigue	

Table 5.4: Subjective attributes of the Five-point 'Likert-type' scale (Likert, 1932) subjective assessment (Tracing Test – Abstract)

In terms of the user participants' profile, their background information is collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years). Handedness is relevant in this experiment. According to Peters and Murphy (1992), *'how consistently individuals use one hand over the other does have an impact'*. Lyle *et al*. (2012) stated that some individuals consistently use the same hand regardless of task, whereas others switch hands between tasks or between performances of the same task.

5.3.4 RESULT ANALYSIS AND DISCUSSION

5.3.4.1 TEST DATA

5.3.4.1.1 Tracing Test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Since the study employed the touch as the base line, it is possible to compare the difference between the touch, mouse and stylus on the tracing test. The hypothesis H1 is based on the fact that study by Helander *et al.* (1997) suggests that using the mouse on a tracing task is faster rather than using touch and stylus. Therefore, it can be expected that the participant movement time (*MT*) using the mouse will be faster compared to the touch and stylus on the tracing test.

Descriptive statistic (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) shows that the touch data is not normally distributed. Thus, a Friedman test was carried out. On the other hand, the mouse and stylus data showed that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Table 5.5 indicates that the movement time (*MT)* for every task of the touch is significantly faster than the other two devices (mouse and stylus). Moreover, the amount of speed does not show statistically significant difference between all the devices.

Device	N	Task				P value
		1	2	3	4	
Mouse	23	24.91 (13.58)	22.08 (10.84)	21.53 (12.34)	20.29 (10.95)	.005*
Stylus	23	16.82 (8.29)	15.94 (8.89)	15.94 (8.96)	14.41 (9.11)	.790*
Touch	23	11.19 (8.02)	10.69 (8.90)	9.32 (8.06)	9.13 (7.41)	.005**

Table 5.5: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test and Friedman test on every task (Tracing Test - Abstract)

** Repeated Measure ANOVA test*

*** Friedman test*

Furthermore, *Figure 5.10* illustrates that the amount of speed using mouse, stylus and touch has a quite similar tendency of movement time (*MT*), where the amount of speed using the three devices was relatively stable throughout all the tasks. All of the input devices' movement time shows a decrease in task 4. Through observation, task 4 has lower movement time because the task is learnable and participants could already predict the next move.

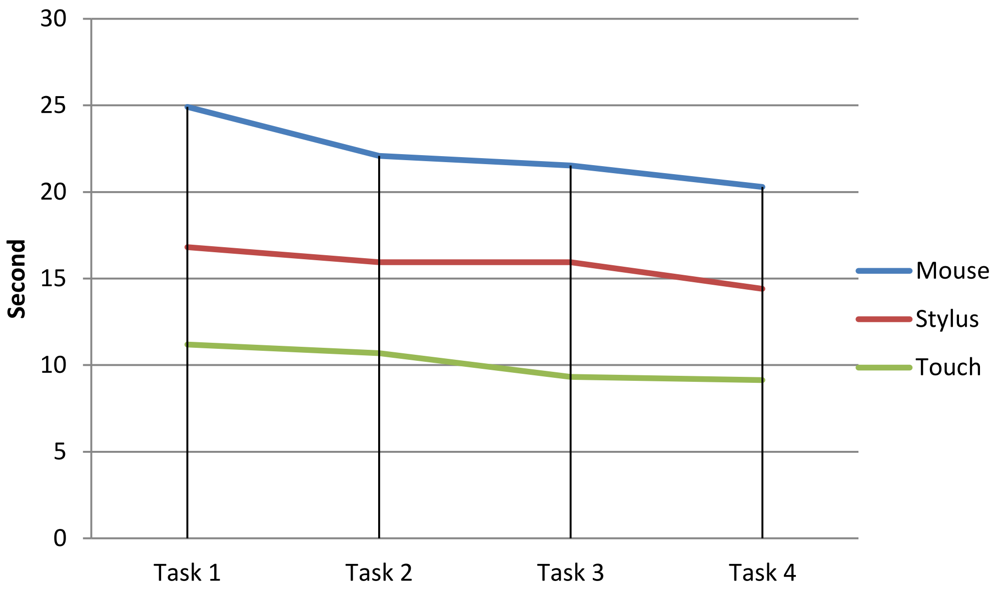


Figure 5.10: The amount of speed based on every task (Tracing Test - Abstract)

In terms of the ***overall movement time,*** a Friedman test was run to determine if there were differences in the amount of speed between 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of speed was statistically significantly different between 3 input devices, $\chi^2(2) = 126.609$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between touch (*Mdn* = 6.88) and stylus (*Mdn* = 13) (*p* < .0005), touch (*Mdn* = 6.88) and mouse (*Mdn* = 18.10) (p < .0005) and stylus (*Mdn* = 13) and mouse (*Mdn* = 18.10) (p < .0005).

Table 5.6 indicates that touch (10.08 sec ±8.50) has the fastest overall movement time (*MT*), that is two-times faster than the mouse (22.20 sec ±11.90) which recorded the slowest, while stylus (15.78 sec ±8.72) has the second fastest overall movement time (*MT*).

Device	N	Mean	Std Deviation
Mouse	23	22.20	11.90
Stylus	23	15.78	8.72
Touch	23	10.08	8.05

Table 5.6: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Tracing Test - Abstract)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 5.7,* the Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (ER) with the touch, mouse and stylus has no statistically significantly difference.

Device	N	Task				P value
		1	2	3	4	
Mouse	23	207.91 (288.98)	174.22 (203.41)	195.87 (247.98)	227.35 (321.94)	.120
Stylus	23	184.48 (261.93)	190.26 (260.25)	187.87 (229.34)	209.04 (239.71)	.009
Touch	23	349.39 (333.70)	322.57 (307.77)	288.65 (293.93)	309.22 (297.64)	.529

Table 5.7: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Friedman test on every task (Tracing Test - Abstract)

Touch, stylus and mouse have different tendency of error rates (*ER*) where mouse shows a decrease in task 2, however continues to increase gradually over the next two tasks. The error rates of participants using stylus was relatively stable during the first three tasks and then shows a slight increase in Task 4. On the other hand, using touch which recorded the highest error rates compared to mouse and stylus on all the tasks, shows a decrease of error rates during the first three tasks and then shows an increase in Task 4.Through observation, participants find the difficulty increases in each level of the task and this leads to high error rates using stylus and mouse. In regards to touch, participant found that in the tracing test using touch is difficult, resulting in the highest error rates compared to other input devices.

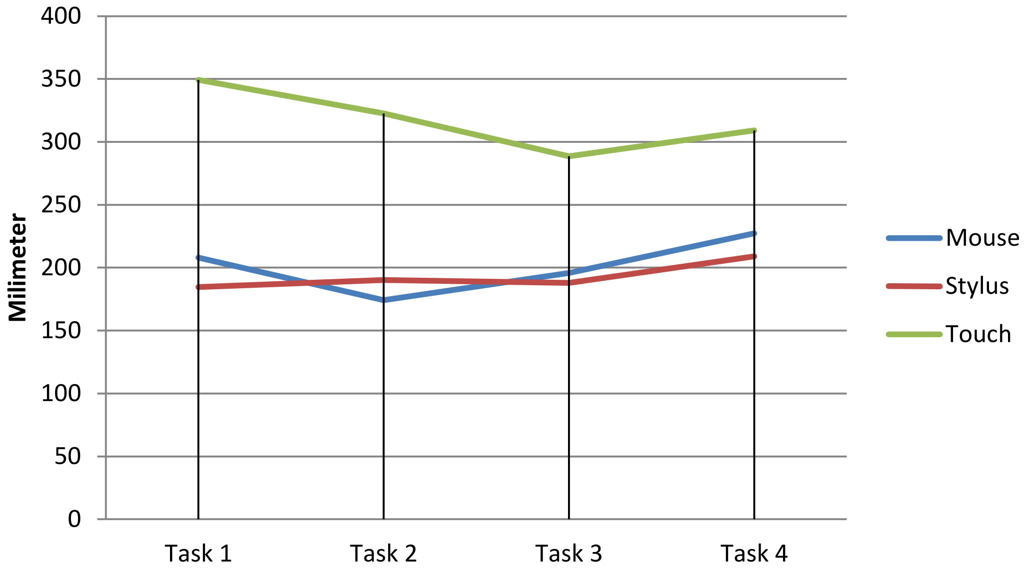


Figure 5.11: Error rates based on every task (Tracing Test - Abstract)

In terms of the ***overall error rate,*** a Friedman test was run to determine if there were differences in the amount of error between 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error was statistically significantly different between 3 input devices, $\chi^2(2) = 97.485$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of error between mouse (*Mdn* = 87) and touch (*Mdn* = 199.5) ($p < .0005$) and stylus (*Mdn* = 91.5) and touch (*Mdn* = 199.5) (p < .0005), but not mouse (*Mdn* = 87) and stylus (*Mdn* = 91.5) (p > .0005).

Table 5.8 indicates that the overall error rate (*ER*) of the touch (317.46mm ±304.327) is almost two-times greater than the stylus (192.91mm ±244.252) which recorded the lowest error among the three input devices, while the mouse (201.34mm ±265.5) shows the second highest error rate (*ER*).

Device	N	Mean	Std Deviation
Mouse	23	201.34	265.5
Stylus	23	192.91	244.25
Touch	23	317.46	304.32

Table 5.8: The effect of the device difference (touch, mouse and stylus) on the overall error rate (Tracing Test - Abstract)

5.3.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input devices based on their experience of using it.

5.3.4.3 TRACING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus interaction.

As can be seen in *Table 5.9*, the inter-reliability test discovered that inter-reliability of the design is very high among all the input devices. Therefore, the comparison can be made for all the input devices in terms of the general indices and the fatigue indices as summarised in *Table 5.10*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.896	.898	.909
Fatigue Indices	.977	.930	.918

Table 5.9: Inter Reliability Statistics with the mouse, stylus and touch (Tracing Test - Abstract)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	4.43	4.17	2.70	.000*
General Indices	Actuation force	23	3.87	3.48	2.87	.000*
	Operation smoothness	23	3.96	3.70	2.61	.000*
	Operation effort	23	3.83	3.61	2.73	.001*
	Accuracy	23	3.77	3.83	2.22	.000*
	Operation speed	23	3.74	3.74	2.96	.007
	General comfort	23	4.17	3.78	2.87	.000*
	Overall operation	23	4.09	3.83	2.74	.000*
Fatigue Indices	Finger fatigue	23	4.13	3.61	2.70	.000*
	Wrist fatigue	23	4.22	3.64	2.87	.000*
	Arm fatigue	23	4.26	3.43	2.74	.000*
	Shoulder fatigue	23	4.22	3.39	3.04	.000*
	Neck fatigue	23	4.35	3.78	2.96	.000*
	Overall operation	23	4.22	3.61	2.91	.000*

Table 5.10: Result analysis of the Five-point Likert scale subjective assessment with the touch, mouse and stylus (Tracing Test - Abstract)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participants considered using the mouse (4.43) on the task given much easier than stylus (4.17) and touch (2.70) ($p<0.05$). This indicates that the mouse device is easier than touch and stylus in the tracing test.

In terms of the general indices, six indicators (i.e. actuation force, operation smoothness, operation effort, operation speed, general comfort and overall operation) are highly rated as being better using mouse, while accuracy indicator recorded that using stylus (3.78) was better. The touch recorded the worst feeling among the participants as the mean shows the lowest among the three input devices on all of the seven indicators.

In regards to the fatigue indices, all indicators (i.e. finger fatigue, wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue, overall operation) show that using the mouse in the tracing test tasks produces a better feeling among the participants compared to stylus and touch, with touch being the least popular.

This led to the conclusion that use of the mouse was clearly the most favoured by participants in the tracing test in relation to general and fatigue indices.

5.4 EXPERIMENT 2: DRAGGING TEST

The second test is the dragging test which evaluating clicking and dragging objects to specific locations, for example; clicking and dragging the pointer down a pull-down menu, and selecting and dragging an object from one window to another. The test object consists of circles with a diameter of 8 mm over a distance of 100 mm and places them in circles with a diameter of 10 mm. The tasks were performed in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.

5.4.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

5.4.1.1 Dragging test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

H2: There is a difference in the amount of error using touch, mouse or stylus.

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus interaction.

5.4.2 EXPERIMENT DESIGN - DRAGGING TEST

The dragging test, on the other hand, aimed to investigate the difference of the three input devices. In order to achieve this, a within users design was deployed. The study assigned twenty-three participants to do the test which had 12 tasks altogether for all of them. The rationale for the design was because the experiment investigates tasks with significant individual differences.

5.4.3 VARIABLES - DRAGGING TEST

5.4.3.1 Independent Variables

The independent variables were:

Factors/Parameters	Level
Target Width	8 mm
Target Distance	100 mm
Target	12 per task
Task	4

Figure 5.12 show the arrangement of dragging test task in which the test object consist of circles with a diameter of 8 mm and over a distance of 100 mm. Participants were tested to place them in circles with a diameter of 10 mm and complete the task in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.

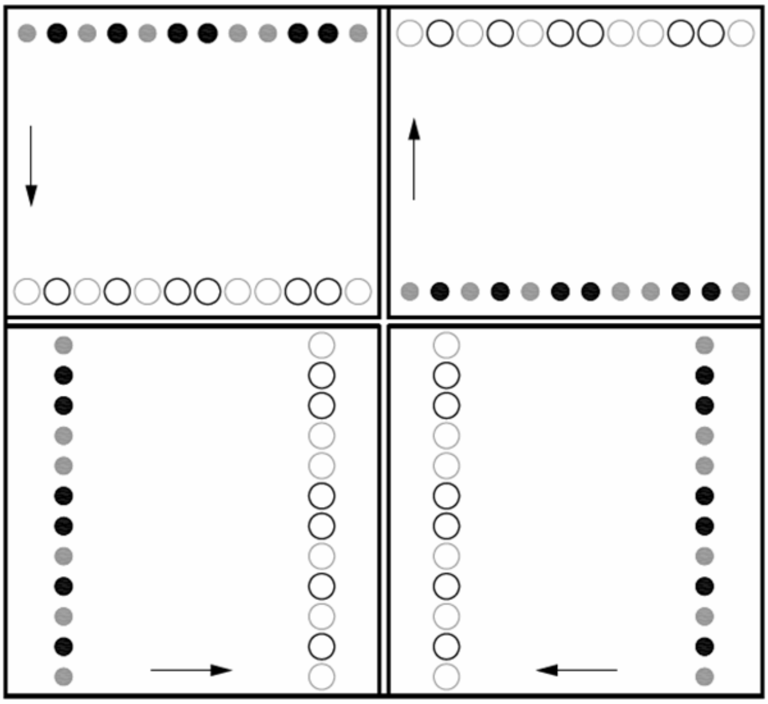


Figure 5.12: Arrangement of the objects for the dragging task for the cardinal directions (Dragging Test - Abstract)
Source: BS EN ISO 9241-420 (2011)

5.4.3.2 Dependent variables

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (**design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* for each task, and error rate (*ER*). Error rate is the millimeter of off target circle that is when it is place away from the perfect hit of the target.

5.4.3.3 Dependent measures

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch (summarized in *Table 5.11*).

Dependent measures	Description
Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 5.11: Objective measures of the human performance (Dragging Test - Context)

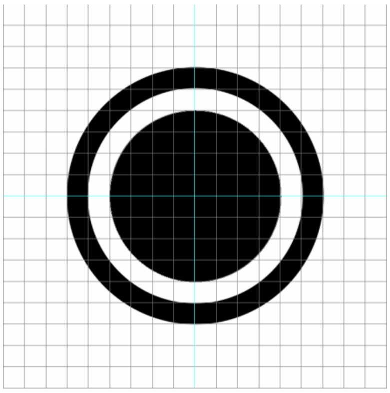


Figure 5.13: Accuracy for the dragging test (Dragging Test - Abstract)

Figure 5.13 shows a perfect hit of a circle placed in the center of the target. The error will be recorded in millimeter (each of the grid scale indicates one mm) if the circle was placed away from the perfect hit of the target.

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire shown in *Table 5.12*:

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both test, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 5.12: Subjective attributes of the Five-point 'Likert-type' scale (Likert, 1932) subjective assessment (Dragging Test - Abstract)

In terms of the user participants profile, their background information is collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

5.4.4 RESULT ANALYSIS AND DISCUSSION

5.4.4.1 TEST DATA

5.4.4.1.1 Dragging Test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistic (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) shows that the touch data is not normally distributed. Thus, the Friedman test was carried out. On the other hand, the mouse and stylus data show that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Table 5.13 indicates the movement time (*MT)* for every task indicated that the stylus is significantly faster than the other two devices (mouse and touch). Moreover the amount of speed of touch and mouse does not show statistically significant difference whereas the stylus indicates statistically significant difference

Devices	N	Task				P value
		1	2	3	4	
Mouse	23	39.47 (15.97)	33.29 (9.10)	32.40 (8.26)	30.36 (8.72)	.048**
Stylus	23	35.13 (14.84)	33.63 (20.90)	31.80 (11.43)	28.83 (6.73)	.002*
Touch	23	43.10 (22)	32.85 (14.31)	36.11 (16.10)	34.18 (17.98)	.022**

Table 5.13: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test and Friedman test on every task (Dragging Test - Abstract)

** Repeated Measure ANOVA test*

*** Friedman test*

Furthermore, *Figure 5.14* illustrates that the amount of speed of mouse and stylus have a similar tendency of movement time (*MT*) where they show a gradual decrease throughout the tasks until they reached their lowest point in Task 4. On the other hand, mouse shows a fluctuation trend of the movement time throughout the task. Through observation, participants show faster movement time in Task 4 because it is learnable and participants could predict their next move.

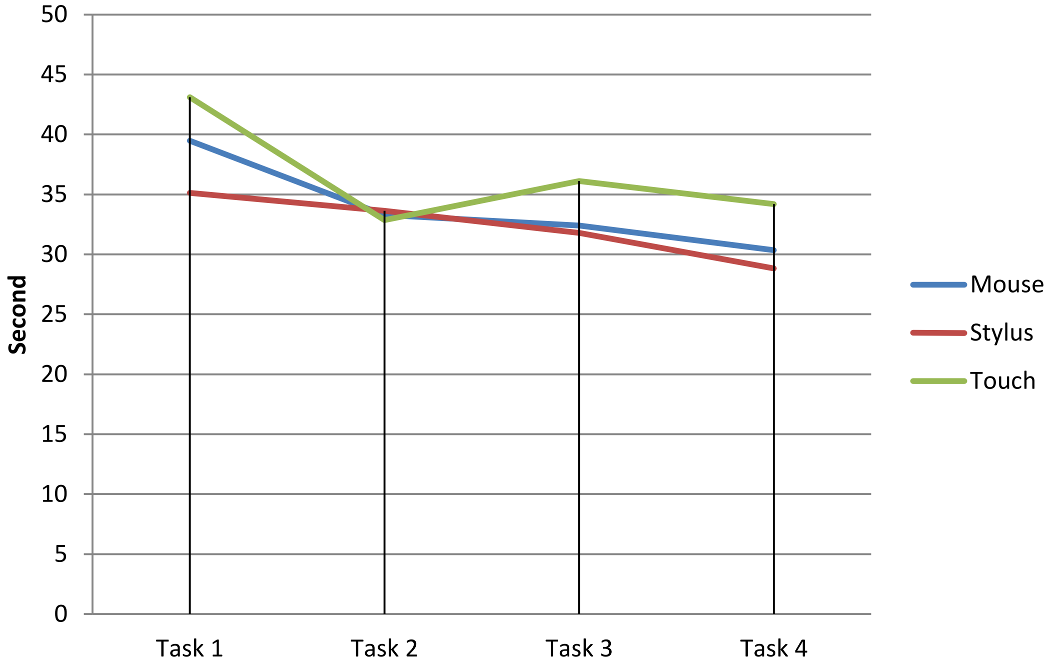


Figure 5.14: The amount of speed based on every task (Dragging Test - Abstract)

In terms of the ***overall movement time,*** a Friedman test was run to determine if there were differences in the amount of speed between 3 inputs devices. The amount of speed decreased between mouse (*Mdn* = 31.17), to touch (*Mdn* = 30.92), to stylus (*Mdn* = 29.07), but the differences were not statistically significant, $\chi^2(2) = 6.447$, $p > .040$

As can be seen in *Table 5.14*, the mean and standard deviation indicate that the slowest overall movement time (*MT)* is the touch (36.56 sec ±17.96), followed by the mouse (33.65 sec ±11.12) and then stylus (32.25 sec ±14.45) which recorded the fastest among the devices.

Device	N	Mean	Std Deviation
Mouse	23	33.65	11.12
Stylus	23	32.25	14.45
Touch	23	36.56	17.96

Table 5.14: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Dragging Test - Abstract)

H2: There is a difference in the amount of error using touch, mouse or stylus.

As can be seen in *Table 5.15*, the Repeated Measure ANOVA and Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (*ER*) with the touch, mouse and stylus has no statistically significantly difference.

Devices	N	Task				P value
		1	2	3	4	
Mouse	23	15.91 (6.93)	18 (9.18)	18.4 (7.18)	19.61 (6.95)	.061**
Stylus	23	27.04 (10.65)	24.91 (10.10)	24.74 (12.33)	24.09 (10.97)	.071*
Touch	23	54.09 (16.42)	49.52 (13.27)	55.48 (17.86)	54.48 (14.54)	.070*

Table 5.15: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Repeated Measure ANOVA test and Friedman test on every task (Dragging Test - Abstract)

** Repeated Measure ANOVA test*

*** Friedman test*

Figure 5.15 illustrate that the error rates of stylus and mouse have quite similar tendency where the error rates remained fairly stable throughout the tasks. On the other hand, touch which has the highest error rates on all of the tasks shows fluctuation from task 1 to task 4. Through observation, participants show higher error rates using touch because participant found that dragging using touch is difficult, hence resulting in the highest error rates compared to other input devices.

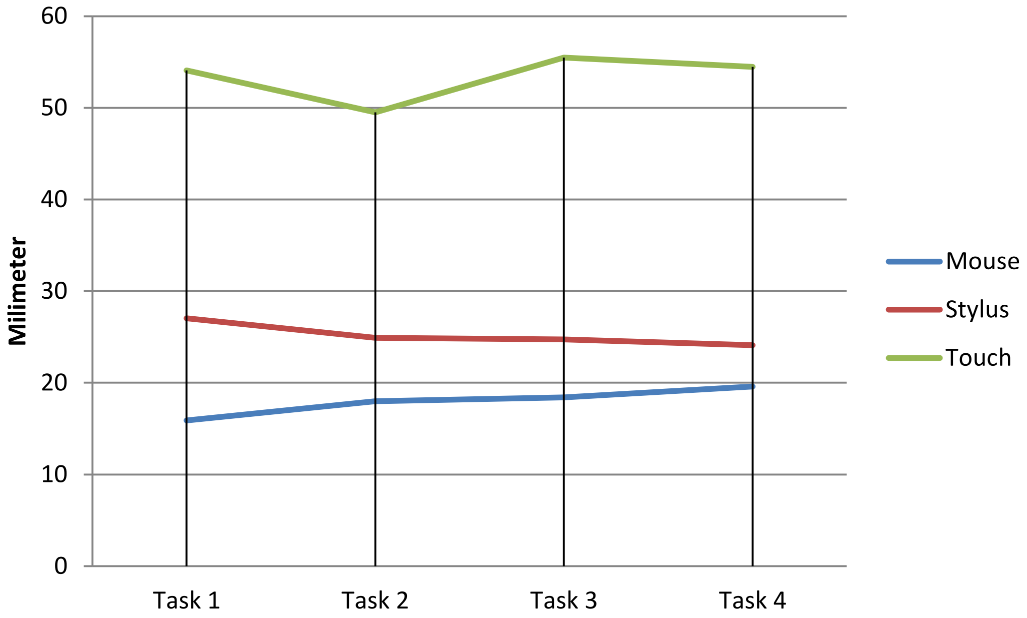


Figure 5.15: The error rates based on every task (Dragging Test - Abstract)

A Friedman test was run to determine if there were differences in the amount of error between 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error was statistically significantly different between 3 input devices, $\chi^2(2) = 146.383$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of error between mouse (*Mdn* = 16.5) and stylus (*Mdn* = 23) ($p < .0005$) mouse (*Mdn* = 16.5) and touch (*Mdn* = 54) (p < .0005) and stylus (*Mdn* = 23) and touch (*Mdn* = 54) (p < .0005)

Device	N	Mean	Std Deviation
Mouse	23	17.89	7.60
Stylus	23	25.20	10.90
Touch	23	53.39	15.53

Table 5.16: The effect of the device difference (touch, mouse and stylus) on the overall error rate (Dragging Test - Abstract)

The mean and standard deviation that has been shown in *Table 5.16* points out that the overall error rate (*ER*) of the touch (53.39 mm ±15.32) is almost two-times higher than the mouse (17.89 mm ±7.60) which recorded the lowest error, while the stylus (25.20 mm ±10.90) shows the second highest error rate (*ER*).

5.4.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input devices based on their experience of using it.

5.4.4.2.1 DRAGGING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus interaction.

As can be seen in *Table 5.17*, the inter-reliability test discovered that inter-reliability of the design is very high among all input devices. Thus, the comparison can be made for all the input devices in terms of the general indices and the fatigue indices as summarised in *Table 5.18*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.965	.950	.969
Fatigue Indices	.971	.959	.968

Table 5.17: Inter Reliability Statistics with the mouse. Stylus and touch (Dragging Test - Abstract)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	4.52	4.09	3.04	.000*
General Indices	Actuation force	23	4.35	4.05	3.13	.000*
	Operation smoothness	23	4.26	4.00	3.04	.000*
	Operation effort	23	4.35	4.04	2.87	.000*
	Accuracy	23	4.17	3.83	2.52	.000*
	Operation speed	23	4.17	4.00	3.09	.000*
	General comfort	23	4.39	4.00	2.96	.000*
	Overall operation	23	4.26	4.04	3.04	.000*
Fatigue Indices	Finger fatigue	23	4.17	3.70	3.09	.001*
	Wrist fatigue	23	4.09	3.70	3.30	.006
	Arm fatigue	23	4.17	3.70	3.09	.000*
	Shoulder fatigue	23	4.04	3.91	3.30	.011
	Neck fatigue	23	4.04	4.00	3.57	.091
	Overall operation	23	4.22	3.83	3.22	.000*

Table 5.18: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the I Friedman test on the raw data of the subjective assessment (Dragging Test - Abstract)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participant considered using the mouse (4.52) on the task given much easier than stylus (4.09) and touch (3.04) ($p<0.05$). This indicates that the mouse device is easier than touch and stylus in the dragging test.

In terms of the general indices, all seven indicators (i.e. actuation force, operation smoothness, operation effort, accuracy, operation speed, general comfort and overall operation) are highly rated as being better using mouse. The touch recorded the worst feeling among the participants as the mean show the lowest among the three input devices on all of the seven indicators.

In regards to the fatigue indices, all indicators (i.e. finger fatigue, wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue, overall operation) show that using the mouse in the dragging test tasks created a better feeling among the participants compared to stylus and touch, with touch being the least favoured.

This led to the conclusion that the mouse was clearly the most popular in the dragging test in relation to general and fatigue indices.

5.5 EXPERIMENT 3: ONE DIRECTION TAPPING

The third test is the one direction tapping test in which the evaluating point is the movement along one axis on a horizontal rubber banding, an insert cursor at points along a character string and selecting information in columns or rows. The test object consists of two rectangles with a defined width in the direction perpendicular to the direction of the movement. The task consists of alternately tapping between the two rectangles.

5.5.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency of use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

5.5.1.1 One-direction tapping test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

H2: There is a difference in the amount of error using touch, mouse or stylus.

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

5.5.2 EXPERIMENT DESIGN - ONE DIRECTION TAPPING TEST

The one direction tapping test aims to find out the difference of three input devices. In this test, the same participants engage in all conditions. As a result, the appropriate design for the test is a within users design. In this test, 23 participants were assigned to carry it out and it included twelve tasks altogether. There were various rationales behind the choice of this design. It was first deployed because it helps in equating groups in an experiment. Consequently, every participant in this test was in a group and this avoided any significant differences since they were the same people (Langston, 2014). The design was also chosen to enhance efficiency because it substantially minimizes the number of participants required. This is due to the fact that such a design only needs a smaller sample size and enhances the quality (Langston, 2014: Bannan-Ritland, 2003). Moreover, Lazar *et al.* (2010) argue that the within-in group design is more appropriate for researchers who have difficulty in finding as well as recruiting qualified participants which is the frequent problem faced by many HCI researchers.

5.5.3 VARIABLES - ONE DIRECTION TAPPING TEST

5.5.3.1 Independent Variables

The independent variables were:

- Target Width (3 mm, 5 mm, 9 mm, 12 mm)
- Target Distance (30 mm, 80 mm, 135 mm, 170 mm)
- Tapping task (1 to 50 per task)
- Task (1 to 4)

In the one-direction case, a task consists of 50 tapping of the same width-distance combination. A total of 600 tappings were run (50 tappings per task × [4 widths & distance] x 3 inputs).

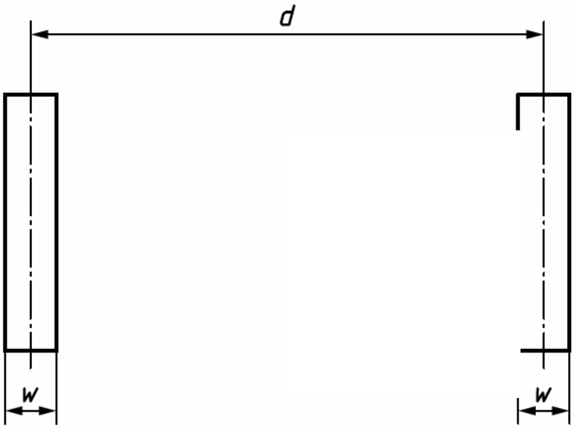


Figure 5.16: One direction tapping task (One Direction Tapping Test - Abstract)
Source: BS EN ISO 9241-420 (2011)

Key

d	target distance
w	target width

5.5.3.2 *Dependent variables*

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (i.e. **design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* for each task, and error rate (*ER*). Error rate will be recorded if the participant taps outside the target object.

5.5.3.3 *Dependent measures*

The dependent measures were that the performance was measured by examining movement time (*MT*) of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse stylus and touch (summarized in *Table 5.19*).

Dependent measures	Description
Error Rate	Error attempt is recorded (Tap away from the target)
Movement Time	When the cursor enters the target, it will be counted.

Table 5.19: Objective measures of the human performance (One Direction Tapping Test - Abstract)

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire, shown in *Table 5.20*:

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both test, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 5.20: Subjective attributes of the Five-point Likert scale subjective assessment (One Direction Tapping Test - Abstract)

In terms of the user participants profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

5.5.4 RESULT ANALYSIS AND DISCUSSION

5.5.4.1 TEST DATA

5.5.4.1.1 ONE DIRECTION TAPPING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistics (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) show that data from all the input devices shows that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

General observation shows that the movement time (*MT*) for every task using touch is significantly faster than the other two devices (mouse and stylus). Moreover, the amount of speed shows a statistically significant difference in the use of mouse and stylus but not for touch.

Devices	N	Task				P value
		1	2	3	4	
Mouse	22	51.13 (12.46)	50.05 (6.09)	55.34 (7.30)	63.85 (10.28)	.000
Stylus	22	52.69 (7.61)	49.27 (6.02)	55.27 (8.51)	70.71 (11.15)	.000
Touch	22	41.19 (6.90)	41.84 (12.01)	46.50 (9.08)	55.59 (12.39)	.456

Table 5.21: The effect of the device difference (touch based, mouse and stylus) on the movement time based on the Repeated Measure ANOVA on every task (One Direction Tapping Test - Abstract)

Table 5.21 illustrates that the amount of speed on each task using mouse, stylus and touch has a similar tendency of movement time (*MT*). During the first two tasks, movement time of all the input devices remained fairly alike. There was a significant increase in the movement time tendency of participants taking the test between Task 3 and Task 4. The movement time of participants using the stylus was at its slowest movement time in Task 4 with a mean of 70.71 seconds. Mouse reached a lowest point of speed of 63.85 seconds in Task 4 while touch 55.59 seconds. Observation showed that participants used slower movements as the difficulty of the tasks increased.

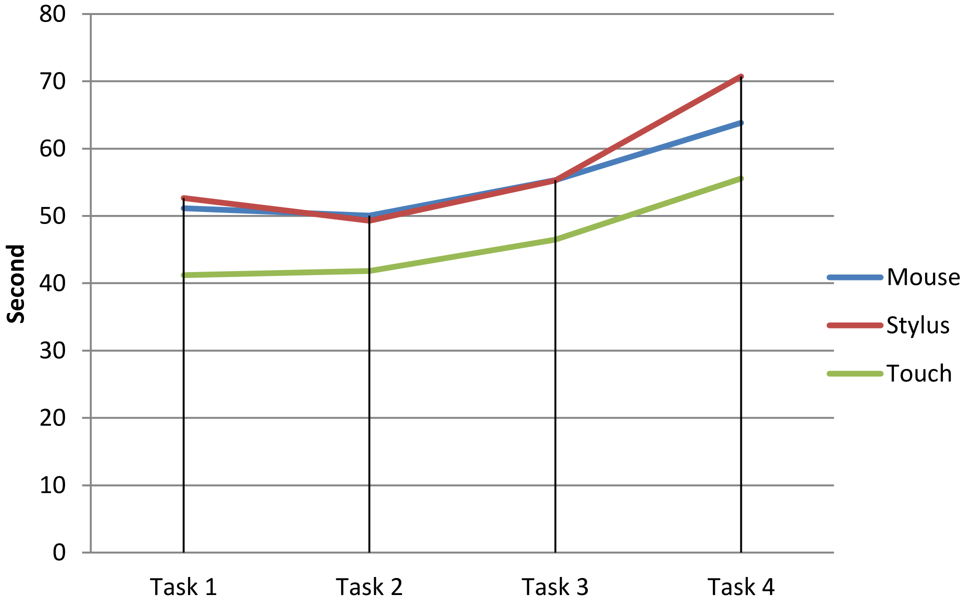


Figure 5.17: The amount of speed based on every task (One Direction Tapping Test - Abstract)

In terms of the overall movement time, a Friedman test was run to determine if there were differences in the amount of speed between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of speed was statistically significantly different between 3 the input devices, $\chi2(2) = 84.636$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between touch (Mdn = 44.79) and stylus (Mdn = 53.83) ($p < .0005$) and touch (Mdn = 44.79) and mouse (Mdn = 54.52) ($p < .0005$), but not stylus (Mdn = 53.83) and mouse (Mdn = 54.52) ($p > .0005$).

Table 5.22 indicate that touch (46.28 sec ±10.18) has the fastest overall movement time (*MT*), that is two-times faster than the stylus (56.98 sec ±17.61) which recorded the slowest, while mouse (55.09 sec ±10.70) has the second fastest overall movement time (*MT*).

Device	N	Mean	Std Deviation
Mouse	22	55.09	10.70
Stylus	22	56.98	17.61
Touch	22	46.28	10.18

Table 5.22: The effect of the device difference (touch, mouse and stylus) on the overall movement time (One Direction Tapping Test - Abstract)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 5.23*, the Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (ER) with the touch, mouse and stylus has statistically significantly differences.

Device	N	Task				P value
		1	2	3	4	
Mouse	22	2.32 (6.90)	4.32 (12.01)	3.91 (9.08)	5.82 (12.36)	.000
Stylus	22	2.05 (6.66)	2.32 (8.68)	3.23 (8.37)	7.45 (10.62)	.000
Touch	22	3.05 (9.51)	2.82 (9.70)	5.59 (14.05)	13.14 (17.26)	.000

Table 5.23: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Friedman test on every task (One Direction Tapping Test - Abstract)

Touch and mouse have a similar tendency of error rates (*ER*) where it fluctuated for tasks 1, 2 and 3, while stylus was relatively stable during the first two tasks and continued to rise steadily over the next task. The error rate of participants in Task 4 shot up dramatically for all the input devices (*see Figure 5.18*). Through observation, participants found Task 4 difficult and this led to high error rates.

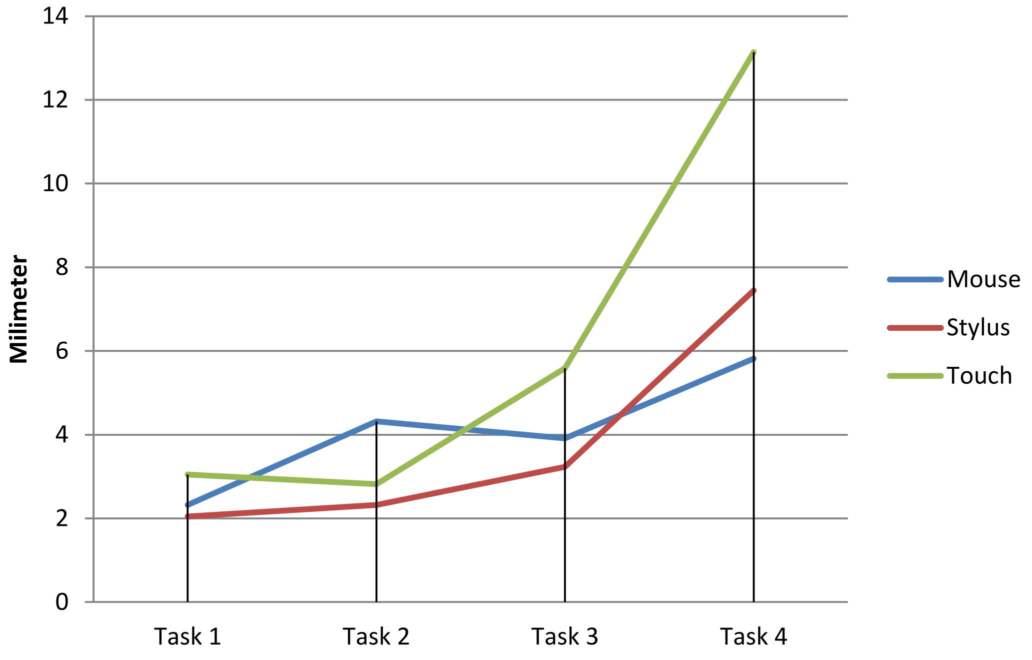


Figure 5.18: The error rates based on every task (One Direction Tapping Test - Abstract)

In terms of the ***overall error rate,*** a Friedman test was run to determine if there were differences in the amount of error between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error was statistically significantly different between the 3 input devices, $\chi^2(2) = 28.962$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between stylus (*Mdn* = 1) and touch (*Mdn* = 2) ($p < .0005$) and mouse (*Mdn* = 1) and touch (*Mdn* = 2) (p < .0005), but not stylus (*Mdn* = 1) and mouse (*Mdn* = 1) (p > .0005).

Table 5.24 indicates that the overall error rate (*ER*) of the touch (6.15 ±13.48) was almost two-times greater than the stylus (3.76 ±8.82) which recorded the lowest errors among the three input devices, while the mouse (4.09 ±10.24) showed the second highest error rate (*ER*).

Device	N	Mean	Std Deviation
Mouse	22	4.09	10.24
Stylus	22	3.76	8.82
Touch	22	6.15	13.48

Table 5.24: The effect of the device difference (touch, mouse and stylus) on the overall error rate (One Direction Tapping Test - Abstract)

5.5.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input device based on their experience of using it.

5.5.4.2.1 ONE DIRECTION TAPPING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 5.25*, the inter-reliability test discovered that inter-reliability about the design is very high with all the input devices, thus the comparison can be made for all devices in terms of the general indices and the fatigue indices, as summarised in *Table 5.26*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.942	.923	.939
Fatigue Indices	.812	.891	.906

Table 5.25: Inter Reliability Statistics with Mouse, Stylus and Touch (One Direction Tapping Test - Abstract)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	22	4.18	3.95	4.18	.667
General Indices	Actuation force	22	3.82	3.68	4.14	.168
	Operation smoothness	22	3.68	3.86	4.27	.084
	Operation effort	22	3.68	3.73	4.23	.240
	Accuracy	22	3.64	3.77	3.95	.219
	Operation speed	22	3.41	3.86	4.32	.055
	General comfort	22	3.86	3.59	4.05	.589
	Overall operation	22	4.05	3.95	4.14	.767
Fatigue Indices	Finger fatigue	22	3.50	3.52	3.64	.620
	Wrist fatigue	22	3.36	3.33	3.64	.430
	Arm fatigue	22	3.50	3.19	3.55	.193
	Shoulder fatigue	22	3.45	3.24	3.45	.911
	Neck fatigue	22	3.68	3.52	3.64	.529
	Overall operation	22	3.82	3.62	3.82	.397

Table 5.26: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the Friedman test on the raw data of the subjective assessment (One Direction Tapping Test - Abstract)

The Friedman test indicates that the participants considered much it easier using the touch (4.18) and mouse (4.18) on the task given rather than the stylus (3.95) ($p>0.05$). This indicates that the mouse and touch are easier than the stylus in the one direction tapping test.

In terms of the general indices, touch was highly rated on all seven indicators (i.e. actuation force, operation smoothness, operation effort, accuracy, operation speed, general comfort and overall operation). . The stylus and mouse recorded mixed feeling among the participants as the mean shows that for three indicators (i.e. actuation force, general comfort and overall operation) the stylus was rated worst while for four indicators (i.e. operation smoothness, operation effort, accuracy, operation speed) the mouse came out worst.

With regards to the fatigue indices, for five indicators (i.e. finger fatigue, wrist fatigue, arm fatigue, shoulder fatigue, overall operation) using touch in the one direction tapping test produced a high level of satisfaction among participants. In addition, mouse and touch shared a good rating in two of the indicators (shoulder fatigue and overall operation). Mouse also scored well among participants, compared to stylus and touch in terms of neck fatigue. Stylus was rated to be the worst input device in five of the indicators (i.e. wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue, overall operation), while the mouse was the worst rated in the finger fatigue indicator.

This led to the conclusion that touch produced the best rating overall in the one direction tapping test in relation to general and fatigue indices.

5.6 EXPERIMENT 4: MULTI-DIRECTIONAL TAPPING

The fourth test is the multi-direction tapping test in which the evaluating points are the movements in many different directions on the repositioning of a pointer at different areas on the screen, cell selection in a spreadsheet and selecting randomly located icons. The test object consists of targets positioned around the circumference of a circle. The targets are to be arranged so that the movements are nearly equal to the diameter of the circle. The box to which the targets should advance is marked with X. Each test session starts after the subject points to the topmost target and ends when the sequence is completed (at the topmost target). The tests were conducted with a range of difficulties in the size and the distance of the target squares.

5.6.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency of use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

5.6.1.1 Multi-directional tapping test

H1: There is a difference in the amount of speed using touch, mouse or stylus.

H2: There is a difference in the amount of error using touch, mouse or stylus.

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

5.6.2 EXPERIMENT DESIGN – MULTI- DIRECTIONAL TAPPING TEST

The multi-directional tapping test also aimed to find out the difference of three input devices. The assumption in this test is that the same participants take part in all conditions. In this regard, the test deployed a within users study design whereby 23 participants were given the test which had 9 tasks (3 inputs x 3 task) altogether.

Similarly, since the test is to investigate the difference of three input devices and the same participants take part in all conditions, a within users design was deployed. Twenty-three participants were assigned to do the test which included 9 tasks (3 inputs x 3 task) altogether.

5.6.3 VARIABLES – MULTI- DIRECTIONAL TAPPING TEST

5.6.3.1 Independent Variables

- Target Width (4mm, 5mm, 6 mm)
- Target Distance (6mm, 7mm, 8.5 mm)
- Target Angle (14.4°, 28.8°, 43.2°, 57.6°, 72°, 86.4°, 100.8°, 115.2°, 129.6°, 144°, 158.4°, 172.8°, 187.2°, 201.6°, 216°, 230°, 244.8°, 259.2°, 273.6°, 288°, 302.4°, 316.8°, 331.2°, 345.6°, 360°)
- Tapping task (1 to 25 per task)
- Task (1 to 3)

Task of the multi-directional task is defined as the 25 fully crossed combinations of target distance, width and angular location from the starting position (3 widths & distances × 25 angles × 3 inputs). Moreover, a total of 225 tappings were run (25 tapping per task × [3 widths & distance] x 3 inputs).

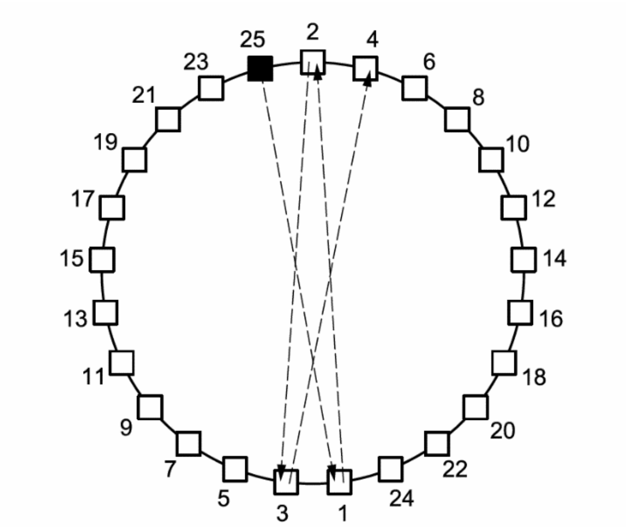


Figure 5.19: Multi-directional tapping test (Multi Direction Tapping Test - Abstract)
Source: BS EN ISO 9241-420 (2011)

5.6.3.2 Dependent variables

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (**design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* for each task, and error rate (*ER*). Error rate will be recorded if the participant taps outside the target object.

5.6.3.3 Dependent measures

The dependent measures were that the performance was measured by examining movement time (*MT*) of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch (summarized in *Table 5.27*).

Dependent measures	Description
Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 5.27: Objective measures of the human performance (Multi Direction Tapping Test - Abstract)

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire, shown in *Table 5.28*:

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both tests, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 5.28: Subjective attributes of the Five-point Likert scale subjective assessment (Multi Direction Tapping Test - Abstract)

In terms of the user participants profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

5.6.4 RESULT ANALYSIS AND DISCUSSION

5.6.4.1 TEST DATA

5.6.4.1.1 MULTI-DIRECTIONAL TAPPING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistics (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) show that the all the input devices data suggest that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Generally, the movement time (*MT*) for every task indicated that touch is faster than the other two devices (mouse and stylus). Moreover the amount of speed of all devices does not show a statistically significant difference (*see Table 5.29*).

Devices	N	Task			P value
		1	2	3	
Mouse	22	27.62 (3.32)	24.94 (2.81)	24.94 (2.86)	.118
Stylus	22	23.87 (4.23)	21.67 (3.61)	22.20 (2.48)	.391
Touch	22	21.17 (3.38)	21.47 (4.11)	21.97 (4.18)	.562

Table 5.29: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test on every task (Multi Direction Tapping Test - Abstract)

Figure 5.20 illustrates that the amount of speed of stylus and touch have a similar tendency of movement time (*MT*) where it decreased in Task 2, while remaining relatively steady in Task 3. On the other hand, touch showed a relatively stable movement time (*MT*) throughout the tasks. The movement time of participants using stylus and mouse was at its fastest movement time in Task 2 and 3. Touch showed the fastest movement time compared to the other input devices in all of the tasks. Through observation, participants show faster movement times using stylus and mouse in Tasks 2 and 3 because it is learnable and participants could predict their next move.

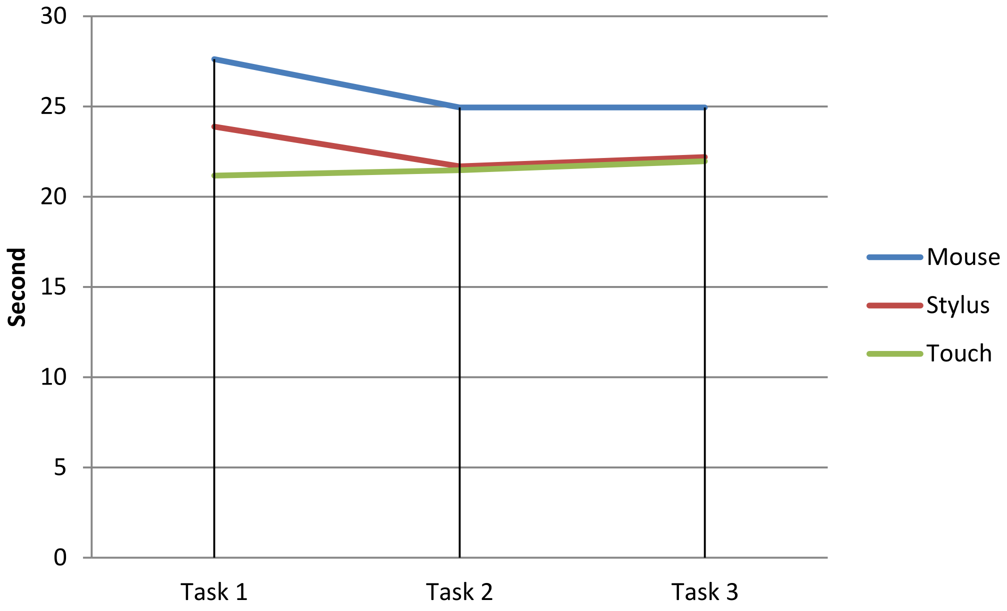


Figure 5.20: The amount of speed based on every task (Multi Direction Tapping Test - Abstract)

In terms of the ***overall movement time,*** a one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences in the amount of speed over the three types of input: mouse, stylus and touch. There were no outliers and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. Mauchly's test of sphericity indicated that the assumption of sphericity had not been violated, $\chi2(2) = 1.003$, $p = .606$.

The three type input (Mouse, Stylus, Touch) elicited statistically significant difference in amount of speed, $F(2, 130) = 32.271$, $p < .0005$, partial $\eta2 = .332$, with amount of speed decreasing from 25.83 ± 3.22 seconds for mouse input to 22.56 ± 3.59 seconds for stylus input and to 21.55 ± 3.86 seconds for touch input. Post hoc analysis with a Bonferroni adjustment revealed that amount of speed was statistically significantly decreased from mouse input to stylus input (3.26 (95% CI, 1.97 to 4.54) seconds, $p < .0005$), and from mouse input to touch input (4.30 (95% CI, 2.88 to 5.72) seconds, $p = .001$), but not from stylus input to touch input (1.04 (95% CI, 0.37 to 2.46) seconds, $p = .054$).

As can be seen in *Table 5.30*, the mean and standard deviation indicates that the slowest overall movement time (*MT)* is the mouse (25.83 sec ±3.22), followed by the stylus

(22.58 sec ±3.59) and then touch (21.53 sec ±3.86) which recorded the fastest among the devices.

Device	N	Mean	Std Deviation
Mouse	22	25.83	3.22
Stylus	22	22.58	3.59
Touch	22	21.53	3.86

Table 5.30: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Multi Direction Tapping Test - Abstract)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 5.31*, the Repeated Measure ANOVA and Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (*ER*) with the touch, mouse and stylus has no statistically significantly difference.

Devices	N	Task			P value
		1	2	3	
Mouse	22	.50 (1.06)	.27 (.88)	.45 (.91)	.862**
Stylus	22	.36 (.79)	.59 (1.14)	1.14 (1.13)	.006**
Touch	22	2.36 (1.92)	2.45 (2.39)	3.91 (4.31)	.562*

Table 5.31: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Repeated Measure ANOVA test and Friedman test on every task (Multi Direction Tapping Test - Abstract)

* Repeated Measure ANOVA test

** Friedman test

Touch and Stylus have a similar tendency of error rates (*ER*) where it gradually increased in Task 2 and continue to show a dramatic rise in Task 3 with touch (3.91) having the highest error rates followed by stylus (1.14). Mouse shows the least error rates in Task 2 (0.27) where it reached its lowest point and went slightly up again in Task 3 (0.45). Through observation, participants show higher error rates using touch because the target becomes smaller towards the end of the tasks and caused the participant to tap out of the target especially if the target view was block by the participant's finger. Stylus and Mouse show minimal error because the stylus has the tip point and the mouse has the pointer which make it easier to select and led to precise performance as mouse and stylus only block to a lesser amount the participant's view of the target.

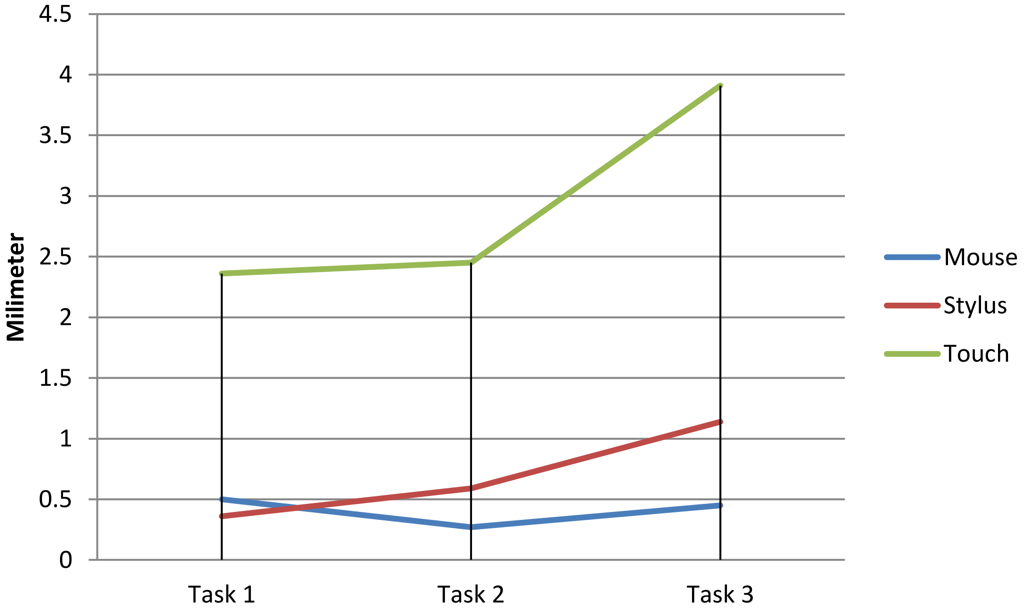


Figure 5.21: The error rates based on every task (Multi Direction Tapping Test - Abstract)

A Friedman test was run to determine if there were differences in the amount of error between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error was statistically significantly different between the 3 input devices, $\chi^2(2) = 72.492$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between mouse (Mdn = 0) and touch (Mdn = 2) (p < .0005) and stylus (Mdn = 0) and touch (Mdn = 2) (p < .0005), but not mouse (Mdn = 0) and stylus (Mdn = 0) (p > .0005).

Device	N	Mean	Std Deviation
Mouse	22	.41	.94
Stylus	22	.70	1.07
Touch	22	2.91	3.09

Table 5.32: The effect of the device difference (touch, mouse and stylus) on the overall error rate (Multi Direction Tapping Test - Abstract)

The mean and standard deviation that has been shown in *Table 5.32* points out that the overall error rate (*ER*) of the touch (2.91 ±3.09) is almost five-times higher than the mouse (.41 ±.94) which recorded the lowest error, while the stylus (.70 ±1.07) shows the second highest error rate (*ER*).

5.6.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input device based on their experience of using it.

5.6.4.2.1 MULTI DIRECTIONAL TAPPING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 5.33*, the inter-reliability test discovered that inter-reliability about the design is very high with the all input devices, thus the comparison can be made for both devices in terms of the general indices and the fatigue indices, as summarised in *Table 5.34*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.972	.952	.972
Fatigue Indices	.947	.915	.940

Table 5.33: Inter Reliability Statistics with Mouse, Stylus and Touch (Multi Direction Tapping Test - Abstract)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	22	3.91	3.86	4.00	.589
General Indices	Actuation force	22	4.05	3.95	3.86	.934
	Operation smoothness	22	4.14	3.95	3.95	.766
	Operation effort	22	4.05	4.00	3.86	.979
	Accuracy	22	4.18	3.73	3.77	.180
	Operation speed	22	4.00	4.05	4.14	.744
	General comfort	22	4.27	3.91	4.00	.673
	Overall operation	22	4.27	4.14	4.18	.739
Fatigue Indices	Finger fatigue	22	3.91	3.91	3.82	.607
	Wrist fatigue	22	3.82	3.77	3.86	.836
	Arm fatigue	22	3.91	3.50	3.73	.299
	Shoulder fatigue	22	3.82	3.41	3.68	.323
	Neck fatigue	22	4.00	3.73	3.77	.521
	Overall operation	22	4.00	3.95	3.73	.689

Table 5.34: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the Friedman test on the raw data of the subjective assessment (Multi Direction Tapping Test - Abstract)

The Friedman test indicates that the participants considered it much easier using touch (4.00) in the tasks given rather than stylus (3.86) and mouse (3.91) (p>0.05). This indicates that touch will be easier than the mouse and stylus in the multi directional tapping test.

In terms of the general indices, for six indicators (i.e. actuation force, operation smoothness, operation effort, accuracy, general comfort and overall operation) the mouse was highly rated, while the operation speed indicator recorded that touch was better (4.14). The stylus and touch recorded mixed feeling among the participants: the mean shows that for four indicators (i.e. operation smoothness, accuracy, general comfort and overall operation) the stylus was rated worst, while for three indicators (i.e. actuation force, operation smoothness, operation effort), touch was worst.

With regard to the fatigue indices, for five indicators (i.e. finger fatigue, arm fatigue, shoulder fatigue, neck fatigue, and overall operation) in the multi directional tapping test participants favoured the mouse. In addition, the mouse and stylus share a better rating in the finger fatigue indicator. For avoiding wrist fatigue, touch is preferable.

The stylus comes out worst in four of the indicators (i.e. wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue), while touch is the worst rated in finger fatigue and overall fatigue indicators.

This led to the conclusion that the mouse was rated highest among participants in the multi directional tapping test in relation to general and fatigue indices.

5.7 SUMMARY

5.7.1 Tracing and Dragging Test

Based on the result analysis, the tracing test shows that although touch has the fastest Movement Time (*MT*) mean with 10.083 seconds, however, it has the greatest Error Rate (*ER*) mean with 317.46 mm compared to stylus and touch.

On the other hand, the dragging test shows that the stylus has the fastest input with 32.249 seconds Movement Time (*MT*) mean if compared to the mouse and touch. Yet again, Touch has the highest Error Rate (*ER*) mean with 53.39 mm among the three input devices.

The reason for the error as specified by the participants feedback in terms of touch was related to the size of fingers (as for example the finger could cover the tracing object thereby making it difficult to tell the end point), an unresponsive touch screen and the fact that using a finger is difficult to do the task accurately.

In regards to satisfaction levels of the tracing test and dragging test results using the Friedman test revealed that overall there is a statistically significant difference in satisfaction level between mouse, stylus and touch inputs.

As regards to all the General indices and Fatigue indices findings, using the mouse is the highest rated and is, therefore the most preferable input device compared to stylus and touch, while touch is the least preferable.

5.7.2 One Direction Tapping and Multi-Directional Tapping Tests

Based on the result analysis, the one direction tapping test indicates that touch (46.28 ±10.18) has the fastest overall movement time (*MT*). Although touch has the fastest overall movement time (*MT*), it also recorded the highest overall error rate (*ER*) (6.15 ±13.48).

Similarly, the multi-directional tapping test shows that touch has also the fastest input with 21.53 seconds Movement Time (*MT*) mean if compared to the mouse and stylus. Yet again, touch has the highest Error Rate (*ER*) mean with 2.91 among the three input devices.

The reason for the error as specified by the participants' feedback in terms of touch was related to the size of fingers (as, for example, the finger could cover the tracing object that makes it difficult to tell the end point) and the fact that using fingers makes it difficult to do the task accurately.

Furthermore, participants showed higher error rates using touch because the target became smaller towards the end of the tasks and caused the participants to tap out of the target especially if the target view was blocked by the participants' finger. The stylus and mouse show minimal error because the stylus has the tip point and the mouse has the pointer which make it easier to select and led to a more precise performance as the mouse and stylus only block to a lesser amount the participants' view of the target.

In relation to participants' satisfaction level between the touch, mouse and stylus for general and fatigue indices, touch was rated the best in the one direction tapping test while the mouse came out best in the multi directional tapping test.

CHAPTER 6: CONTEXT EXPERIMENTS

6.1 INTRODUCTION

In this chapter, experiments are intentionally developed as contextual tasks with the purpose of analysing the users' ability in a real world context. Similarly, Chapter 5 consists of tracing, dragging, one direction tapping test and a multi directional tapping tests, however in contextual undertaking.

This experiment is deliberately developed as context tasks with the purpose of analysing the user's ability in a real world context. Generally, the aim was also to discover if there are differences in mouse, stylus and touch on the tracing test and dragging test with different levels of difficulty that could affect users' performance and satisfaction.

Pietro: I think overall you should discuss in more depth the issues/reasons for having out of context and in context tasks. You should probably refer to some literature to do with cognition or similar. This discussion I am referring to needs to be strategically placed somewhere in the thesis that will help to cover all your experiments as it is a fundamental approach that you have used throughout and therefore needs more depth.

Similarly, findings from pilot studies were taken into account in the experiment. The use of apparatus and materials was revised. The design of the prototype and questionnaire is based on the guidance of The International Organization for Standardization known as ISO (ISO 9241-420, 2011).

6.1.1 USERS

Participants were paid volunteers who were recruited through posters and personal contact. They were given BND $10 (*equivalent to £5*) each for completing both tests.

The initial recruitment of the participants took place by means of a recruitment questionnaire. Since the test involved testing aspects of using touch-based technology, it was deemed important to have participants with a certain amount of experience in using touchscreen devices and computers in general. This is because if there happened to be a number of beginners to such activities, these could lead to compromised results. Therefore:

- All participants had computer experience.
- All participants had experience with a touch-based device.
- All the participants' uncorrected visual problems or physical limitations that would inhibit their use of the mouse, stylus and touch as an input device were accounted for.
- *Table 6.1* is the demographic of the sample group

		CONTEXT			
		Tracing	**Dragging**	**One Direction Tapping**	**Multi Direction Tapping**
TOTAL USER		24	24	23	23
Gender	Male	8	8	14	14
	Female	16	16	9	9
Age	18 – 25	23	23	22	22
	26 – 39	1	1	1	1
	40 and above	0	0	0	0
Hand Use	Right	19	19	21	21
	Left	5	5	2	2
Academic Level	Undergraduate	23	23	21	21
	Postgarduate	1	1	1	1
	Others	0	0	1	1

Table 6.1: Demographics of sample group (Context)

They all signed an informed consent document informing them of the goals and activities of the study, their rights to terminate, and the confidentiality of their performance.

6.1.2 APPARATUS AND MATERIALS

6.1.2.1 Testing apparatus

The room used for the experiment was a small office room on the Second Floor in the Multimedia and Technology Centre, Seri Begawan Religious Teachers University College, Brunei. The max capability of the room allows one participant to be assessed in a single shot, shown in *Figure 6.1*:

Figure 6.1: Experiment Area (Tracing Test and Dragging Test - Context)

This experiment was conducted based on the following equipment/tools:

- A convertible Samsung computer tablet with keyboard (Samsung ATIV Smart PC XE700T1C Tablet with Keyboard)
- Intel Core i5-3317U Dual Core Processor,
- Microsoft Windows 8 64bit, 64GB Storage, 4GB DDR3 RAM
- 11.6" Full HD Touch Screen

- 1920x1080 Full HD Resolution
- Wireless Optical PC five button mouse with 1000 dpi, manufactured by Samsung
- S-Pen stylus
- A Five-point 'Likert-type' scale (Likert, 1932) questionnaire (*see Appendix A7*), used to collect the user profile (i.e. age, gender, etc.) and subjective feeling about the device design and the discomfort in the particular body region;

6.1.2.2 Task

Experimental tasks were presented by two different tests: Test 1 is tracing and test 2 is dragging. This prototype was written in adobe flash and runs under Windows 8.

Test 1 (Tracing Test) consists of a four route map, each with a diameter of 80 mm. The participants were instructed to trace a free-hand line to measure the distance the route map using touch, stylus and mouse around each of the circles (*see Figure 6.2*) — in the clockwise direction.

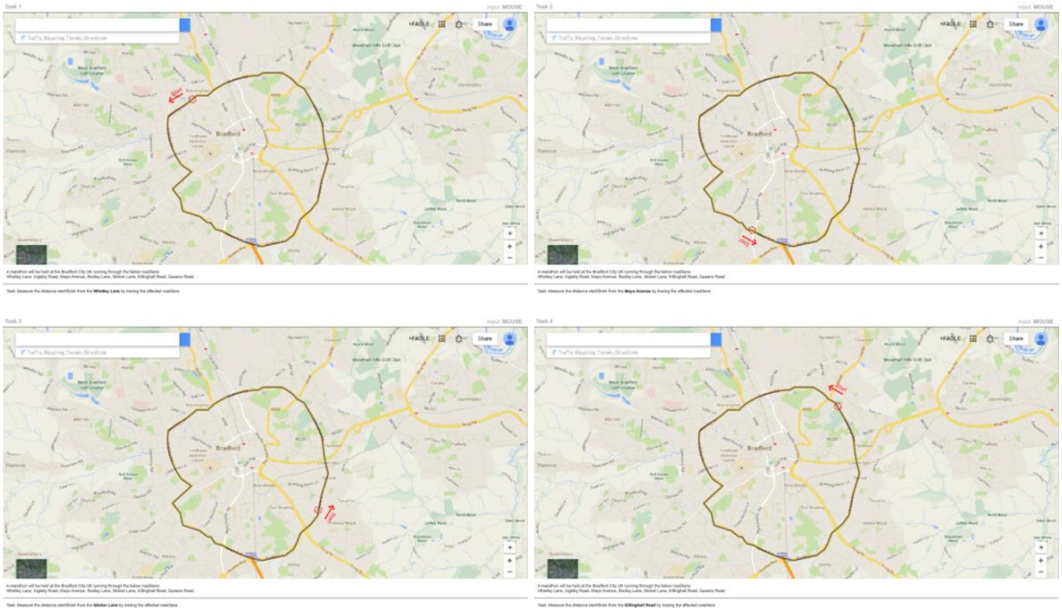
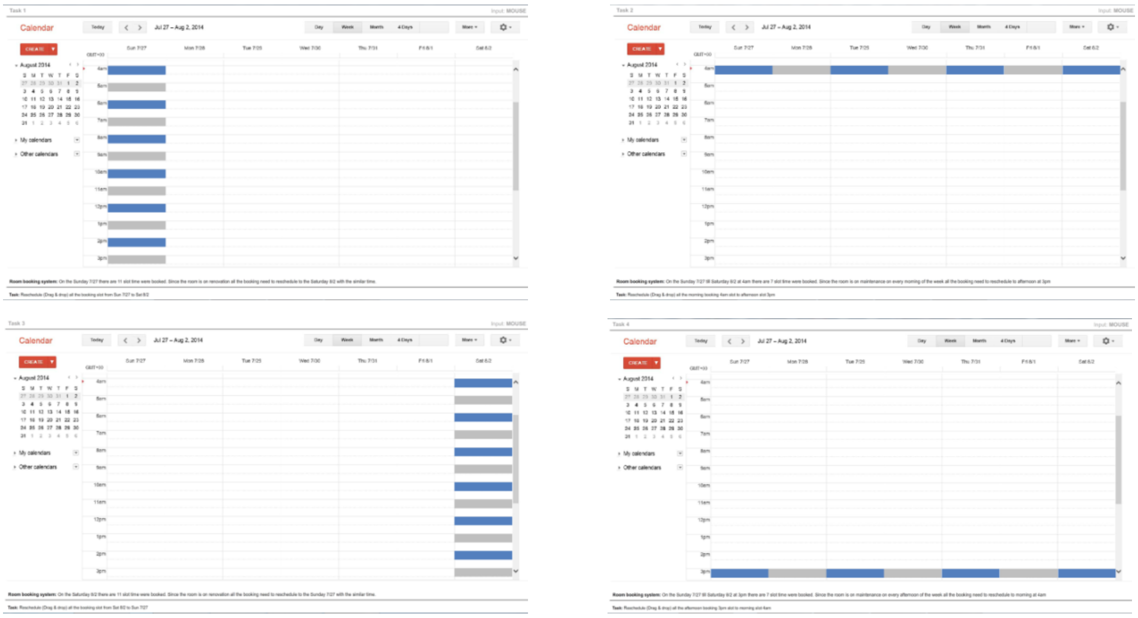


Figure 6.2: Test object and direction of movement (Tracing Test - Context)

Test 2 is a drag test on an object, namely a rectangle with a width of 30 mm over a distance of 140 mm and participants were asked to place it on a new target place that have been specified (*see Figure 6.3*). The participants were also instructed to perform the task in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.



Room booking system: On the Sunday 7/27 there are 11 slot time were booked. Since the room is on renovation all the booking need to reschedule to the Saturday 8/2 with the similar time.

Task: Reschedule (Drag & drop) all the booking slot from Sun 7/27 to Sat 8/2

Room booking system: On the Sunday 7/27 till Saturday 8/2 at 4am there are 7 slot time were booked. Since the room is on maintenance on every morning of the week all the booking need to reschedule to afternoon at 3pm

Task: Reschedule (Drag & drop) all the morning booking 4am slot to afternoon slot 3pm

Room booking system: On the Saturday 8/2 there are 11 slot time were booked. Since the room is on renovation all the booking need to reschedule to the Sunday 7/27 with the similar time.

Task: Reschedule (Drag & drop) all the booking slot from Sat 8/2 to Sun 7/27

Room booking system: On the Sunday 7/27 till Saturday 8/2 at 3pm there are 7 slot time were booked. Since the room is on maintenance on every afternoon of the week all the booking need to reschedule to morning at 4am

Task: Reschedule (Drag & drop) all the afternoon booking 3pm slot to morning slot 4am

Figure 6.3: Arrangement of the objects for the dragging task for the cardinal directions (Dragging Test - Context)

Test 3 is one-direction tapping and test 2 is multi-directional tapping. These tests were written in adobe flash and run under Windows 8.

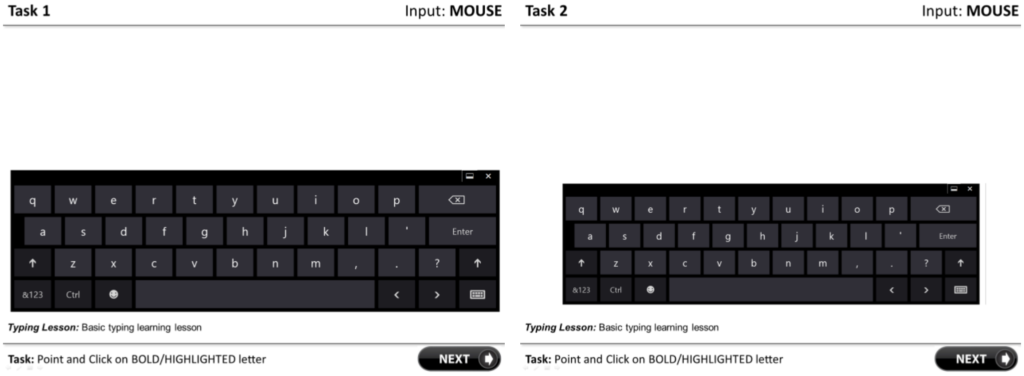
Since this one direction tapping test is a context task with the purpose of analysing the users' ability in a real world context, thus a virtual keyboard has been used (*see Figure 6.4*). The task consists of alternately tapping between the two keyboard keys. The participants were instructed to point and click, along one axis, within each key for 25 times using Touch, Stylus and Mouse. Each test session starts when the user first moves the pointer into a key and actuates it. This allows the participant to move quickly back and forth between the two keys.

This experiment consists of four tasks altogether in which the level of difficulty increases at each stage in which the target becomes smaller and the distance greater.

Figure 6.4: One-direction tapping task (Context)

For test 4 (Multi-directional tapping test), a virtual keyboard has also been used (*see Figure 6.5*). The participants were instructed to point and click a keyboard key by tapping on a random highlighted key for 25 times using *Touch*, *Stylus* and *Mouse*. Each test session starts when the user first moves the pointer onto a key and actuates it. This allows the participant to move quickly back and forth between the two keys.

This experiment consists of three tasks in which the target becomes smaller thereby making each level of the task increasingly difficult.



Typing Lesson: Basic typing learning lesson

Task: Point and Click on BOLD/HIGHLIGHTED letter

Typing Lesson: Basic typing learning lesson

Task: Point and Click on BOLD/HIGHLIGHTED letter


Task 3

Input: **MOUSE**

Typing Lesson: Basic typing learning lesson

Task: Point and Click on BOLD/HIGHLIGHTED letter

Figure 6.5: Multi-directional tapping test (Context)

6.1.3 PROCEDURE

Participant was asked to present themselves at a specific room in the institution set aside for the experiment. During the experiment each participant was seated at the desk in the room with the laptop facing them and the researcher sat opposite the participant.

There were two sections in the experiment: in section 1 of the SOP (*see Figure 6.6*), the experimenter introduced the SOP to participants and demonstrated each task to familiarize the participants with the task and the laboratory environment. After that, participants were asked to sign a Consent Form to ensure commitment to the experiment. Participants were then interviewed and filled out 'personal information' to gather demographic data, i.e. age, gender, preferred hand, and experiential data such as computer experience.

In section 2 of the SOP (*see Figure 6.7*), participants were instructed to perform each task "*as accurately as possible and as fast as possible*" (Zhai *et al*., 2004). There were **four** different tests: test 1 the tracing test, test 2 the dragging test, test 3 One-direction tapping and test 4 Multi-directional tapping. The participants were instructed to use touch, mouse and a stylus in order to perform the test.

The task was explained and demonstrated to the participant. They were instructed to work as fast as possible while still maintaining high accuracy. Participants were also instructed to continue without trying to correct errors. Moreover, the prototype would record variables such as movement time (*MT)* and the error rates (*ER*).

After completion of the tracing task, participants rested for a few minutes before receiving instruction on the next task.

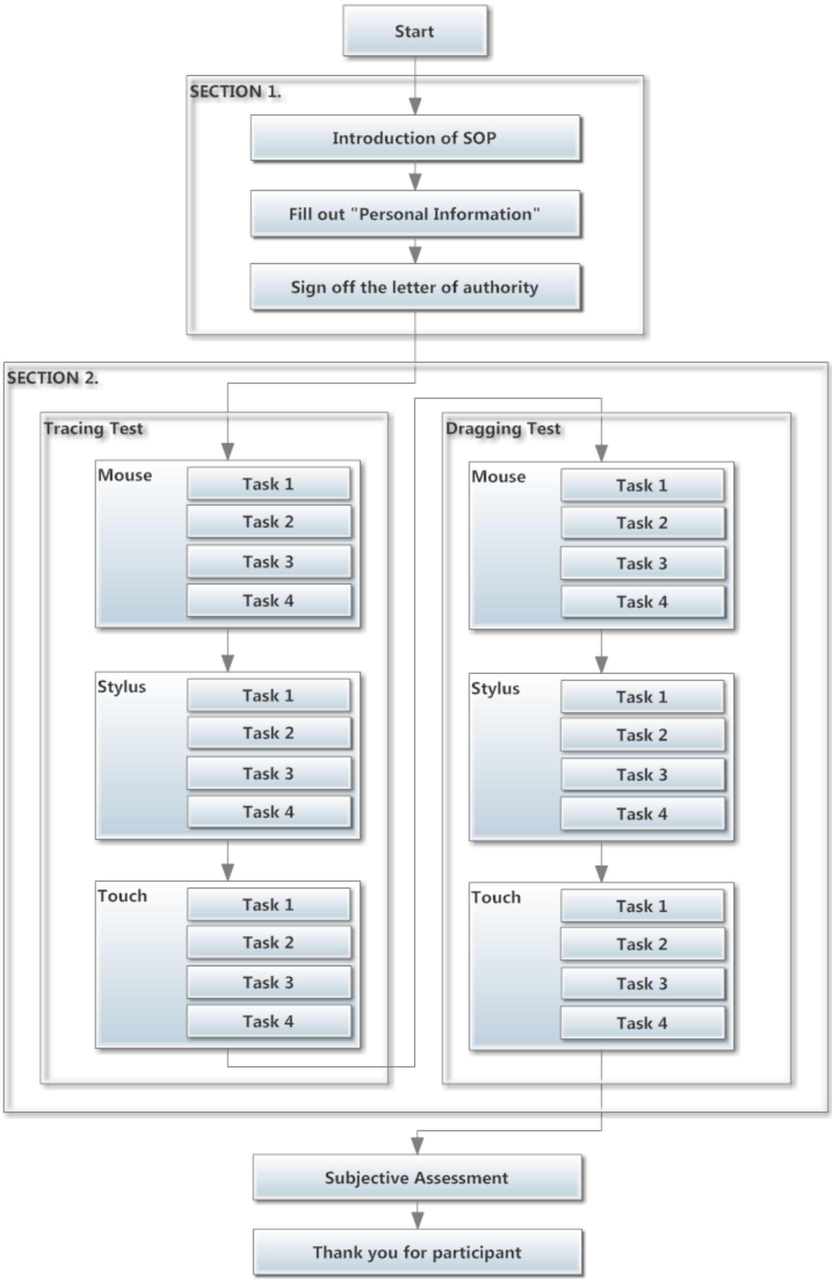


Figure 6.6: Standard Operation Procedure (SOP) (Tracing Test and Dragging Test - Context)

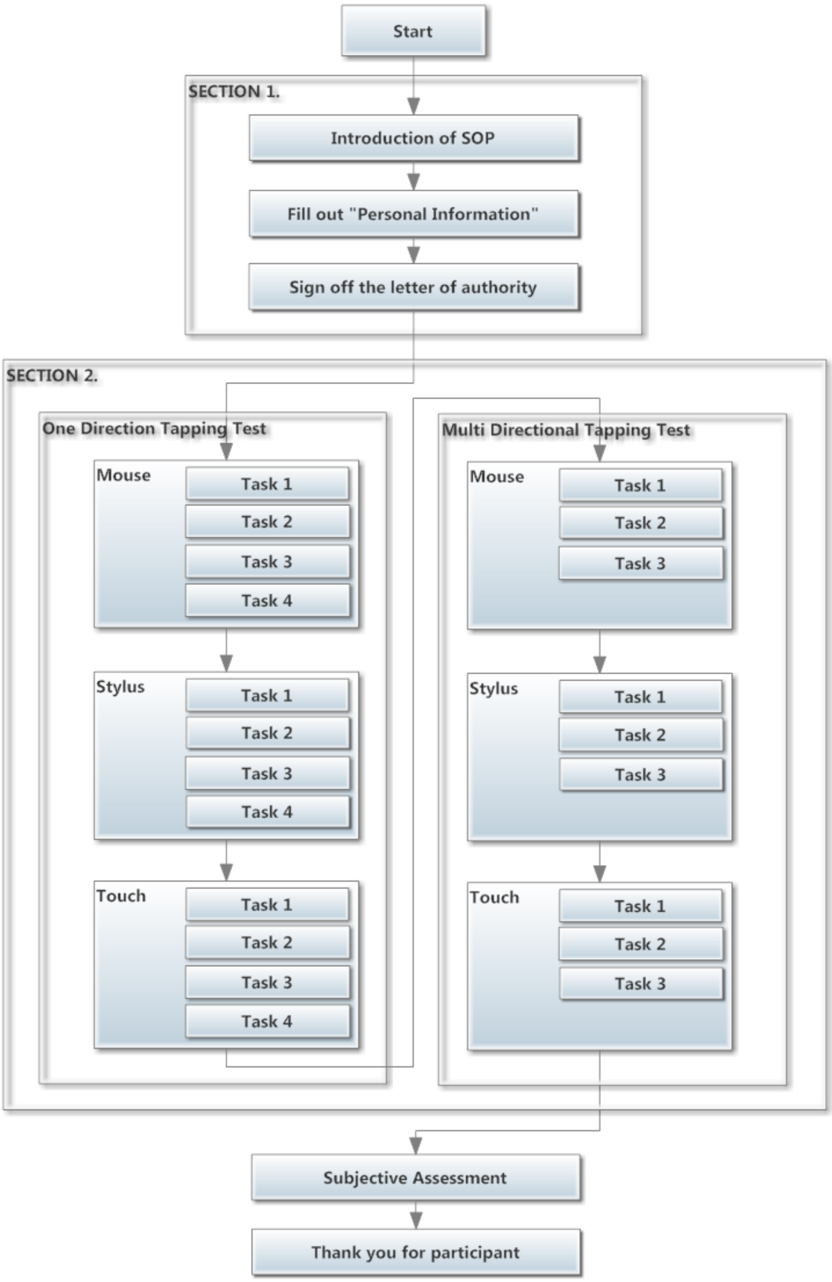


Figure 6.7: Standard Operation Procedure (SOP) (One Direction Tapping Test and Multi Direction Tapping Test - Context)

At the conclusion of the performance portion of the experiment, participants were asked to respond to a written questionnaire asking them to rate their experience in using the device. The questionnaire consisted of forty-seven questions covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

In this questionnaire, the first 5 questions asked the subjects about general questions. The next 21 questions were asking about the first prototype tests and followed by another 21 questions asking the second prototype. In each experiments there are two prototype to be test by the subjects i.e. Test 1 (Tracing and Dragging), Test 2 (One Direction Tapping and Multi-Direction Tapping). Researcher then categorised the responses into fourteen data points as been listed in *Table 6.2* illustrates this device assessment questionnaire that illustrates this device assessment questionnaire.

They were also told in advance that they can inquire if they have problem understanding the terms. Researcher also provided explanation of the meaning of the various surveys terms and even translated the incomprehensible and misunderstood terms to Malay language to ensure that participants fully understood each term.

General indices

(Please (√) Tick Appropriate Box)

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

Fatigue indices

(Please (√) Tick Appropriate Box)

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

Table 6.2: Device Assessment Questionnaire (Tracing Test and Dragging Test - Context)
Source: BS EN ISO 9241-420 (2011)

The total time spent by each participant ranged from slightly less than an hour to one hour and 30 minutes. The performance section took between 45 minutes to one hour to complete.

6.1.3.1 *Statistical Analysis*

The data for movement time (*MT)* of the tracing and dragging tests and the Error Rate (*ER*) of tracing test were collected directly by the prototype which presented the experimental tasks. The data was then prepared for further statistical analysis by computing values for movement time (*MT*) and Error Rate (*ER*). Descriptive Statistics and Inferential Statistics were performed using SPSS.

6.1.3.2 *Device Assessment Questionnaire*

The mean and standard deviation of the ratings for each of the forty-seven questions was computed. Given the ordinal nature of the data, the Friedman test non-parametric statistic was computed to test for significant differences between participants in the three device groups.

6.2 EXPERIMENT 1: TRACING TEST

The first test is the tracing test which is to evaluate and to measure a map route, free-hand input. The test map consists of a four route map, each with a route diameter of 80 mm. The participants are instructed to draw a free-hand line (to measure distance) around each of the routes in a clockwise direction.

6.2.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

6.2.1.1 Tracing test

H1: There is a difference in the amount of speed using touch, mouse or stylus.
H2: There is a difference in the amount of error using touch, mouse or stylus.
H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

6.2.2 EXPERIMENT DESIGN - TRACING TEST

The tracing test's main objective was to carry out an investigation of the difference of three input devices. It was assumed that the same participants take part in all conditions. Due to this assumption, the test deployed a within users study design whereby 23 respondents were given the task which has 12 tasks altogether. The design was chosen to equate the groups taking part in the experiment because every participant is in a group hence there cannot be any significant differences due to their similarities. According to Langston (2014), the within study design significantly reduces the number of required participants because it requires a smaller size of sample, and conducts the same kind of quality research. It also helps in alienating issues associated with finding and recruiting qualified participants (Lazar *et al*, 2010).

6.2.3 VARIABLES - TRACING TEST

6.2.3.1 Independent Variables

The independent variables were:

Factors/Parameters	Level
Route diameter	80 mm
Task	4

The test object consists of a route map, with a diameter of 80 mm. The participants were instructed to draw a free-hand line tracing to measure the route map.

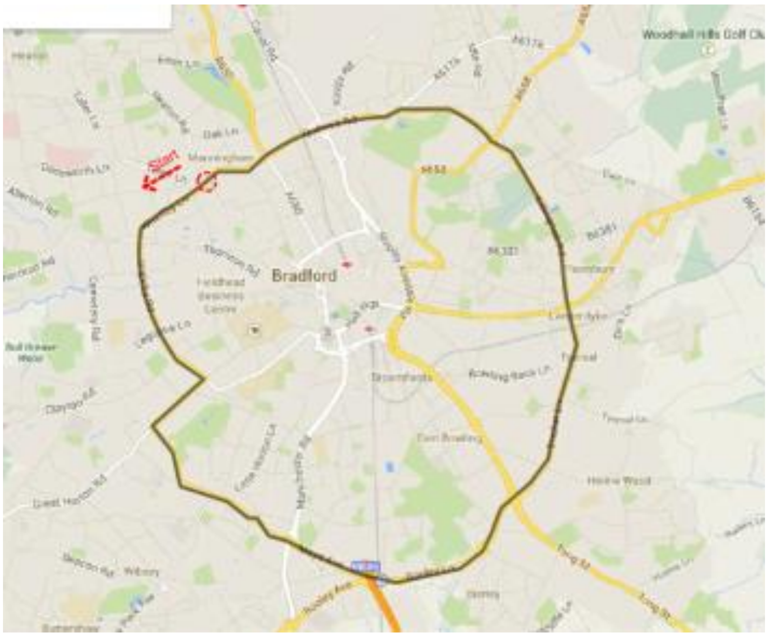


Figure 6.8: Route map and direction to start the measurement (Tracing Test - Context)

6.2.3.2 *Dependent variables*

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (i.e. **design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT*) for each task, and error rate (*ER*). Error rate was determined by the distance in millimeters by which the pointer missed its target.

6.2.3.3 *Dependent measures*

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch summarized in *Table 6.3*.

Dependent measures	Description
Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 6.3: Objective measures of the human performance (Tracing Test - Context)

Figure 6.9 show the distance measurement of the test object (each of the circles) and the free-hand-drawn line, in full millimeters, at 36 locations.

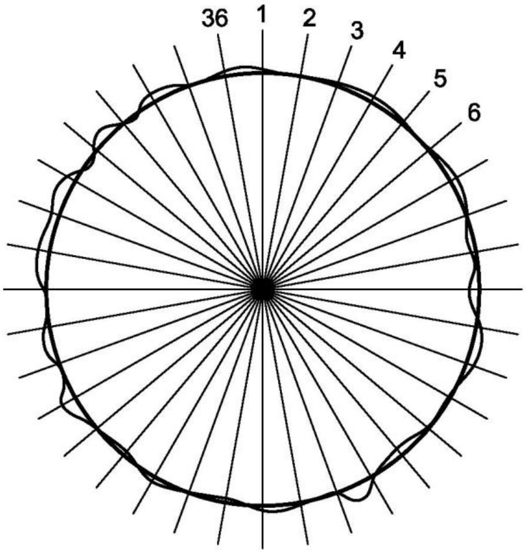


Figure 6.9: Test object and points of measurement for deviations (Tracing Test - Context)
Source: BS EN ISO 9241-420 (2011)

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire, shown in *Table 6.4*.

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both tests, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 6.4: Subjective attributes of the Five-point Likert scale subjective assessment (Tracing Test - Context)

In terms of the user participants' profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

6.2.4 RESULT ANALYSIS AND DISCUSSION

6.2.4.1 TEST DATA

6.2.4.1.1 TRACING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistics (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) shows that the data from all the input devices suggests that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Table 6.5 indicates that the movement time (*MT)* for every task of the touch is significantly faster than the other two devices (mouse and stylus). Moreover, the amount of speed shows a statistically significant difference for the mouse and stylus but not for touch.

Device	N	Task				P value
		1	2	3	4	
Mouse	23	49.14 (24.97)	41.24 (17.32)	37.69 (17.62)	35.70 (15.55)	.000
Stylus	23	52.69 (25.87)	49.27 (6.39)	55.27 (8.85)	70.71 (14.62)	.000
Touch	23	41.19 (7.61)	41.84 (6.02)	46.50 (8.51)	55.59 (11.15)	.456

Table 6.5: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test on every task (Tracing Test - Context)

Figure 6.10 illustrates that the amount of speed of the stylus and touch have a similar rate of movement time (*MT*) except in Task 2 where the stylus shows a slight decrease and touch shows a slight increase. However, their amount of speed continued to increase steadily over the next two tasks until it reached their highest point in Task 4. On the other hand, the mouse showed a downward trend of the movement time over the next level of each task. Through observation, participants showed a faster movement time using touch and stylus in Task 1 and a slower rate over the next level of task because participants showed their alertness and concentration which led to an increase in speed. However, participants recorded a gradual decrease in speed throughout the task because participants were accustomed to using a mouse.

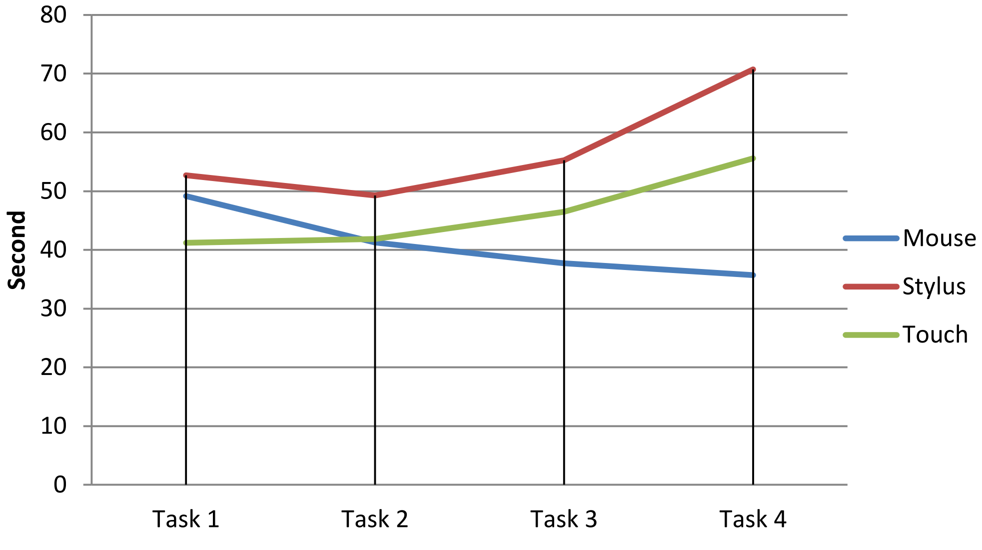


Figure 6.10: The amount of speed based on every task (Tracing Test - Context)

In terms of the overall movement time, a Friedman test was run to determine if there were differences in the amount of speed between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of speed was statistically significantly different between the 3 input devices, $\chi^2(2) = 138.562$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between touch (*Mdn* = 18.79) and stylus (*Mdn* = 22.36) ($p < .0005$), touch (*Mdn* = 18.79) and mouse (*Mdn* = 37.82) (p < .0005) and stylus (*Mdn* = 22.36) and mouse (*Mdn* = 37.82) (p < .0005).

Table 6.6 indicates that touch (22.63 sec ±15.09) has the fastest overall movement time (*MT),* that is two-times faster than the mouse (40.95sec ±19.59) which recorded the slowest, while the stylus (25.83 sec ±10.59) has the second fastest overall movement time (*MT).*

Device	N	Mean	Std Deviation
Mouse	23	40.95	19.59
Stylus	23	25.83	10.59
Touch	23	22.63	15.09

Table 6.6: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Tracing Test - Context)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 6.7*, the Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (*ER*) with the touch, mouse and stylus does not show a statistically significant difference.

Device	N	Task				P value
		1	2	3	4	
Mouse	23	169.17 (82.35)	152.67 (76.95)	151 (54.87)	159.96 (63.14)	.069
Stylus	23	104.08 (32.40)	109.79 (32.45)	106.92 (42.33)	111.38 (36.48)	.112
Touch	23	220.63 (71.29)	239.75 (78.74)	220.29 (74.30)	240.58 (67.89)	.065

Table 6.7: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Friedman test on every task (Tracing Test - Context)

Figure 6.11 illustrates that the error rates of stylus and mouse have a similar tendency except in Task 2 where the stylus shows a slight increase and touch shows a slight decrease, but their error rates increase insignificantly in the last task. On the other hand, touch shows fluctuation in error rates from task 1 to task 4. Through observation, participants show higher error rates using touch because the tracing line is blocked by the participant's finger that leads to inaccuracy.

However, the stylus show lowest error rates because it is a pen-shaped instrument that has a sharp point which leads to accuracy in doing the tracing test compared to the other input devices.

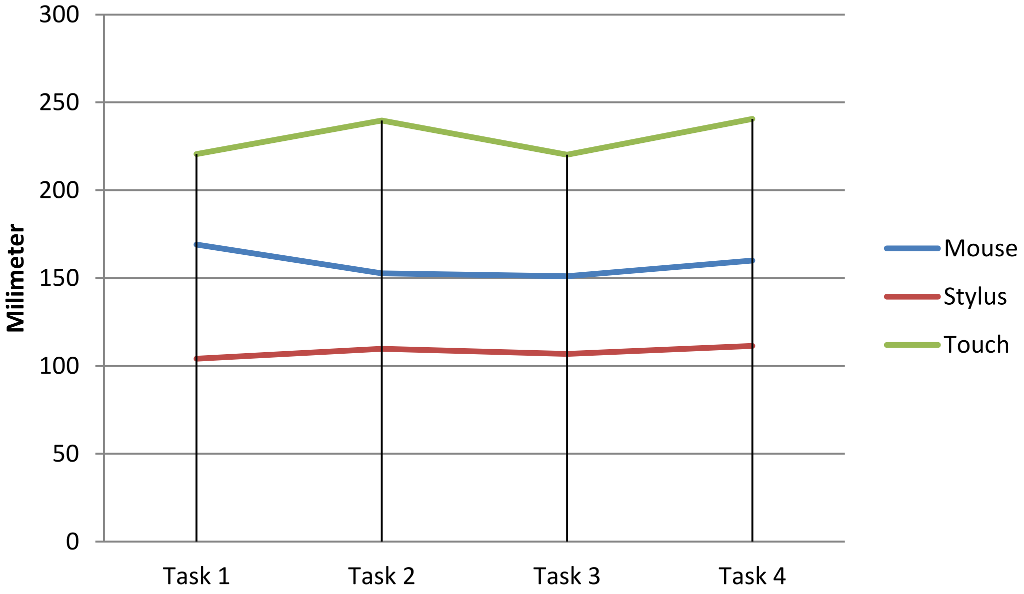


Figure 6.11: The error rates based on every task (Tracing Test - Context)

In terms of the overall error rate, a Friedman test was run to determine if there were differences in the amount of error between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error was statistically significantly different between the 3 input devices, $\chi^2(2) = 128.089$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of error between the stylus (*Mdn* = 99) and mouse (*Mdn* = 147.5) ($p < .0005$), stylus (*Mdn* = 99) and touch (*Mdn* = 226.5) (p < .0005) and mouse (*Mdn* = 147.5) and touch (*Mdn* = 226.5) (p < .0005).

Table 6.8 indicates that the overall error rate (*ER*) of touch (230.31 mm ±72.68) is almost two-times greater than the stylus (108.04 mm ±35.68) which recorded the lowest error among the three input devices, while the mouse (158.20 mm ±69.44) showed the second highest error rate (*ER*).

Device	N	Mean	Std Deviation
Mouse	23	158.20	69.44
Stylus	23	108.04	35.68
Touch	23	230.31	72.68

Table 6.8: The effect of the device difference (touch, mouse and stylus) on the overall error rate (Tracing Test - Context)

6.2.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input device based on their experience of using it.

6.2.4.2.1 TRACING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 6.9*, the inter-reliability test discovered that inter-reliability of the design is very high among all the input devices. Therefore, the comparison can be made for all the input devices in terms of the general indices and the fatigue indices as summarised in *Table 6.10*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.9.25	.960	.962
Fatigue Indices	.961	.959	.823

Table 6.9: Inter Reliability Statistics with the mouse, stylus and touch

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	3.96	4.26	3.30	.003*
General Indices	Actuation force	23	3.83	3.87	3.17	.007
	Operation smoothness	23	4.04	4.04	3.09	.000*
	Operation effort	23	4.00	3.96	3.17	.002*
	Accuracy	23	3.74	4.13	2.87	.000*
	Operation speed	23	3.87	4.22	3.22	.000*
	General comfort	23	4.04	3.96	3.17	.004*
	Overall operation	23	4.09	4.04	3.17	.001*
Fatigue Indices	Finger fatigue	23	4.04	4.09	3.13	.001*
	Wrist fatigue	23	4.04	3.87	3.39	.014
	Arm fatigue	23	4.04	3.96	3.39	.003*
	Shoulder fatigue	23	4.00	4.00	3.26	.001*
	Neck fatigue	23	3.96	4.04	3.39	.002*
	Overall operation	23	4.13	4.00	3.22	.002*

Table 6.10: Result analysis of the Five-point Likert scale subjective assessment with the touch, mouse and stylus (Tracing Test - Context)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participants considered it much easier using the stylus (4.26) on the task given rather than touch (3.30) and mouse (3.96) ($p<0.05$). This indicates that the stylus device is easier to use than touch and stylus in the tracing test (context).

In terms of the general indices, for four indicators (i.e. actuation force, operation smoothness, accuracy and operation speed), the stylus was are highly rated, while the mouse performed better in terms of operation smoothness, operation effort, general comfort and general overall operation indicators. The touch recorded the worst result as the mean showed the lowest level among the three input devices on all of the seven indicators.

In regards to the fatigue indices, for four indicators (i.e. wrist, arm, shoulder and fatigue overall operation) the mouse performed well in the tracing test tasks, whereas the stylus was rated highly on the areas of finger, shoulder and neck. Touch, on the other hand scored badly for use in the tracing test.

This led to the conclusion that both stylus and mouse were rated highly in the tracing test (context) in respect of general and fatigue indices.

6.3 EXPERIMENT 2: DRAGGING TEST

The second test is the dragging test which evaluates clicking and dragging objects to specific locations, for example; clicking and dragging the pointer down a pull-down menu, and selecting and dragging an object from one window to another. The test object consists of rectangles with a width of 30 mm over a distance of 140 mm and places them in circles with a diameter of 10 mm. The task was performed in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.

6.3.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

6.3.1.1 Dragging test

H1: There is a difference in the amount of speed using touch, mouse or stylus.
H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

6.3.2 EXPERIMENT DESIGN - DRAGGING TEST

Similarly, the dragging test aims to carry out an investigation of the difference between three input devices. It is assumed that the same participants engage in all conditions hence the most appropriate study design for this test is the within users design. In this test, 23 participants were assigned to perform the test which has 12 tasks altogether.

Similarly, since the test is to investigate the difference between three input devices and the same participants take part in all conditions, a within users design was deployed. Twenty-three participants were assigned to do the test which included 12 tasks (3 inputs x 4 task) altogether.

6.3.3 VARIABLES - DRAGGING TEST

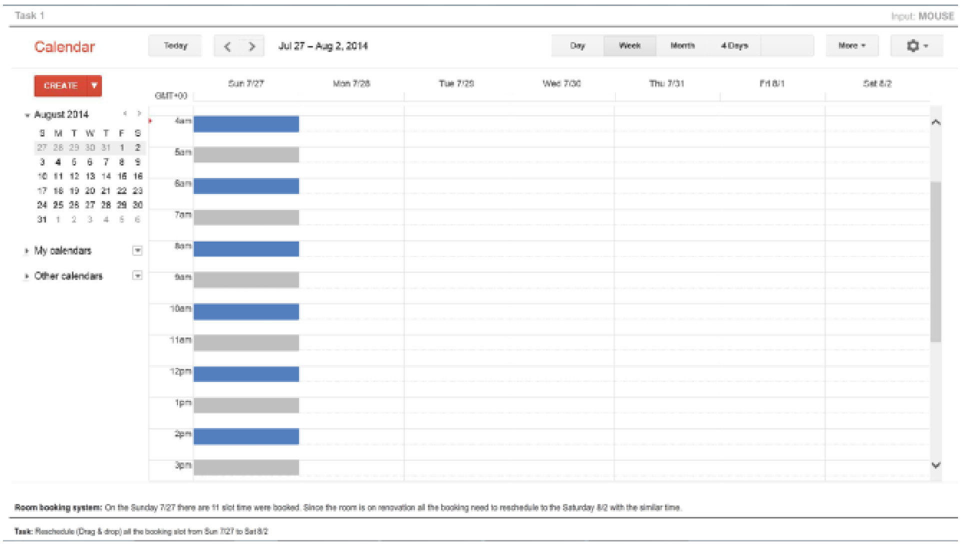
6.3.3.1 Independent Variables

The independent variables were:

Factors/Parameters	Level
Target Width	30 mm
Target Distance	140 mm
Target	12 per task
Task	4

Figure 6.12 shows the arrangement of the dragging test task in which the test object consisted of a rectangle with a width of 30 mm and over a distance of 140 mm. Participants were to drag the rectangle to a specific place and complete the task in all

four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.



Room booking system: On the Sunday 7/27 there are 11 slot time were booked. Since the room is on renovation all the booking need to reschedule to the Saturday 8/2 with the similar time.

Task: Reschedule (Drag & drop) all the booking slot from Sun 7/27 to Sat 8/2

Figure 6.12: Arrangement of the objects for the dragging task for the cardinal directions (Dragging Test - Context)

6.3.3.2 Dependent variables

The dependent variables consisted of the following two clusters: the **objective human performance** (i.e. movement time (*MT*) and error rate (*ER*) and **the subjective feelings** about the device (i.e. **design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

Since in a real world context of a Google a Calendar, dragging task of merging an object is set to automatically merge, the error rate for this task was not recorded.

6.3.3.3 Dependent measures

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch (summarized in *Table 6.11*).

Dependent measures	Description
*Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 6.11: Objective measures of the human performance (Dragging Test - Context)
**Tracing test only*

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire shown in *Table 6.12*:

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both tests, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 6.12: Subjective attributes of the Five-point 'Likert-type' scale (Likert, 1932) subjective assessment (Dragging Test - Context)

In terms of the user participants profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

6.3.4 RESULT ANALYSIS AND DISCUSSION

6.3.4.1 TEST DATA

6.3.4.1.1 DRAGGING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistics (i.e Q-Q plot, Normal probability plot and Shapiro-Wilk test) show that the touch data is not normally distributed. Thus, the Friedman test was carried out. On the other hand, the mouse and stylus data show that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Table 6.13 indicates that the movement time (*MT)* for every task of the stylus is significantly faster than the other two devices (mouse and touch). Moreover, the amount of speed shows a statistically significant difference on mouse and touch but not for stylus.

Devices	N	Task				P value
		1	2	3	4	
Mouse	23	42.16 (19.97)	18 (2.94)	28.93 (4.58)	17.42 (3.63)	.000*
Stylus	23	27.71 (6.23)	17.08 (5.20)	28.42 (7.47)	7.45 (4.05)	.735*
Touch	23	31.56 (14.47)	26.15 (17.09)	30.98 (9.13)	20.07 (10.74)	.000**

Table 6.13: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test and Friedman test on every task (Dragging Test - Context)

** Repeated Measure ANOVA test*

*** Friedman test*

Figure 6.13 illustrates that the amount of speed on each task using mouse, stylus and touch have a similar tendency of movement time (*MT*). It went up and down widely as the tasks progressed. It can be concluded that participants showed faster amounts of speed in Tasks 2 and 4, while they took much of their time in Task 1 and Task 3. Through observation the movement time (*MT*) of task 1 and task 3 showed high movement time because dragging a rectangle from left-to-right and right-to-left, to a new target place is difficult. Tasks 2 and 3 show less amount of movement time because dragging a rectangle up and down to the target place is easier.

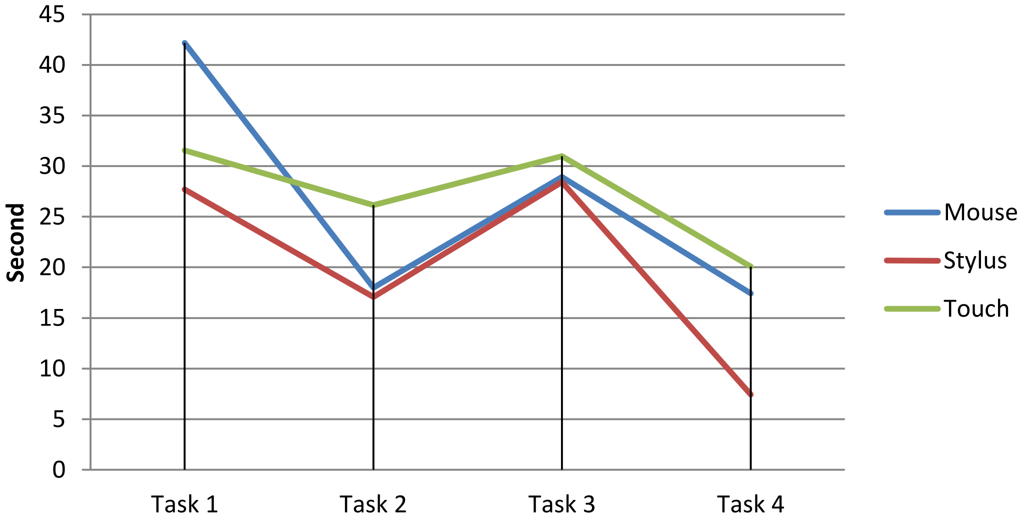


Figure 6.13: The amount of speed based on every task (Dragging Test - Context)

A Friedman test was run to determine if there were differences in the amount of speed between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of speed was statistically significantly different between the 3 input devices, $\chi^2(2) = 23.583$, $p < .0005$. Post hoc analysis revealed statistically significant differences in the amount of speed between the stylus (*Mdn* = 20.84) and touch (*Mdn* = 23.15) ($p < .0005$) and stylus (*Mdn* = 20.84) and mouse (*Mdn* = 23.83) (p < .0005), but not touch (*Mdn* = 23.15) and mouse (*Mdn* = 23.83) (p > .0005).

As can be seen in *Table 6.14*, the mean and standard deviation indicate that the slowest overall movement time (*MT)* is touch (27.19 sec ±13.82), followed by the mouse (26.63 sec ±14.47) and then stylus (21.81 sec ±8.61) which recorded the fastest among the devices.

Device	N	Mean	Std Deviation
Mouse	23	26.63	14.47
Stylus	23	21.81	8.61
Touch	23	27.19	13.82

Table 6.14: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Dragging Test - Context)

6.3.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Participants were given forty-seven questions related to the functionality of different input devices and asked to rate each input device based on their experience of using it.

6.3.4.2.1 DRAGGING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 6.15*, the inter-reliability test discovered that inter-reliability of the design is very high among all input devices. Thus, comparison can be made for all the input devices in terms of the general indices and the fatigue indices as summarised in *Table 6.16*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.984	.972	.960
Fatigue Indices	.994	.978	.945

Table 6.15: Inter Reliability Statistics with Mouse, Stylus and touch (Dragging Test - Context)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	4.39	4.39	3.65	.001*
General Indices	Actuation force	23	4.32	4.30	3.70	.017
	Operation smoothness	23	4.35	4.35	3.70	.002*
	Operation effort	23	4.43	4.30	3.52	.000*
	Accuracy	23	4.48	4.48	3.70	.000*
	Operation speed	23	4.39	4.52	3.74	.004*
	General comfort	23	4.48	4.39	3.74	.002*
	Overall operation	23	4.43	4.43	3.83	.004*
Fatigue Indices	Finger fatigue	23	4.30	4.26	3.61	.008
	Wrist fatigue	23	4.30	4.17	3.52	.011
	Arm fatigue	23	4.30	4.04	3.48	.008
	Shoulder fatigue	23	4.35	4.13	3.57	.011
	Neck fatigue	23	4.22	4.26	3.65	.008
	Overall operation	23	4.35	4.13	3.61	.010

Table 6.16: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the Friedman test on the raw data of the subjective assessment (Dragging Test - Context)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participants considered it much easier using the mouse (4.39) and stylus (4.39) on the task given, rather than touch 3.65) ($p<0.05$). This indicates that the mouse and stylus devices are easier than touch in the dragging test (context).

In terms of the general indices, the actuation force, effort and comfort were highly rated when the mouse was used. For smoothness, accuracy and general overall operation the mouse and stylus were similarly highly rated. The touch recorded the worst rating among the participants as the mean showed it as the lowest among the three input devices on all of the seven indicators.

With regard to the fatigue indices, all indicators showed that the mouse in the dragging test tasks performed best in all areas except the neck fatigue, compared to stylus and touch in which touch again scored the worst overall.

This led to the conclusion that the mouse and stylus were considered much more user-friendly in the dragging test (context) in relation to general and fatigue indices.

6.4 EXPERIMENT 3: ONE DIRECTION TAPPING

The third test is the one-direction tapping test in which the evaluating point is the movement along one axis on a horizontal rubber banding, an insert cursor at points along a character string and selecting information in columns or rows. The test object consists of two rectangles with a defined width in the direction perpendicular to the direction of the movement. The task consists of alternately tapping between the two rectangles.

6.4.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency of use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

6.4.1.1 One-direction tapping test

H1: There is a difference in the amount of speed using touch, mouse or stylus.
H2: There is a difference in the amount of error using touch, mouse or stylus.
H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

6.4.2 EXPERIMENT DESIGN - ONE DIRECTION TAPPING TEST

The one direction tapping test aims to find out the difference of three input devices. In the test, the same participants were given a test which had 12 tasks (3 inputs x 4 task) altogether. Due to the nature of the test, a within user design was adopted to equate the groups in the experiment since every participant is in every group. In addition, there cannot be any differences because they are all similar. Besides, the design was chosen because it promotes efficiency by significantly reducing the number of participants required (Langston, 2014). This is because a within study design only requires a small sample size and performs the same type of quality research. Moreover, Lazar *et al* (2010) postulate that many HCI researchers find difficulty in recruiting qualified participants. This design helps to eliminate this aspect.

6.4.3 VARIABLES - ONE DIRECTION TAPPING TEST

6.4.3.1 Independent Variables

The independent variables were:

- Target Width (20 mm, 17 mm, 10 mm, 8 mm)
- Target Distance (44 mm, 96 mm, 163 mm, 210 mm)
- Tapping (1 to 50 per task)
- Task (1 to 4)

6.4.3.2 Dependent variable

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (i.e. **design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* and error rate (*ER*) for each task. Error rate will be recorded if the participant taps outside the target object.

6.4.3.3 Dependent Measure

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with a mouse, stylus and touch (summarized in *Table 6.17*).

Dependent measures	Description
Error Rate	Error attempt is recorded (Tap away from the target)
Movement Time	When the cursor enters the target, it will be counted.

Table 6.17: Objective measures of the human performance (One Direction Tapping Test - Context)

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire, shown in *Table 6.18*.

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both test, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 6.18: Subjective attributes of the Five-point 'Likert-type' scale (Likert, 1932) subjective assessment (One Direction Tapping Test - Context)

In terms of the user participants' profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

6.4.4 RESULT ANALYSIS AND DISCUSSION

6.4.4.1 TEST DATA

6.4.4.1.1 ONE DIRECTION TAPPING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

Descriptive statistics (i.e. Q-Q plot, Normal probability plot and Shapiro-Wilk test) show that the data from all the input devices shows that they are normally distributed and therefore the Repeated Measure ANOVA test was conducted.

Table 6.19 indicates the movement time (*MT*) for every task; touch is significantly faster than the other two devices (mouse and stylus). Moreover, the amount of speed does not show a statistically significant difference for the mouse and stylus, but not for touch.

Devices	N	Task				P value
		1	2	3	4	
Mouse	23	51.63 (12.99)	48.79 (6.98)	53.32 (7.49)	54.55 (7.49)	.095
Stylus	23	44.87 (6.96)	45.56 (5.40)	52.10 (6.29)	52.17 (7.07)	.021
Touch	23	35.88 (3.79)	36.54 (4.90)	40.44 (9.77)	40.48 (6.68)	.000

Table 6.19: The effect of the device difference (touch based, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test on every task (One Direction Tapping Test - Context)

Figure 6.14 illustrates that the amount of speed using mouse, stylus and touch has a similar tendency of movement time (*MT*) except for the mouse in Task 2 where the mouse showed a slight decrease while touch and stylus showed a slight increase. All of the input devices' movement times gradually increased in the last two tasks. Through observation, task 4 had a higher amount of speed compared to other tasks because the difficulty increased and this led to slower movement time. The movement time of touch is faster because it is easier and faster tapping with a finger. However, the mouse showed slowest amount of speed because tapping using a mouse takes more time as the participant needs to control the mouse to the targeted area and then click it.

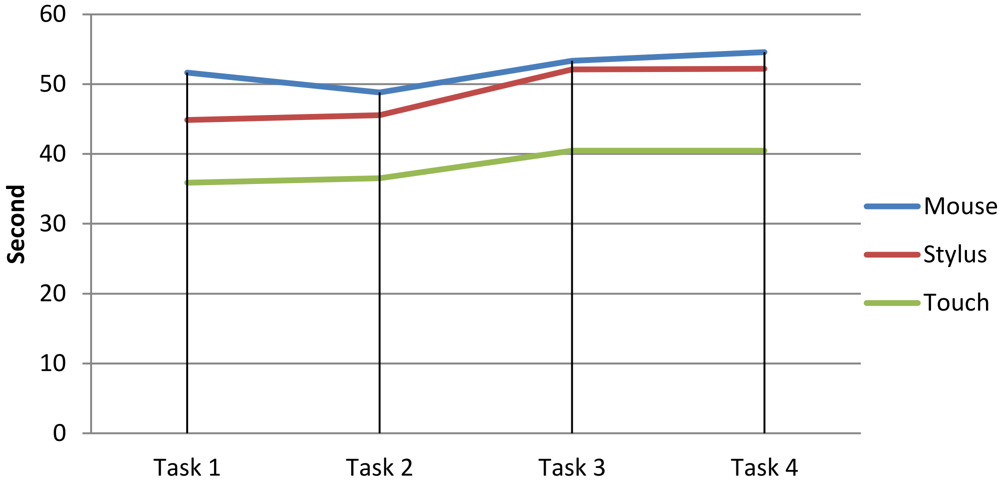


Figure 6.14: The amount of speed based on every task (One Direction Tapping Test - Context)

In terms of the overall movement time, a one-way repeated measures ANOVA was conducted to determine whether there were statistically significant differences in amount of speed over the three types of input: mouse, stylus and touch. There were no outliers and the data was normally distributed, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, χ^2 (2) = 1.347, p = 0.510. Therefore, a Greenhouse-Geisser correction was applied (ε = 0.985).

The three types of input (Mouse, Stylus, Touch) elicited statistically significant changes in amount of speed over time, F(1.971, 179.337) = 105.390, p < .0005, partial η2 = 0.54, with amount of speed decreasing from 52.07 ± 9.19 seconds for mouse input to 48.67 ± 0.57 seconds for stylus input and to 38.41 ± 6.89 seconds for touch input.

Post hoc analysis with a Bonferroni adjustment revealed that the amount of speed was statistically significantly decreased for mouse input to stylus input (3.40 (95% CI, 0.97 to 5.83) second, *p = .003*), and from mouse to touch input (13.66 (95% CI, 11.18 to

16.15) second, $p < .0005$), but not from stylus input to touch input (10.26 (95% CI, 8.02 to 12.51) second, $p < .0005$).

Table 6.20 indicates that touch (37.75 sec ±51.06) has the fastest overall movement time (*MT*), that is almost two-times faster than the mouse (51.45 sec ±45.35) which recorded the slowest, while the stylus (47.93 sec ±29.68) has the second fastest overall movement time (*MT*).

Device	N	Mean	Std Deviation
Mouse	23	51.45	45.35
Stylus	23	47.93	29.68
Touch	23	37.75	51.06

Table 6.20: The effect of the device difference (touch, mouse and stylus) on the overall movement time (One Direction Tapping Test - Context)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 6.21*, the Friedman test is applied on the raw material data to examine the significance of the difference. As a result, it indicates that the error rate (*ER*) of the touch, mouse and stylus has no statistically significantly difference.

Device	N	Task				P value
		1	2	3	4	
Mouse	23	.61 (.891)	.61 (.839)	1.35 (1.58)	.87 (.97)	.062
Stylus	23	.48 (.947)	.43 (.662)	.91 (1.44)	2.09 (3.67)	.000
Touch	23	.61 (.89)	1.09 (1.31)	1.57 (2.84)	2 (3.36)	.063

Table 6.21: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Friedman test on every task (One Direction Tapping Test - Context)

Touch, stylus and mouse have a different tendency of error rates (*ER*) where the mouse showed steady error rates for Tasks 1 and 2 and fluctuated over the next two tasks. There was a slight decrease of error rates using the mouse in Task 2 however, they continued to increase steadily over the next two tasks. On the other hand, using touch showed an increase of error rates in each level of tasks until it reached its highest point in Task 4 (*see Figure 6.15*). Through observation, participants found the difficulty increases in each level of the task and this led to high error rates using touch and stylus. With regards to the mouse, participants were more careful during the first two tasks, however, in Task 3 they recorded high error rates as they became more careless, having done well in the first two tasks, while in Task 4 they concentrated more and did the task carefully.

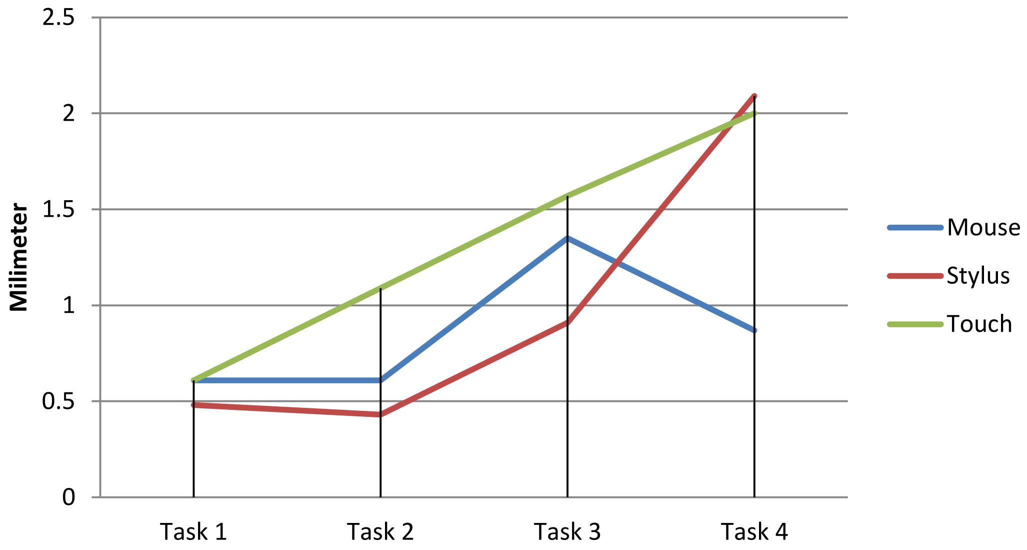


Figure 6.15: The error rates based on every task. (One Direction Tapping Test - Context)

In terms of the overall error rate, a Friedman test was run to determine if there were differences in the amount of error between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. The amount of error decreased from mouse (Mdn = 1) and touch (Mdn = 1) to stylus (Mdn = 0), but the differences were not statistically significant, $\chi^2(2) = 2.986$, $p > .225$

Table 6.22 indicates that the overall error rate (*ER*) of touch (1.32 ±2.36) was almost two-times greater than with the mouse (.86 ±1.14) which recorded the lowest error rate among the three input devices, while the stylus (.98 ±2.13) showed the second highest error rate (*ER*).

Device	N	Mean	Std Deviation
Mouse	23	.86	1.14
Stylus	23	.98	2.13
Touch	23	1.32	2.36

Table 6.22: The effect of the device difference (touch, mouse and stylus) on the overall error rate (One Direction Tapping Test - Context)

6.4.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Users were given forty-seven questions related to the functionality of the different input devices and asked to rate each input devices based on their experience of using it. Since 5 point 'Likert-type' scale (Likert, 1932) were used, a Friedman non-parametric test was done for the analysis it is the best tool to analyse the result based on three different input devices.

6.4.4.2.1 ONE DIRECTION TAPPING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 6.23*, the inter-reliability test discovered that inter-reliability about the design is very high with all the input devices, thus the comparison can be made for all devices in terms of the general indices and the fatigue indices, as summarised in *Table 6.24*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.930	.973	.972
Fatigue Indices	.921	.961	.918

Table 6.23: Inter Reliability Statistics with the mouse. Stylus and touch (One Direction Tapping Test - Context)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	4.52	3.91	4.61	.002*
General Indices	Actuation force	23	3.87	3.43	4.26	.045
	Operation smoothness	23	3.96	3.61	4.30	.019
	Operation effort	23	4.17	3.78	4.30	.020
	Accuracy	23	4.17	3.41	4.48	.001*
	Operation speed	23	3.96	3.70	4.52	.016
	General comfort	23	4.17	3.57	4.35	.004*
	Overall operation	23	4.32	3.70	4.43	.014
Fatigue Indices	Finger fatigue	23	3.96	3.70	4.57	.000*
	Wrist fatigue	23	4.00	3.61	4.48	.001*
	Arm fatigue	23	3.83	3.70	4.13	.019
	Shoulder fatigue	23	4.00	3.48	4.04	.032
	Neck fatigue	23	4.00	3.65	4.35	.008
	Overall operation	23	4.09	3.59	4.35	.022

Table 6.24: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the I Friedman test on the raw data of the subjective assessment (One Direction Tapping Test - Context)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participants considered it much easier using touch (4.61) on the task given rather than the mouse (4.52) and stylus (3.91) (p>0.05). This indicates that the touch device is easier than the mouse and stylus in the one direction tapping test.

In terms of the general indices, for all seven indicators (i.e. actuation force, operation smoothness, operation effort, accuracy, operation speed, general comfort and overall operation) touch was rated highest. The stylus recorded the worst rating among the participants, as the mean shows, for all seven indicators.

Similarly with regard to the fatigue indices, for all six indicators (i.e. finger fatigue, wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue, overall operation) touch gave the most satisfaction in the one direction tapping test. The stylus was rated to be the worst input device for all of the indicators.

This led to the conclusion that touch came out best overall in the one direction tapping test in relation to general and fatigue indices.

6.5 EXPERIMENT 4: MULTI-DIRECTIONAL TAPPING

The fourth test is the multi-direction tapping test in which the evaluating point is the movements in many different directions in the repositioning of a pointer at different areas on the screen, cell selection in a spreadsheet and selecting randomly located icons. The test object consists of targets positioned around the circumference of a circle. The targets are arranged so that the movements are nearly equal to the diameter of the circle. A box marked with X is the target to which the subject should advance. Each test session starts after the subject points to the topmost target and ends when the sequence is completed (at the topmost target). The tests were conducted with a range of difficulties in the size and the distance of the target squares.

6.5.1 HYPOTHESIS

Several hypotheses were devised around the area of efficiency of use and user satisfaction for the purposes of this experiment. In all cases, statistically significant differences in the data to be collected were looked at.

6.5.1.1 Multi-directional tapping test

H1: There is a difference in the amount of speed using touch, mouse or stylus.
H2: There is a difference in the amount of error using touch, mouse or stylus.
H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

6.5.2 EXPERIMENT DESIGN – MULTI- DIRECTIONAL TAPPING TEST

The multi-directional tapping test also aimed to investigate the difference of three input devices. In this test, the same participants take part in all conditions hence a within users design was deployed. 23 participants were provided with the test which had 9 tasks (3 inputs x 3 task) altogether.

Similarly, since the test is to investigate the difference of three input devices and the same participants take part in all conditions, a within users design was deployed. Twenty-three participants were assigned to do the test which included 9 tasks (3 inputs x 3 task) altogether.

6.5.3 VARIABLES – MULTI- DIRECTIONAL TAPPING TEST

6.5.3.1 Independent Variables

- Target Width (11 mm, 16 mm, 18 mm)
- Tapping (1 to 25 per task)
- Task (1 to 3)

6.5.3.2 Dependent variables

The dependent variables consisted of the following two clusters: the **objective human performance** and **the subjective feelings** about the device (**design**-general indices and the **discomfort-**fatigue indices in the particular body regions), while also taking into account the participant profile.

The objective **human performance dependent** variables are movement time (*MT)* and error rate (*ER*) for each task. Error rate will be recorded if the participant taps outside the target object.

6.5.3.3 Dependent measures

The dependent measures were that the performance was measured by examining movement time (*MT)* of the tasks and the number of errors made. In regards to the objective human performance, these objective measures were collected during the experiment with an ordinary mouse, stylus and touch (summarized in *Table 6.25*).

Dependent measures	Description
Error Rate	Error attempt is recorded
Movement Time	When the cursor enters the target, it will be counted.

Table 6.25: Objective measures of the human performance (Multi Direction Tapping Test - Context)

As for the participant feelings, these subjective attributes were collected by using a Five-point 'Likert-type' scale (Likert, 1932) questionnaire shown in *Table 6.26*:

The subjective opinions were measured by means of a post-experiment questionnaire. Participants were asked to rate various aspects of the user interface using a 'Likert-type' scale (Likert, 1932). The main areas covered by the post-experiment questionnaire were opinions about the overall evaluation on both test, covering issues of physical operation, fatigue and comfort, speed and accuracy, and overall usability. Participants were asked to respond to each question with a rating from worse to better.

Cluster/Level	Factor	Current studies
	Consider the task of using?	ISO 9241-420 (2011)
General Indices	Actuation force	
	Operation smoothness	
	Operation effort	
	Accuracy	
	Operation speed	
	General comfort	
	Overall operation	
Fatigue Indices	Finger fatigue	
	Wrist fatigue	
	Arm fatigue	
	Shoulder fatigue	
	Finger fatigue	

Table 6.26: Subjective attributes of the Five-point Likert scale subjective assessment (Multi Direction Tapping Test - Context)

In terms of the user participants profile, their background information was collected, including gender, age, handedness (i.e. preferred domain right hand or left hand) and experience in using a touch base, mouse and stylus (i.e. number of years).

6.5.4 RESULT ANALYSIS AND DISCUSSION

6.5.4.1 TEST DATA

6.5.4.1.1 MULTI DIRECTIONAL DIRECTION TAPPING TEST

H1: There is a difference in the amount of speed using touch, mouse or stylus.

These data sets indicate outliers. Researcher experience challenges as one of the subjects felt asleep as the experiments is held on the orientation day. Consequently, these data are legitimately discarded as researcher is not interested in studying the subjects' reactions to particular phenomenon.

As stated by Osborne & Overbay (2004), "*unusual phenomena include construction noise outside a research lab or an experimenter feeling particularly grouchy, or even events outside the context of the research lab, such as a student protest, a rape or murder on campus, observations in a classroom the day before a big holiday recess, and so on can produce outliers. Faulty or non-calibrated equipment is another common cause of outliers that can be legitimately discarded if the researchers are not interested in studying the particular phenomenon in question".*

Table 6.27 indicates the Repeated Measure ANOVA and Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the movement time (*MT)* with all input devices (Mouse, Stylus and Touch) has no statistically significantly.

Devices	N	Task			P value
		1	2	3	
Mouse	22	24.93 (5.24)	24.52 (8.73)	22.81 (3.36)	.360**
Stylus	22	23.56 (3.98)	22.54 (2.09)	22.66 (4.93)	.310**
Touch	22	20.67 (1.72)	20.63 (1.92)	20.02 (1.79)	.186*

Table 6.27: The effect of the device difference (touch, mouse and stylus) on the movement time based on the Repeated Measure ANOVA test and Friedman test on every task (Multi Direction Tapping Test - Context)

** Repeated Measure ANOVA test*

*** Friedman test*

Figure 6.16 illustrate that the amount of speed of mouse and stylus have quite a similar tendency of movement time (*MT*) where it was fairly stable throughout the entire tasks. On the other hand, the touch recorded the fastest speed in all tasks. Through observation, participants show faster movement times using touch in the entire task because it was easier using touch in the tapping tasks.

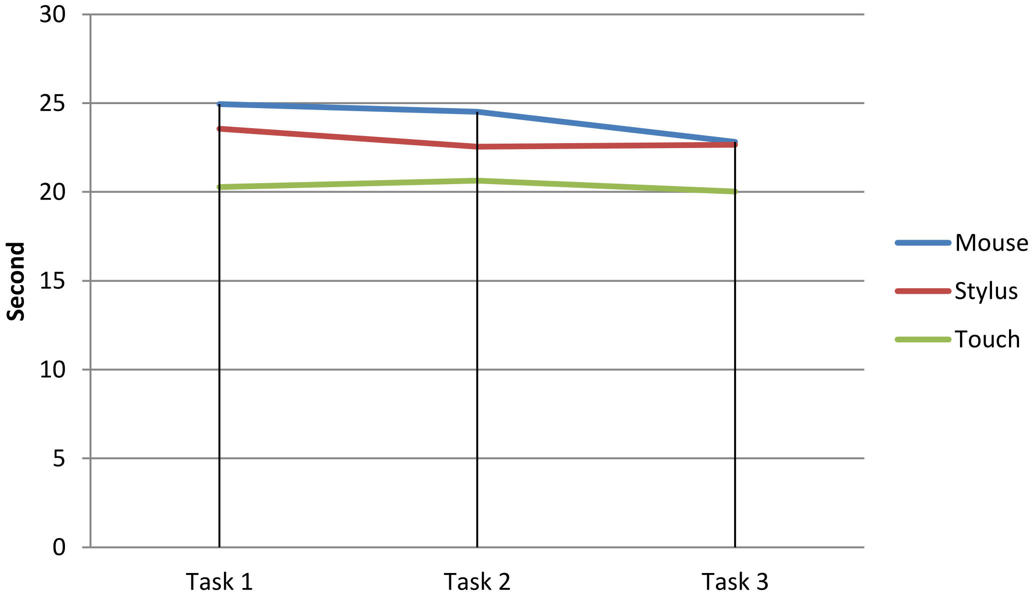


Figure 6.16: The amount of speed based on every task (Multi Direction Tapping Test - Context)

In terms of the overall movement time, a Friedman test was run to determine if there were differences in the amount of speed between the 3 inputs devices. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons.

The amount of speed was from mouse (*Mdn* = 22.59), to stylus (*Mdn* = 22.29), to touch (*Mdn* = 20.11), but the differences were not statistically significant, $\chi2(2) = 53.420$, $p < .0005$.

As can be seen in *Table 6.28*, the mean and standard deviation indicate that the fastest overall movement time (*MT)* is touch (20.31 sec ±1.80), followed by the stylus (22.92sec ±3.82) and then the mouse (24.09 sec ±6.16) which recorded the slowest among the devices.

Device	N	Mean	Std Deviation
Mouse	22	24.09	6.16
Stylus	22	22.92	3.82
Touch	22	20.31	1.80

Table 6.28: The effect of the device difference (touch, mouse and stylus) on the overall movement time (Multi Direction Tapping Test - Context)

H2: There is a difference in the amount of error using touch, mouse or stylus

As can be seen in *Table 6.29*, the Friedman test is applied on the raw material to examine the significance of the difference. As a result, it indicates that the error rate (*ER*) with the touch, mouse and stylus has no statistically significantly difference.

Devices	N	Task			P value
		1	2	3	
Mouse	22	.19 (.09)	.68 (.26)	.27 (.10)	.112
Stylus	22	.24 (12)	.27 (.18)	.27 (.18)	.741
Touch	22	.57 (.22)	.45 (.16)	.23 (.09)	.441

Table 6.29: The effect of the device difference (touch, mouse and stylus) on the error rates based on the Friedman test on every task (Multi Direction Tapping Test - Context)

Figure 6.17 shows that touch, stylus and mouse have a dissimilar tendency of error rates (*ER*).

There was a considerable fall in the error rates of participants using touch from Task 1 to Task 3. The mouse that recorded the least error rates among all the devices in Task 1 saw a significant rise of error in Task 2, producing the highest error rate, before falling steadily in Task 3. The stylus showed a slight increase of error rates in Task 2 and became relatively stable in the next two tasks. Through observation, participants showed a decrease in error at the end of the task using touch because the task is learnable and participants could predict the next move. Using the mouse, participants recorded the highest rate of error in task 2 because they were doing the task quickly and this led to inaccuracy; however, they managed to make less error in task 3. The stylus shows just slight differences in error rates in the tasks as participants were more careful and the fact that using stylus which has a sharp point, leads to accuracy.

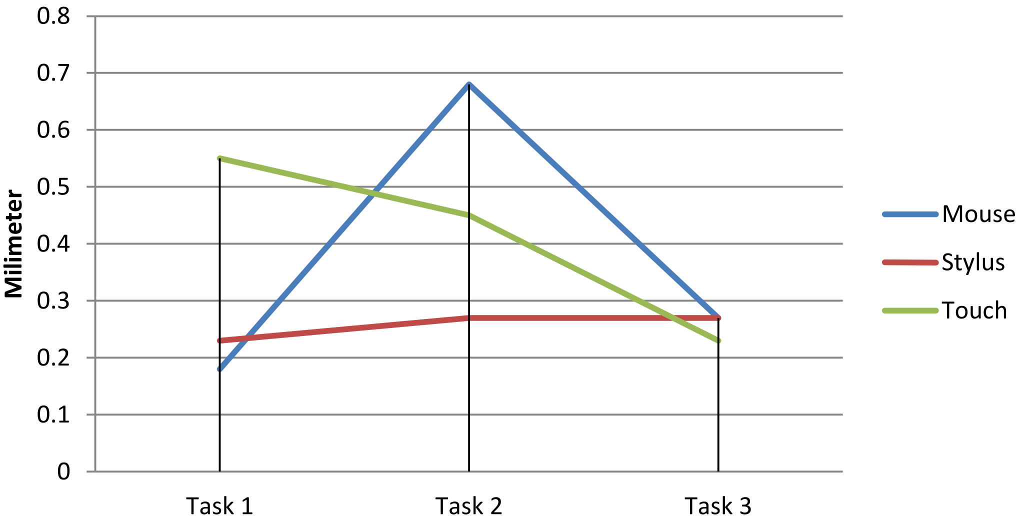


Figure 6.17: The error rates based on every task

A Friedman test was run to determine if there were differences in the amount of error between the 3 inputs devices. The median amount of error for all devices was 0 and the differences were not statistically significant, $\chi^2(2) = 2.714$, $p > .257$

Device	N	Mean	Std Deviation
Mouse	22	.38	.10
Stylus	22	.26	.07
Touch	22	.41	.09

Table 6.30: The effect of the device difference (touch, mouse and stylus) on the overall error rate (Multi Direction Tapping Test - Context)

The mean and standard deviation that has been shown in *Table 6.30* points out that the overall error rate (*ER*) of touch (.41 ±.09) is almost two-times higher than the mouse (.38 ±..10) which recorded the lowest error, while the stylus (.26 ±.07) showed the second highest error rate (*ER*).

6.5.4.2 USER SATISFACTION LEVEL

In this experiment a brief survey was conducted. Users were given forty-seven questions related to the functionality of the different input devices and asked to rate each input devices based on their experience of using it. Since 5 point 'Likert-type' scale (Likert, 1932) were used, a Friedman non-parametric test was done for the analysis as it is the best tool to analyse the result based on three different input devices.

6.5.4.2.1 MULTI DIRECTIONAL TAPPING TEST

H3: There is a difference in the participants' satisfaction level between the touch, mouse or stylus based interaction.

As can be seen in *Table 6.31*, the inter-reliability test discovered that inter-reliability about the design is very high with all the input devices, thus the comparison can be made for both devices in terms of the general indices and the fatigue indices, as summarised in *Table 6.32*.

Subjective feeling	Cronbach's Alpha		
	Mouse	Stylus	Touch
General Indices	.971	.979	.974
Fatigue Indices	.962	.968	.956

Table 6.31: Inter Reliability Statistics with Mouse, Stylus and Touch (Multi Direction Tapping Test - Context)

	Subjective Feeling	N	Devices			P value
			Mouse	Stylus	Touch	
	How do you consider the task of using?	23	4.52	3.78	4.87	.014
General Indices	Actuation force	23	4.35	3.70	4.48	.023
	Operation smoothness	23	4.30	3.70	4.48	.009
	Operation effort	23	4.39	3.83	4.48	.052
	Accuracy	23	4.17	3.61	4.52	.004*
	Operation speed	23	4.30	3.57	4.57	.002*
	General comfort	23	4.43	3.74	4.52	.006
	Overall operation	23	4.41	3.73	4.59	.006
Fatigue Indices	Finger fatigue	23	4.09	3.70	4.30	.006
	Wrist fatigue	23	4.22	3.74	4.43	.009
	Arm fatigue	23	4.00	3.52	4.13	.052
	Shoulder fatigue	23	3.87	3.48	4.04	.177
	Neck fatigue	23	4.04	3.48	4.17	.011
	Overall operation	23	3.95	3.59	4.18	.089

Table 6.32: The effect of the device difference (mouse, stylus and touch) on the subjective feelings based on the Friedman test on the raw data of the subjective assessment (Multi Direction Tapping Test - Context)

* The difference between the devices is statistically significant.

The Friedman test indicates that the participants considered it much easier using the touch (4.87) on the task given rather than the mouse (4.52) and stylus (3.78) ($p>0.05$). This indicates that touch is easier to use than the stylus and mouse in the multi directional tapping test.

In terms of the general indices, for all seven indicators (i.e. actuation force, operation smoothness, operation effort, accuracy, operation speed, general comfort and overall operation), touch was the highest rated. The stylus recorded the worst ratings, as the mean shows, for all seven indicators.

Similarly with regard to the fatigue indices, for all six indicators (i.e. finger fatigue, wrist fatigue, arm fatigue, shoulder fatigue, neck fatigue, overall operation), touch was favoured by the participants in the multi directional tapping test. The stylus was rated to be the worst input device for all of the indicators.

This led to the conclusion that touch came out best in the multi directional tapping test in relation to general and fatigue indices.

6.5.5 SUMMARY

6.5.5.1 Tracing and Dragging Tests

Based on the results analysis, the tracing test shows that although touch has the fastest Movement Time (*MT*) mean with 22.63 seconds, however, it has the highest Error Rate (*ER*) mean with 230.31 mm compared to the stylus and mouse.

The reason for the error as specified by the participants' feedback in terms of touch was related to the size of fingers, which might prove to be a hindrance in completing the task. This was backed up by participants showing higher error rates using touch because the tracing line was blocked by participants' finger that resulted in inaccuracy.

However, the dragging test shows that using the stylus has the fastest input with 21.81 seconds of Movement Time (*MT*) mean compared to the mouse and touch.

In relation to participants' satisfaction level between the touch, mouse and stylus for general and fatigue indices, both the stylus and mouse recorded better results in the tracing test (context), while in the dragging test (context), the mouse came out best. However, touch is the worst rated among all input devices in tracing and dragging tests (context). This suggests that both the stylus and mouse were rated highly in the tracing task, while touch is the least preferable.

6.5.5.2 One Direction Tapping and Multi-Directional Tapping Tests

Based on the result analysis, the one direction tapping test shows that although touch has the fastest Movement Time (*MT*) mean with 37.75 seconds, however, it has the highest Error Rate (*ER*) mean with 1.32 compared to the stylus and mouse.

Similarly, the multi-directional tapping test also shows that touch has the fastest input with 20.31 seconds Movement Time (*MT*) mean compared to the mouse and stylus. Yet again, touch has the highest Error Rate (*ER*) mean with 0.41 among the three input devices.

Through observation, participants' movement time of touch is faster because it is easier and faster tapping with a finger. The reason for the error was due to the fact that participants found the tasks easy and did them quickly which led to inaccuracy.

In relation to participants' satisfaction level between the touch, mouse and stylus for general and fatigue indices, touch came out best in both the one direction tapping test and the multi directional tapping test. However, the stylus had the worst rating among all input devices. This suggests that using touch is the most preferable input device in tapping tasks, while the stylus is the least preferable.

CHAPTER 7: DISCUSSION AND CONCLUSSIONS

7.1 SUMMARY OF THE RESEARCH AND MAIN FINDINGS

This study aimed to investigate the usability of touchscreen interfaces in terms of efficiency, effectiveness and user satisfaction via a series of empirical experiments that involved comparisons of three input devices: mouse, stylus and touch. To measure efficiency and effectiveness, movement time and error rates of participants were examined. To measure user satisfaction, their feelings regarding the design and their discomfort in the particular body regions were researched.

7.1.1 Experiment Task

The experiment had two tasks: an abstract task and a context task. The purpose of the abstract task was to analyse the users' ability on simple tasks without a real world context, while the context task's aim was to examine the users' ability in a real world context.

7.1.1.1 Abstract

The Tracing test consists of four circles, each with a diameter of 100 mm. The participants were instructed to draw a free-hand line using Touch, Stylus and Mouse around each of the circles

The Dragging test consists of circles with a diameter of 8 mm over a distance of 100 mm. Participants were asked to place them in circles with a diameter of 10 mm and perform the task in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.

The One direction tapping test consists of two rectangles with a defined width in the direction perpendicular to the direction of the movement. The task consists of alternately tapping between the two rectangles. The participants were instructed to point and click, along one axis, within each rectangle 25 times using Touch, Stylus and Mouse. This test consists of four tasks altogether in which the target becomes smaller and the distance greater as the difficulty of the task increases.

The Multi-directional tapping test consists of targets positioned around the circumference of a circle. The task requires alternately tapping around the circumference of a circle with twenty-five squares. The participants were instructed to point and click, along the circumference of a circle with tapping each of the squares using Touch, Stylus and Mouse. The targets to which the participant should advance were marked with X. This test consists of three tasks in which the target becomes smaller in relation to the increasing level of difficulty of each task.

7.1.1.2 Contextual

The Tracing test consists of a four route map, each with a diameter of 80 mm. The participants were instructed to draw a free-hand line in a mean to measure the distance of the route using touch, stylus and mouse around each of the circles in a clockwise direction.

The Dragging test is a drag test on an object namely a rectangle with a diameter of 30 mm over a distance of 140 mm and participants were asked to place them in a new target place The participants were also instructed to perform the task in all four cardinal directions (left-to-right, right-to-left, down, up), ten times in each direction.

The One direction tapping test requires alternately tapping between the two keyboard keys. The participants were instructed to point and click, along one axis, on each key for 25 times using Touch, Stylus and Mouse. This test consists of four tasks altogether in which at each level of task, the target becomes smaller and the distance greater in relation to the increase in difficulty.

The Multi-directional tapping test requires randomly tapping keyboard keys. The participants were instructed to point and click a keyboard key on a random highlighted key for 25 times using Touch, Stylus and Mouse. This test consists of three tasks in which the target becomes smaller as the task becomes more difficult.

7.1.2 Experiment Findings

7.1.2.1 Movement Time (Abstract Tasks)

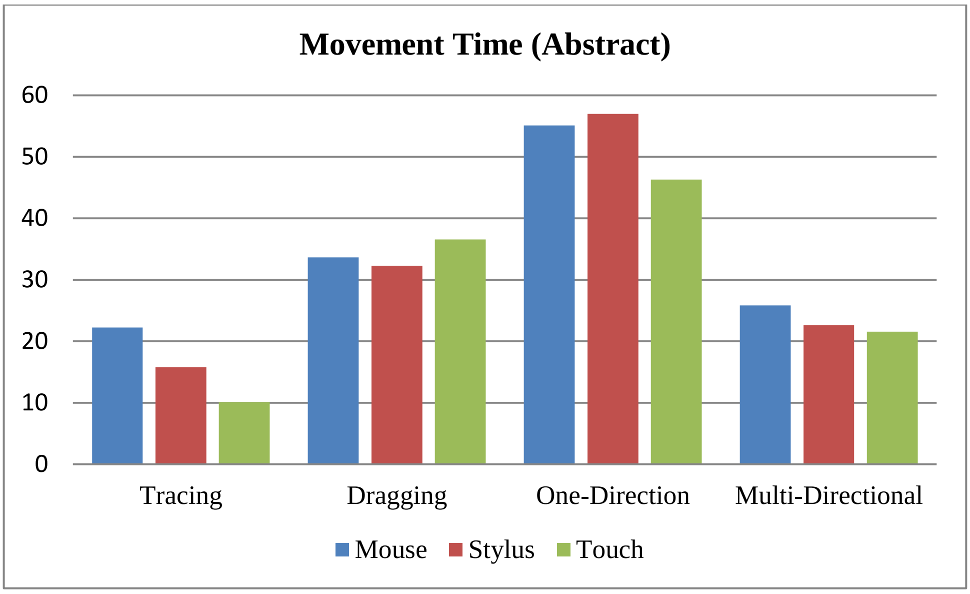


Figure 7.1: Movement Time (Abstract Experiment)

**low graph indicate the fastest movement time*

Figure 7.1 indicates that touch (10.08 sec ±8.50) has the fastest overall movement time (*MT*) in tracing test, that is two-times faster than the mouse (22.20 sec ±11.90) which recorded the slowest, while stylus (15.78 sec ±8.72) has the second fastest overall movement time (*MT*).

The slowest overall movement time (*MT*) in the dragging test is with touch (36.56 sec ±17.96), followed by the mouse (33.65 sec ±11.12) and then the stylus (32.25 sec ±14.45) which recorded the fastest among the devices.

The table also indicates that touch (46.28 sec ±10.18) has the fastest overall movement time (*MT*) in the one direction tapping test, that is two-times faster than the stylus

(56.98 sec ±17.61) which recorded the slowest, while the mouse (55.09 sec ±10.70) has the second fastest overall movement time (*MT*).

The slowest overall movement time (*MT*) in the multi-directional tapping test is the mouse (25.83 sec ±3.22), followed by the stylus (22.58 sec ±3.59) and then touch (21.53 sec ±3.86) which recorded the fastest among the devices.

7.1.2.2 Error Rate (Abstract Tasks)

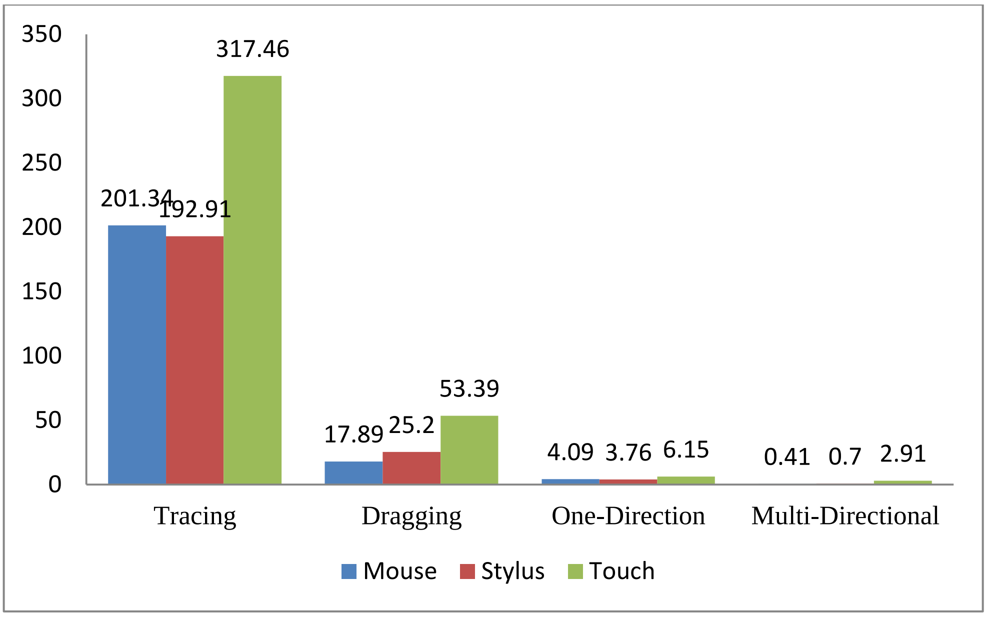


Figure 7.2: Error Rates (Abstract Experiment)

Figure 7.2 indicates that the overall error rate (*ER*) of touch (317.46mm ±304.327) in the tracing test is almost two-times greater than the stylus (192.91mm ±244.252) which recorded the lowest error among the three input devices, while the mouse (201.34mm ±265.5) showed the second highest error rate (*ER*).

The overall error rate (*ER*) of touch in the dragging test (53.39 mm ±15.32) is almost two-times higher than the mouse (17.89 mm ±7.60) which recorded the lowest error, while the stylus (25.20 mm ±10.90) showed the second highest error rate (*ER*).

The table also indicates that the overall error rate (*ER*) of touch (6.15 ±13.48) in the one direction tapping test is almost two-times greater than the stylus (3.76 ±8.82) which recorded the lowest error among the three input devices, while the mouse (4.09 ±10.24) showed the second highest error rate (*ER*).

The overall error rate (*ER*) of touch (2.91 ±3.09) in the multi-directional tapping test is almost five-times higher than the mouse (.41 ±.94) which recorded the lowest error, while the stylus (.70 ±1.07) showed the second highest error rate (*ER*).

7.1.2.3 *User Satisfaction (Abstract Tasks)*

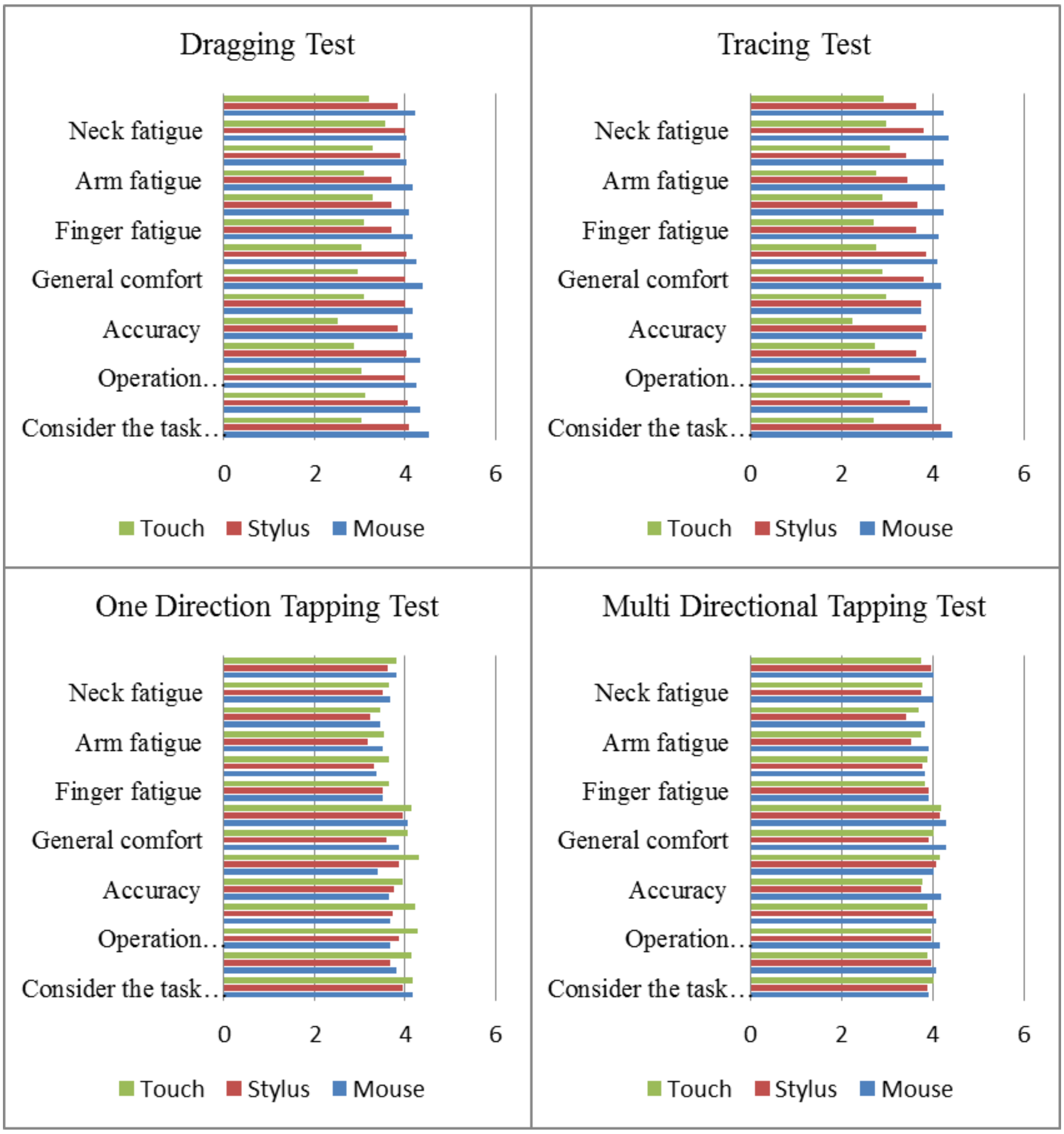


Figure 7.3: User Satisfaction (All Tests - Abstract)

In abstract tasks of the tracing and dragging tests, the mouse has greater user satisfaction. Participants indicated that it is more comfortable, faster, accurate, requires less effort, is a smooth process and uses less force except in the tracing test where the mouse and stylus have comparable user satisfaction in terms of accuracy and speed. The touch recorded the worst rated input in the tracing and dragging tests on all of the general indicators.

With regard to the fatigue indices, the results suggest using the mouse produces less fatigue for participants' wrist, arm, shoulder, finger, shoulder and neck. Users appear to get use to the wireless optical mouse used in the experiment. Their arms, hands, and wrists also are more relaxed while using mousepad. The results show that touch was rated the worst device for causing fatigue in all of the body parts indicators.

In abstract tasks of the one direction tapping test, touch has greater user satisfaction. Participants indicated that it is more comfortable, faster, accurate, uses less effort, is a smooth process and requires less force. On the other hand, in the multi directional tapping test the mouse has better user satisfaction as it is more comfortable, accurate, uses less effort, is a smooth process and requires less force. However in terms of speed participants think using the mouse is slower in the multi directional tapping test. The stylus and mouse comparably recorded the worst results in the one direction tapping test while stylus and touch comparably came out worst in the multi directional tapping test.

When considering the fatigue indices, the results suggest that using touch in the one direction tapping test produces the lowest level of fatigue for participants' finger, wrist, arm and shoulder. Participants think that using the mouse is less tiring for their shoulder and neck. The results also show that the stylus was rated the worst device in terms of causing fatigue in the wrist, arm, shoulder and neck, while the mouse causes finger fatigue.

In the multi directional tapping test, using the mouse is indicated as causing less fatigue for participants' finger, arm, shoulder and neck. The stylus is rated to be the worst input device for wrist, arm, shoulder and neck fatigue, while touch is the worst rated in finger fatigue.

7.1.2.4 Movement Time (Context Tasks)

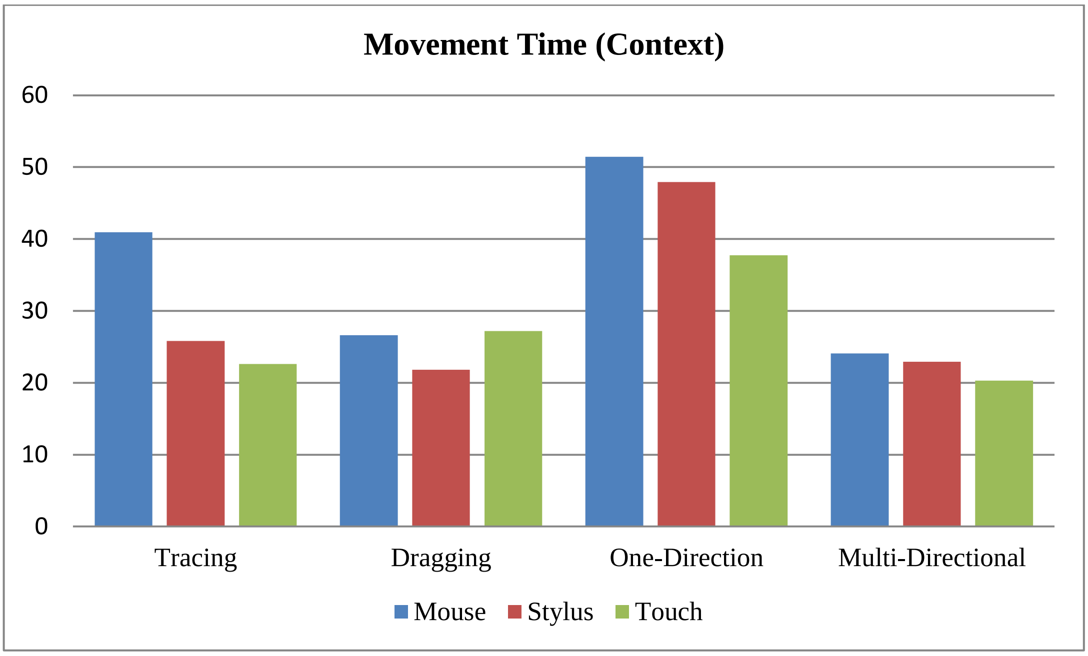


Figure 7.4: Movement Time (Context Experiment)

**low graph indicate the fastest movement time*

Figure 7.4 indicates that touch (22.63 sec ±15.09) has the fastest overall movement time (*MT*) in the tracing test, that is two-times faster than the mouse (40.95sec ±19.59) which recorded the slowest, while stylus (25.83 sec ±10.59) has the second fastest overall movement time (*MT*).

The slowest overall movement time (*MT*) is with touch (27.19 sec ±13.82) in the dragging test, followed by the mouse (26.63 sec ±14.47) and then the stylus (21.81 sec ±8.61) which recorded the fastest among the devices.

The table also indicate that touch (37.75 sec ±51.06) has the fastest overall movement time (*MT*) in the one direction tapping test, that is almost two-times faster than the mouse (51.45 sec ±45.35) which recorded the slowest, while the stylus (47.93 sec ±29.68) has the second fastest overall movement time (*MT*).

The fastest overall movement time (*MT*) is with touch (20.49 sec ±6.16) in the multi-directional tapping test, followed by the stylus (22.92 sec ±3.82) and then the mouse (20.31 sec ±1.80) which recorded the slowest among the devices.

7.1.2.5 Error Rate (Context Tasks)

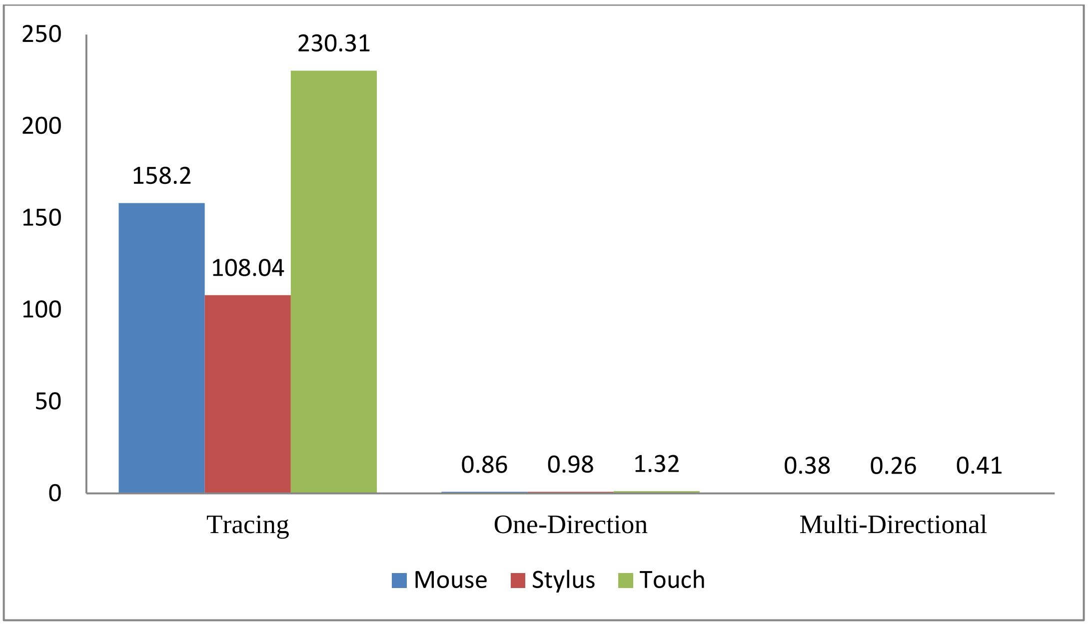


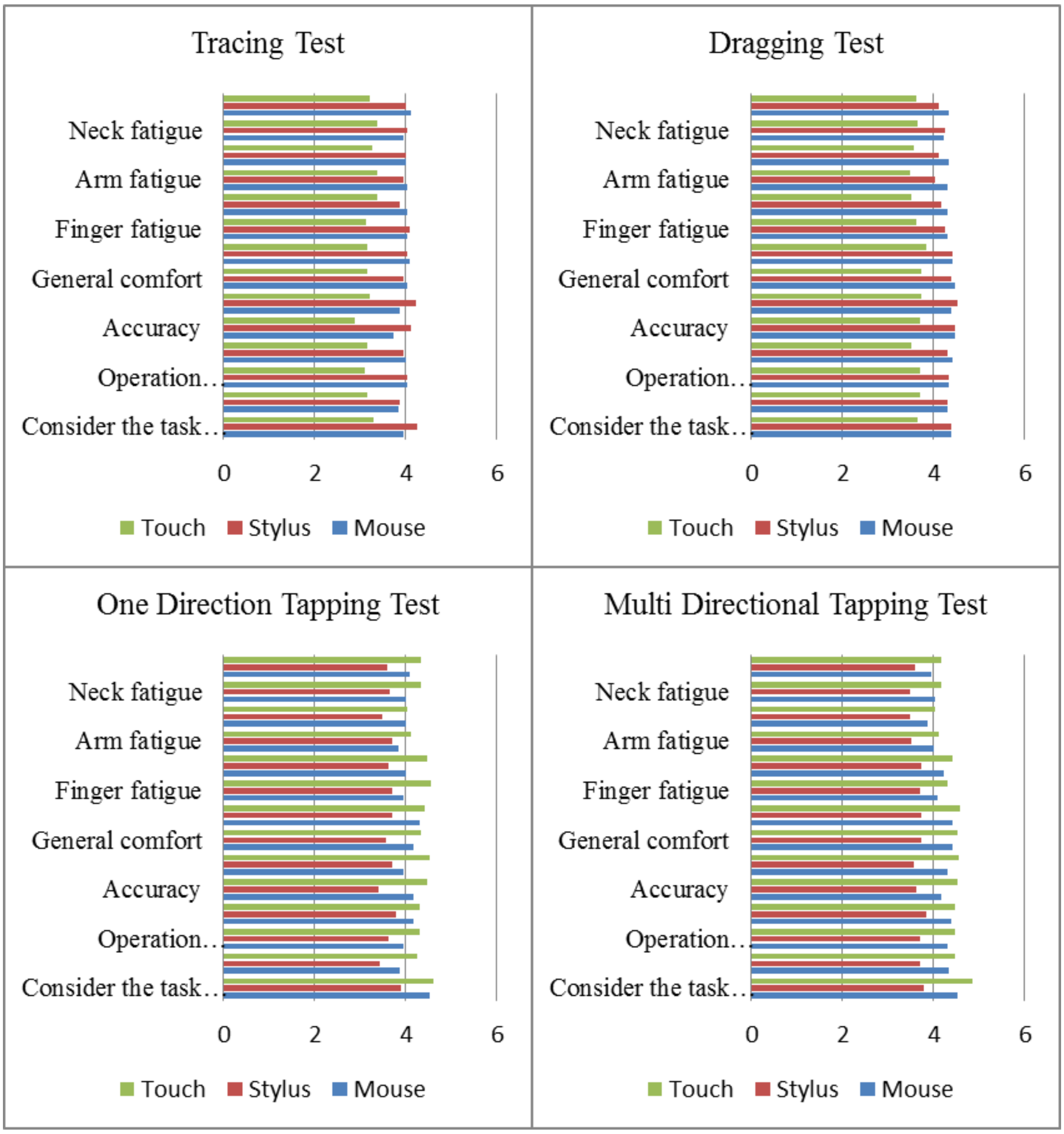
Figure 7.5: Error Rates (Context Experiment)

Figure 7.5 indicate that the overall error rate (*ER*) of touch (230.31 mm ±72.68) in the tracing test is almost two-times greater than the stylus (108.04 mm ±35.68) which recorded the lowest error among the three input devices, while the mouse (158.20 mm ±69.44) shows the second highest error rate (*ER*).

The overall error rate (*ER*) of touch (1.32 ±2.36) in the one direction tapping test is almost two-times greater than the mouse (.86 ±1.14) which recorded the lowest error among the three input devices, while the stylus (.98 ±2.13) showed the second highest error rate (*ER*).

The overall error rate (*ER*) of touch (.41 ±.09) in the multi-directional tapping test is almost two-times higher than the stylus (.26 ±.07) which recorded the lowest error, while the mouse (.38 ±.10) showed the second highest error rate (*ER*).

7.1.2.6 User Satisfaction (Context Tasks)


Tracing Test

Dragging Test

One Direction Tapping Test

Multi Directional Tapping Test

Figure 7.6: User Satisfaction (All Tests - Context)

In contextual tasks of the tracing and dragging tests, the mouse and stylus have comparable user satisfaction. Participants indicated that they are more comfortable, faster, accurate, use less effort, are a smooth process and require less force. The touch recorded the worst rated input to be used in the tracing and dragging test on all of the seven general indicators.

With regard to the fatigue indices, similarly the results suggested comparable fatigue using the mouse and stylus. Participants indicated that there was less fatigue for their wrist, arm, shoulder, finger, shoulder and neck using the mouse and stylus. However, the results show otherwise for touch as it was rated the worst for all fatigue indicators.

In the contextual task of the one direction tapping and multi directional tapping, the results specified a contrast in user satisfaction, as touch was highly rated as being better compared to the mouse and stylus. Participants indicated that in the tapping test using touch was more comfortable, faster, accurate, used less effort, was a smooth process and required less force. The stylus came out as the worst device in tapping tasks.

As for the fatigue indices, the results suggest less fatigue on participants' wrist, arm, shoulder, finger, shoulder and neck when using touch. The stylus is rated to be the worst input device for causing fatigue in the tapping tasks.

7.2 DISCUSSION

7.2.1 Human Performance

7.2.1.1 Tracing test

The findings in the tracing test indicate that touch input outperforms stylus and mouse in movement time (*MT*) for both abstract and context tasks. The touch is overall the fastest device however; it is also the most inaccurate of all devices tested in most of the tasks. It is not surprising in light of previous work by Zabramski *et al.* (2013) that indicate participants show faster movement time using touch as the tasks are learnable forwarding to the next level and participants could predict their next move which explains the fastest movement time, however leading to carelessness. It is likely that the task formulation forcing participants to perform "as fast and as accurate as possible" is responsible for creating different operational bias (Zhou & Ren, 2010). The finding that touch input performs worst of all input methods tested when it comes to participants' error can be explained by certain occurrences that take place during the experiment test. For example, participants' finger caused a big occlusion of the tracing area with the drawing finger occluding the most crucial area where the shape took place. Additionally, participants' distance from the screen may be as it is placed quite far away on the table requiring participants to have to reach for it. Furthermore, another reason for error, as specified by a few participants in their feedback in questionnaires was the unresponsive screen. However, there was no observable system latency but any potential effect of hardware/software's latency was balanced by the fact that the same PC tablet was used for all input methods, so it may be assumed that all results were affected equally.

The mouse on the other hand performed the worst in terms of movement time (*MT*) in both abstract and context tasks as it is harder to operate mouse movements in the tracing test, even though the participants reported their highest daily experience with it.

The findings also illustrate that the Stylus has the least error rate (*ER*) in abstract and context tracing tests as it is a pen-shaped instrument that has a sharp point which results in high accuracy in doing the tracing test compared to other input devices.

7.2.1.2 Dragging Test

Of the three devices tested in the dragging test for both abstract and context tasks, the stylus outperforms the mouse and touch in movement time (*MT*). The stylus is overall the fastest device and has the second highest inaccuracy input of all devices tested in abstract tasks. Performance of touch was poor for the dragging task in both abstract and context tasks and it is also the most inaccurate of all devices tested in most of the tasks. This can be explained by noting it was particularly difficult to drag the circles with a diameter of 8 mm over a distance of 100 mm and place them in circles with a diameter of 10 mm with a finger. Participants' finger caused a big occlusion of the circle with the finger occluding most of the diameter of the circle to be dragged, making it easier to accidentally place the circle on the wrong target. This was not the case with the stylus and mouse which have comparable accuracies as the stylus is a pen-shaped instrument that has a sharp point while the mouse has an arrow pointer which made it easier to navigate as it does not occlude the diameter of the circle.

7.2.1.3 One Direction Tapping

Findings show that in the one direction tapping test touch outperforms the stylus and mouse in movement time (*MT*) for both abstract and context task performance. The touch is overall the fastest device; it is also the least accurate of all devices tested in most of the tasks. The movement time of touch is faster because it is easier tapping with a finger in addition to the learnable and predictable task which explains the fastest movement time. The mouse shows the slowest amount of speed because tapping using the mouse could take more of the participants' time as they need to control the mouse to point the targeted area and then click it. This experiment also confirms that touch performs worst of all input methods tested for both abstract and context tasks in terms of participants' error although the error rate was comparable using the stylus and mouse. The reason for error can be explained by the level of difficulty that increases in each level of the task that leads to comparable error rates using the three inputs. Additionally, participants show carelessness in the following level of tasks, as they did well in the previous task.

7.2.1.4 Multi Direction Tapping

This experiment indicates that touch outperforms the mouse and stylus for the multi direction tapping test in context and abstract tasks, although the touch performance was comparable to using a stylus and mouse. The comparable performance could be explained that in abstract and context tasks participants were more relaxed using the three inputs as the moves were predictable. However, touch still outperforms the other two devices tested in tapping tasks as it is naturally faster and easier to move quickly back and forth using a finger.

Likewise, although the error rate was comparable using all the three inputs, touch is the least accurate input of all devices tested in most of the tasks. In abstract tasks participants show higher error rates using touch because the target circumference come to be smaller towards the end of the tasks and causes the participant to tap out of the target especially if participants' finger caused a big occlusion of the targeted square. This is not in the case of Stylus and Mouse which show least error because the stylus has a pointed tip and the mouse has an arrow pointer which make it easier to point and lead to precise performance as they only occlude a small circumference of the targeted square. With regard to context tasks participants show carelessness in pointing and clicking the keyboard key as they can predict the next move and this leads them to move quickly back and forth between the two keys.

7.2.2 User Satisfaction

It can be concluded from the questionnaire data that the IS0 subjective comfort assessment shows diverse results in the tracing, dragging, one direction tapping and multi directional tapping tests using the three devices. For example,

- Mouse has better user satisfaction in abstract tasks of the tracing and dragging test
- Touch has better user satisfaction in abstract tasks of the one direction tapping test
- Mouse has better user satisfaction in abstract tasks of the multi directional tapping test
- Mouse and stylus have comparable user satisfaction in contextual tasks of the tracing and dragging test
- Touch has better user satisfaction in contextual tasks of the one direction tapping and multi directional tapping test

Given that different tests favour different output, it might be suggested that the difference was reliable enough and can reflect the differences in participants' satisfaction.

However, the most common input device that tended to be rated the worst in terms of user satisfaction is touch. For example,

- Touch recorded the worst rated input to be used in the tracing and dragging test (abstract) on all of the general indicators

- Touch was rated the worst device that caused fatigue in all of the body parts indicator in the tracing and dragging test (abstract)
- Stylus and touch comparably recorded the worst as input devices in the multi directional tapping test (abstract)
- Touch was the worst rated in finger fatigue in the multi directional tapping test (abstract).
- Touch recorded the worst rating in the tracing and dragging test (context) on all of general indicators
- Touch was rated the worst for all fatigue indicators in the tracing and dragging test (context)

It is suggested that the touchscreen and participant position when using touch in the experiment may be responsible for the worst results in terms of fatigue. Touchscreen and participant position biases may exist and impact users' performance. Ahlström *et al*. (1992) have studied that touchscreen inclination (0°, 22,5°, 30° 45°, 60° and 90° from horizontal) can result in a different level of fatigue and adversely affect user preference using touch. Ahlström *et al* (1992) concluded that 22.5° was least fatiguing and the inclination of 30° got the highest preference ratings, however, 90° was the poorest inclination with respect to fatigue.

In this experiment the touchscreen inclination was 60° and this explained the worst satisfaction level and fatigue for those using touch.

7.2.3 ISO Standard (IS0 9241)

One of the aims of using ISO standard in this experiment was to compare the performance of mouse, stylus and touch as it provides the usability requirement and human-system interaction.

As regards the questionnaire, an additional open-ended questionnaire to ascertain the problems faced by the participants during the tests is recommended. This will allow researchers to evaluate specific problems during the experiment.

In chapter 4, accuracy for the dragging test that shows a perfect hit of a circle placed in the centre of the target in which the error will be recorded in millimetres (each of the grid scale indicates one mm), if the circle were placed away from the perfect hit of the target is not an ISO recommended measure. It is because it is more useful to collect the error data in detail rather than assess the accuracy using score.

The IS0 standard does not make clear the calculation of overall usability of types of input devices in consideration as stated in ISO standard (ISO 9241-420, 2011 pg.15). Given these limitations, it is useful to have standardized guidance to evaluate usability.

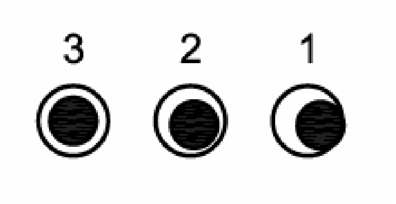


Figure 7.7: ISO recommendation for assessing accuracy: 3" for a perfect hit, "2" for a near-miss and "1" for a miss

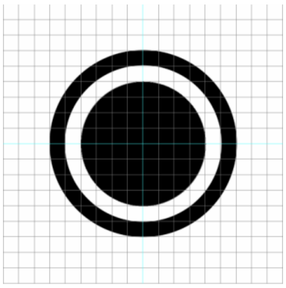
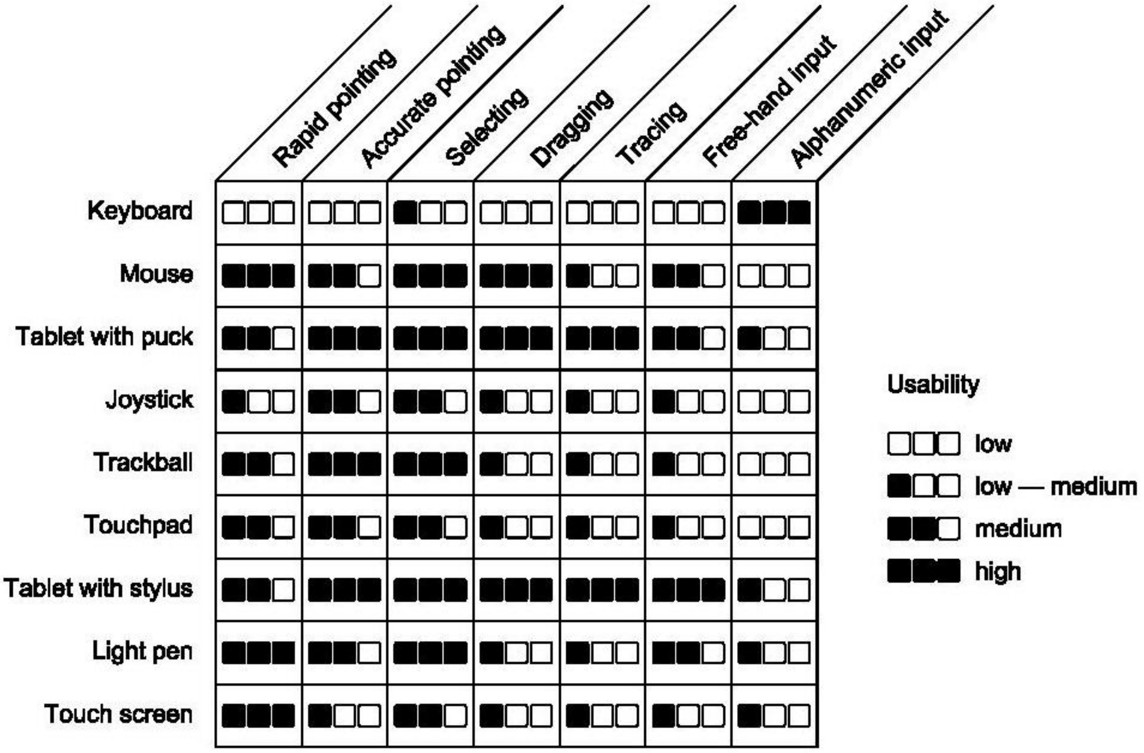


Figure 7.8: Assessing accuracy for the dragging test (each of the grid scale indicates one mm)

This study contribution also appears to be a well-evidenced recommendation to change the ISO 9241-420 (2011) diagram below (*see Figure 7.9*). This diagram could be compared with the findings from the study. In this study, findings are divided into three tables for clearer result: 1) Efficiency: Movement Time (*MT*) (*see Figure 7.10*), 2) Effectiveness: Error Rates (*ER*) (*see Figure 7.11*) and 3) Satisfaction: User Preference (*see Figure 7.12*).



	Rapid pointing	Accurate pointing	Selecting	Dragging	Tracing	Free-hand input	Alphanumeric input
Keyboard	□□□	□□□	■□□	□□□	□□□	□□□	■■■
Mouse	■■■	■■□	■■■	■■■	■□□	■■□	□□□
Tablet with puck	■■□	■■■	■■■	■■■	■■■	■■□	■□□
Joystick	■□□	■■□	■■□	■□□	■□□	■□□	□□□
Trackball	■■□	■■■	■■■	■□□	■□□	■□□	□□□
Touchpad	■■□	■■□	■■□	■□□	■□□	■□□	□□□
Tablet with stylus	■■□	■■■	■■■	■■■	■■■	■■■	■□□
Light pen	■■■	■■□	■■■	■□□	■□□	■■□	■□□
Touch screen	■■■	■□□	■■□	■□□	■□□	■□□	■□□

Usability

□□□ low

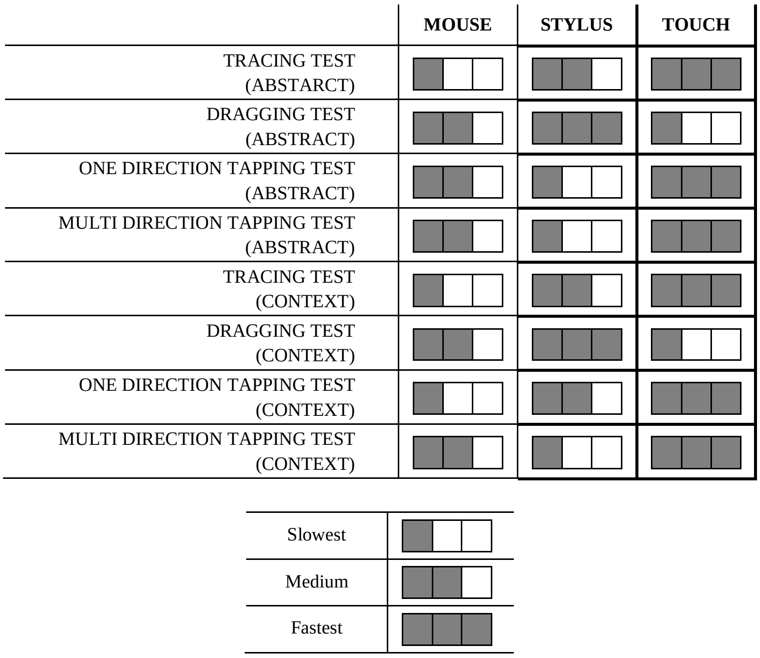
■□□ low — medium

■■□ medium

■■■ high

Figure 7.9: Overall usability of types of input devices in consideration of task principles and relevant aspects
Source: BS EN ISO 9241-420 (2011)

MOVEMENT TIME



	MOUSE	STYLUS	TOUCH
TRACING TEST (ABSTARCT)	■□□	■■□	■■■
DRAGGING TEST (ABSTRACT)	■■□	■■■	■□□
ONE DIRECTION TAPPING TEST (ABSTRACT)	■■□	■□□	■■■
MULTI DIRECTION TAPPING TEST (ABSTRACT)	■■□	■□□	■■■
TRACING TEST (CONTEXT)	■□□	■■□	■■■
DRAGGING TEST (CONTEXT)	■■□	■■■	■□□
ONE DIRECTION TAPPING TEST (CONTEXT)	■□□	■■□	■■■
MULTI DIRECTION TAPPING TEST (CONTEXT)	■■□	■□□	■■■

Slowest	■□□
Medium	■■□
Fastest	■■■

Figure 7.10: Overall Movement Time (MT) of types of input devices

ERROR RATE

	MOUSE	**STYLUS**	**TOUCH**
TRACING TEST (ABSTARCT)	■■□	■□□	■■■
DRAGGING TEST (ABSTRACT)	■■□	■□□	■■■
ONE DIRECTION TAPPING TEST (ABSTRACT)	■■□	■□□	■■■
MULTI DIRECTION TAPPING TEST (ABSTRACT)	■□□	■■□	■■■
TRACING TEST (CONTEXT)	■■□	■□□	■■■
DRAGGING TEST (CONTEXT)	■□□	■■□	■■■
ONE DIRECTION TAPPING TEST (CONTEXT)	■□□	■■□	■■■
MULTI DIRECTION TAPPING TEST (CONTEXT)	■■□	■□□	■■■

Least	■□□
Medium	■■□
Most	■■■

Figure 7.11: Overall Move Error of types (ER) of input devices

USER SATISFACTION

	MOUSE	**STYLUS**	**TOUCH**
TRACING TEST (ABSTARCT)	■■■	■■□	■□□
DRAGGING TEST (ABSTRACT)	■■■	■■□	■□□
ONE DIRECTION TAPPING TEST (ABSTRACT)	■■□	■□□	■■■
MULTI DIRECTION TAPPING TEST (ABSTRACT)	■■■	■□□	■■□
TRACING TEST (CONTEXT)	■■□	■■■	■□□
DRAGGING TEST (CONTEXT)	■■■	■■□	■□□
ONE DIRECTION TAPPING TEST (CONTEXT)	■■□	■□□	■■■
MULTI DIRECTION TAPPING TEST (CONTEXT)	■■■	■□□	■■□

Most Preferable	■□□
Medium	■■□
Least Preferable	■■■

Figure 7.12: Overall User Satisfaction of types of input devices

This study finding also recommends change to *Figure 7.13*. in selecting the correct device as recommended by ISO 9241-420 (2011). *Figure 7.14* is the structogram for selecting input devices as a result of this study.

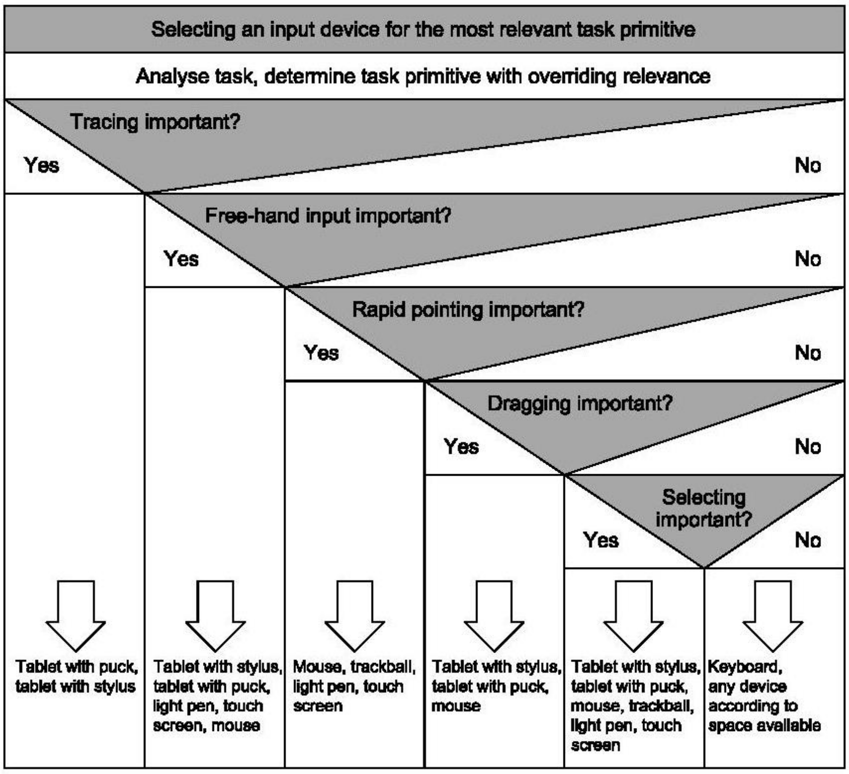


Figure 7.13: Structogram for selecting input devices in consideration of most relevant task primitive
Source: BS EN ISO 9241-420 (2011)

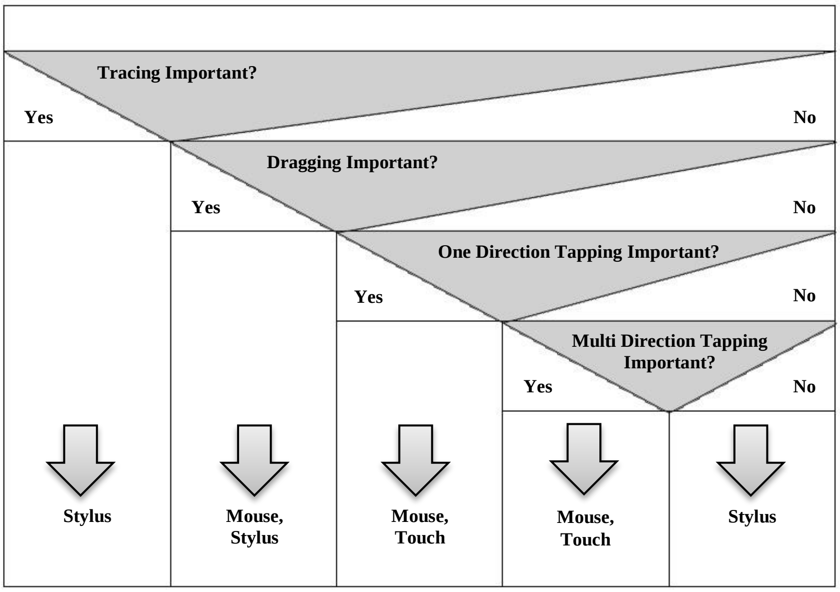


Figure 7.14: Structogram for selecting input devices in this study

This study also finds out that, ISO 9241-420 (2011) as it currently stands does not incorporate multi-touch, yet users habitually use this in most touch-based interactions. The tests of and measurements for satisfaction is less sufficient for current users' actual activities. In addition, ISO 9241-410 (2008) does not specify the categories that are appropriate for devices as according to the concept of usability (i.e. a product has no inherent usability).

CHAPTER 8: EVALUATION AND IDENTIFICATION OF FUTURE WORK

8.1 STRENGTHS AND LIMITATIONS OF THE STUDY

This chapter discusses the strengths and limitations of the study. It further identifies new research potential. This study can also be used as a direct source of reference and guidance for users and system designers to use during the design process.

8.1.1 Strengths

The strength of the studies lies in its contribution. This section discusses the main contribution of this research: theoretical, methodological and professional.

8.1.1.1 Theoretical Contribution

The findings of the study have contributed to an enriched understanding of the ISO. This is the main strength of the current study as it provides findings that are based on ISO standard (ISO 9241-420, 2011). It provided a useful guideline as guidance to users and system designers that can be further developed and applied to other research in this area.

Furthermore, there have been few studies of human performance and user satisfaction in literature that compare three input devices (stylus, mouse and touch) in tracing, dragging, one direction tapping and multi directional tapping test for both abstract and context tasks. This is the first study that has been carried out and therefore contributes to the up-to-date HCI literature.

8.1.1.2 Methodological Contribution

In my research analysis of the data, this study uses the three-way comparison of three input devices (stylus, mouse and touch). It is part of the contribution of this study as no research has done this way. This study can be as an additional methodological contribution to the existing literature as this thesis adopts an experiment from well-designed experiment research approach.

8.1.1.3 Professional Contribution

There have been very limited academic studies of HCI conducted previously in Brunei Darussalam. Clearly, the main strength of the current study is that it provides findings from well-designed implemented research on HCI in real-world-type tasks and scenarios. The study has also yielded up-to-date information on whether touch-based interaction is more effective and preferred by users.

8.1.2 Limitation

This study supported the objective of this study; however this effort did have limitations.

1. Younger Student Age Group: Limitations of the research probably lay in the scope of volunteers as research participants. This study had almost all subjects in the younger age group that is between 18 – 25 and 26 – 39, which has implications for generalisability. This is due to the reason that in this research, it was difficult to identify people who would participate in the experiment. This issue also delayed the data collection progress.

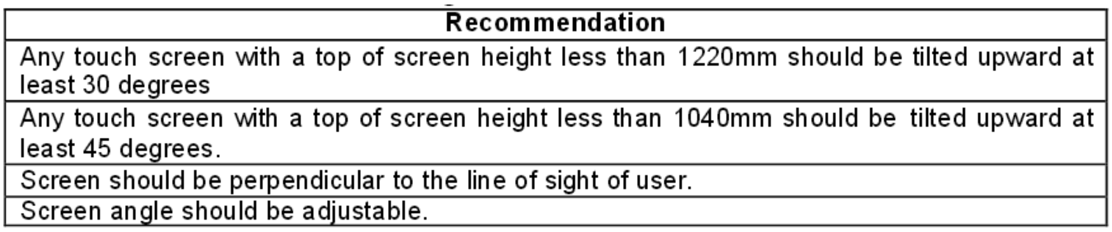
2. Lack of Randomisation in Subjects' Activities: Due to some technical limitations with the experimental software (i.e. 2 experiments in one prototype), the order in which the subjects did the experiments was not randomized. The ordering could have had an impact to reduce *Order Effect*.

3. Lack of Stylus Experience: Another main limitation to be acknowledged in this experiment is that these experiments fail to consider lack of stylus experience. This study should have considered this in the same way as touch experience.

4. Inconsistent with Relevant Ergonomic Guidelines: This study has several measures that are inconsistent with relevant ergonomic guidelines.

The written questionnaire asking participants to rate their experience using the device is by means of 5 point Likert, rather than 7 point as specified in the ISO Standard. Researcher thought that a higher point Likert scale makes it more time consuming for the subject answering the 47 questions to take decision especially they are answering immediately after the experiment were held. This is also considered as limitation as it does not following the ISO standard guidelines.

Another limitation is that this study used the touchscreen device at angles of 60-degree laptop mode for touchscreen which is inconsistent with relevant ergonomic guidelines (*see Table 8.1*).



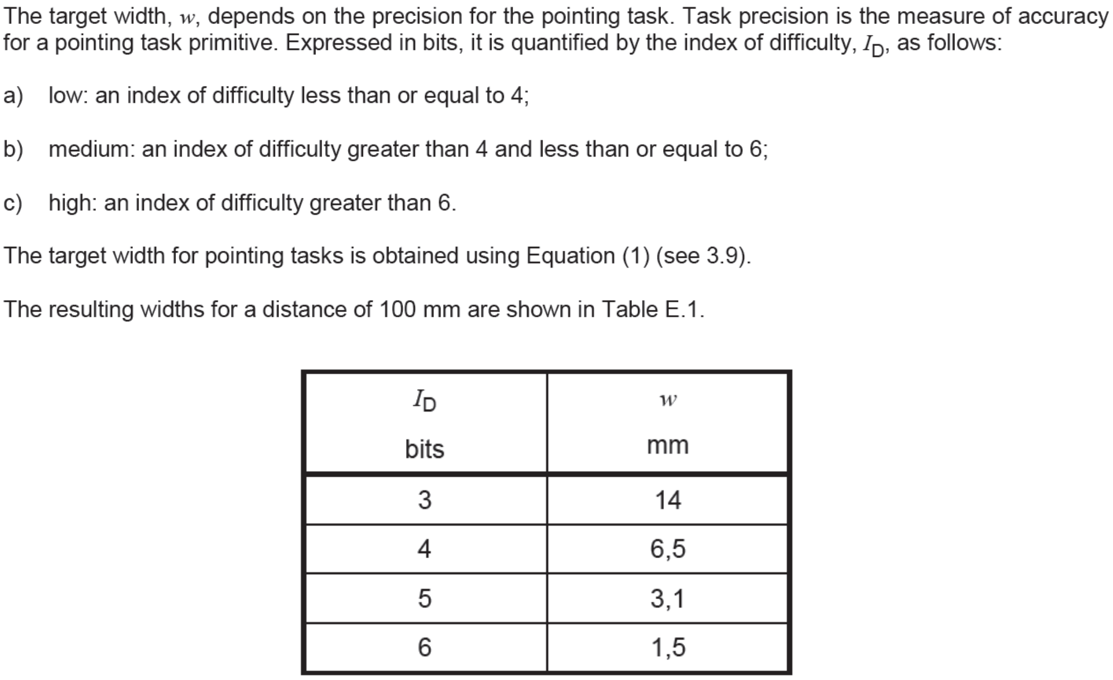
Recommendation
Any touch screen with a top of screen height less than 1220mm should be tilted upward at least 30 degrees
Any touch screen with a top of screen height less than 1040mm should be tilted upward at least 45 degrees.
Screen should be perpendicular to the line of sight of user.
Screen angle should be adjustable.

Table 8.1: Recommendations for angle of view for touchscreens
Source: Swann, M. (2006).

In tracing test, the ISO standard refers to a mix of clockwise and counter-clockwise. However, in this experiment, it only did experiment for the counter-clockwise as researcher overlooked the counter clockwise. Therefore, this is acknowledges as limitation.

In one direction tapping, researcher initially used the I_D equation to calculate the target sizes as stated by ISO standard *(see Table 8.2*), however, due to the resolution of the screen used in this study, researcher modified the target width. Researcher also increases target distance to create more difficulty level which are not stated in

ISO standard (*see Table 8.3*). This is likewise acknowledges as another limitation inconsistent with relevant ergonomic guidelines.



The target width, w, depends on the precision for the pointing task. Task precision is the measure of accuracy for a pointing task primitive. Expressed in bits, it is quantified by the index of difficulty, I_D, as follows:

a) low: an index of difficulty less than or equal to 4;

b) medium: an index of difficulty greater than 4 and less than or equal to 6;

c) high: an index of difficulty greater than 6.

The target width for pointing tasks is obtained using Equation (1) (see 3.9).

The resulting widths for a distance of 100 mm are shown in Table E.1.

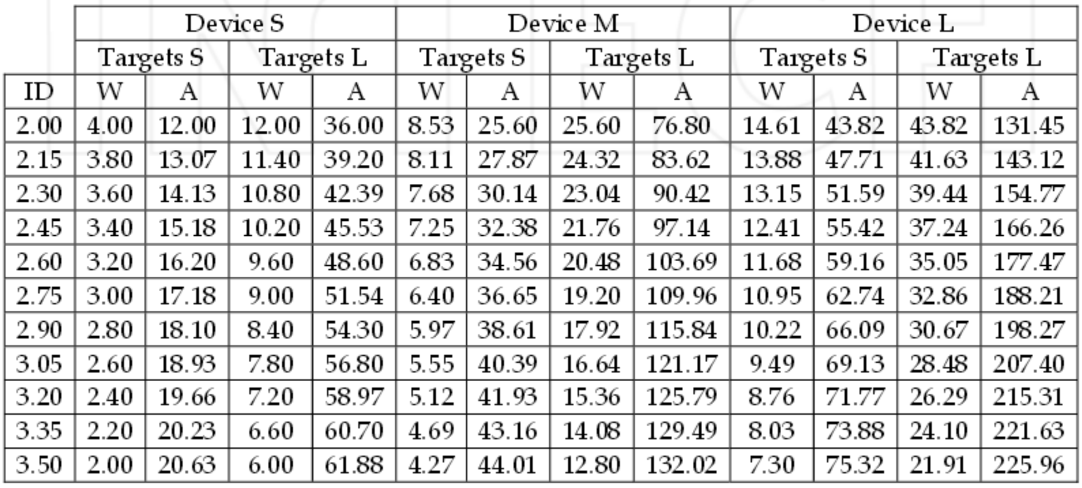
I_D bits	w mm
3	14
4	6,5
5	3,1
6	1,5

Table 8.2: Index of difficulty and target width for a distance of 100 mm

I_D bits	W mm	D mm
3	3	30
4	5	80
5	9	135
6	12	170

Table 8.3: Index of difficulty and target width in one direction tapping

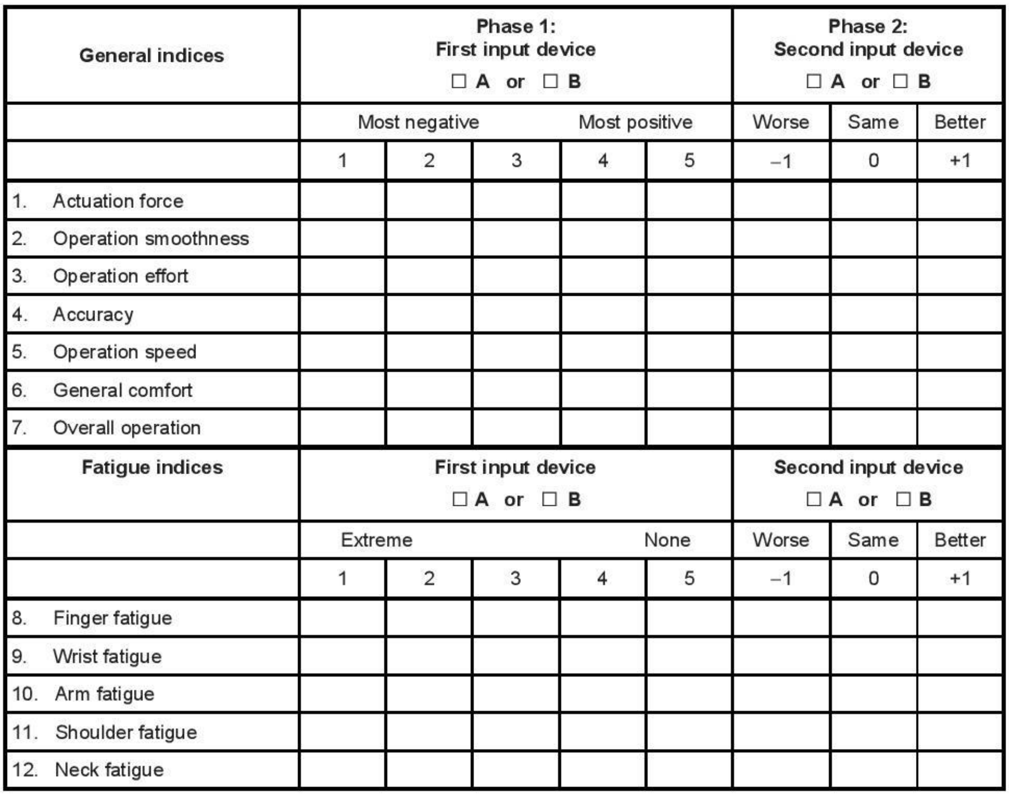
The ISO standard recommends plotting I_D against time for both one and multi tap. Researcher should consult ISO/TR 9241-411 for more refined tests. Researcher should correspondingly refer to research studies that include Index of Difficulty and Screen Size Variations such as study by Okada & Akiba (2010) that discuss *“Fitts’ Law Index of Difficulty Evaluated and Extended for Screen Size Variations”*. *Table 8.4* is the recommended target sizes and distances specified by Okada & Akiba (2010).



	Device S				Device M				Device L			
	Targets S		Targets L		Targets S		Targets L		Targets S		Targets L	
ID	W	A	W	A	W	A	W	A	W	A	W	A
2.00	4.00	12.00	12.00	36.00	8.53	25.60	25.60	76.80	14.61	43.82	43.82	131.45
2.15	3.80	13.07	11.40	39.20	8.11	27.87	24.32	83.62	13.88	47.71	41.63	143.12
2.30	3.60	14.13	10.80	42.39	7.68	30.14	23.04	90.42	13.15	51.59	39.44	154.77
2.45	3.40	15.18	10.20	45.53	7.25	32.38	21.76	97.14	12.41	55.42	37.24	166.26
2.60	3.20	16.20	9.60	48.60	6.83	34.56	20.48	103.69	11.68	59.16	35.05	177.47
2.75	3.00	17.18	9.00	51.54	6.40	36.65	19.20	109.96	10.95	62.74	32.86	188.21
2.90	2.80	18.10	8.40	54.30	5.97	38.61	17.92	115.84	10.22	66.09	30.67	198.27
3.05	2.60	18.93	7.80	56.80	5.55	40.39	16.64	121.17	9.49	69.13	28.48	207.40
3.20	2.40	19.66	7.20	58.97	5.12	41.93	15.36	125.79	8.76	71.77	26.29	215.31
3.35	2.20	20.23	6.60	60.70	4.69	43.16	14.08	129.49	8.03	73.88	24.10	221.63
3.50	2.00	20.63	6.00	61.88	4.27	44.01	12.80	132.02	7.30	75.32	21.91	225.96

Table 8.4: Target sizes and distances
Source: Okada, & Akiba (2010)

Researcher should also have three parallel comparisons contrasting two. It is also indicated as another limitation and as an item to address in future work, and present the analysis of the data without simply averaging. Researcher will ensure that those statistical analyses do not "regress to the mean"



General indices	Phase 1: First input device ☐ A or ☐ B					Phase 2: Second input device ☐ A or ☐ B		
	Most negative			Most positive		Worse	Same	Better
	1	2	3	4	5	–1	0	+1
1. Actuation force								
2. Operation smoothness								
3. Operation effort								
4. Accuracy								
5. Operation speed								
6. General comfort								
7. Overall operation								
Fatigue indices	**First input device** ☐ **A** or ☐ **B**					**Second input device** ☐ **A** or ☐ **B**		
	Extreme				None	Worse	Same	Better
	1	2	3	4	5	–1	0	+1
8. Finger fatigue								
9. Wrist fatigue								
10. Arm fatigue								
11. Shoulder fatigue								
12. Neck fatigue								

Figure 8.1: Dependent rating scale
Source: BS EN ISO 9241-420 (2011)

8.2 FUTURE WORK

The experiment showed a clear difference with devices in movement time in tracing, dragging, one direction tapping and multi direction tapping tasks. Clearly, the work is not complete, and issues such as extending ISO standard experiment testing to accommodate researchers need further investigation.

The identified previous limitations in this study will be the proposed further detailed studies such as:

1. Taking consideration of body posture and touchscreen position when using touch input, especially since there might be additional muscle strain due to the lack of proper support for the user’s arms during the use of touch devices.
2. Adding open-ended questions to understand better the users' previous experience and the context of use when they used each of the technologies or equivalents. I will suggest adequate wording for these questions.

3. Control for *Order Effect*
4. Carry out 3x2 comparison studies (as specified in the ISO standard) rather than three-way
5. Supply more meaningful real world activities for the contextual studies. For example researcher mentioned that other studies researchers such as Ezor (2010) had used gamification to motivate users to carry out more typical actions.
6. Include observational studies to focus on specific interface problems in real world situations and, from those, define activities for the contextual studies.
7. Include a more diverse group of participation with different levels of age and organization.
8. Consider touchpad since users might typically use that more often than a separate mouse
9. Consider Index of Difficulty and Screen Size Variations and consult ISO/TR 9241-411 for more refined tests regarding plotting I_D against time for both one and multi tap tests.

APPENDIX A: EXPERIMENTAL MATERIALS

A1: Experiment Questionnaire Participants' Selections Cover (Pilot Experiment)

University of
Salford
MANCHESTER

Participants' Selection Questionnaire

The purpose of the questionnaire is to select 20 – 30 participants to participate in touch-based experiment.

The objective of this experiment is to investigate the effectiveness and user satisfaction of using touch screen interfaces. This includes determining the direct input for selection and dragging operations on a touch-based PC.

The experiment task is designed using a simple tapping-dragging task to determine the speed and accuracy of the movement object. There will be 2 boxes with 3 different sizes (small, medium, large) and the task will be tap plus drag for the primary box and place on to target box, then release. Each size of the boxes will be tested on 6 different target box locations, which means 3 different sized boxes x 6 different target box locations equalling 18 tests. The participant will use Finger tips, Stylus and Mouse in order to perform the test.

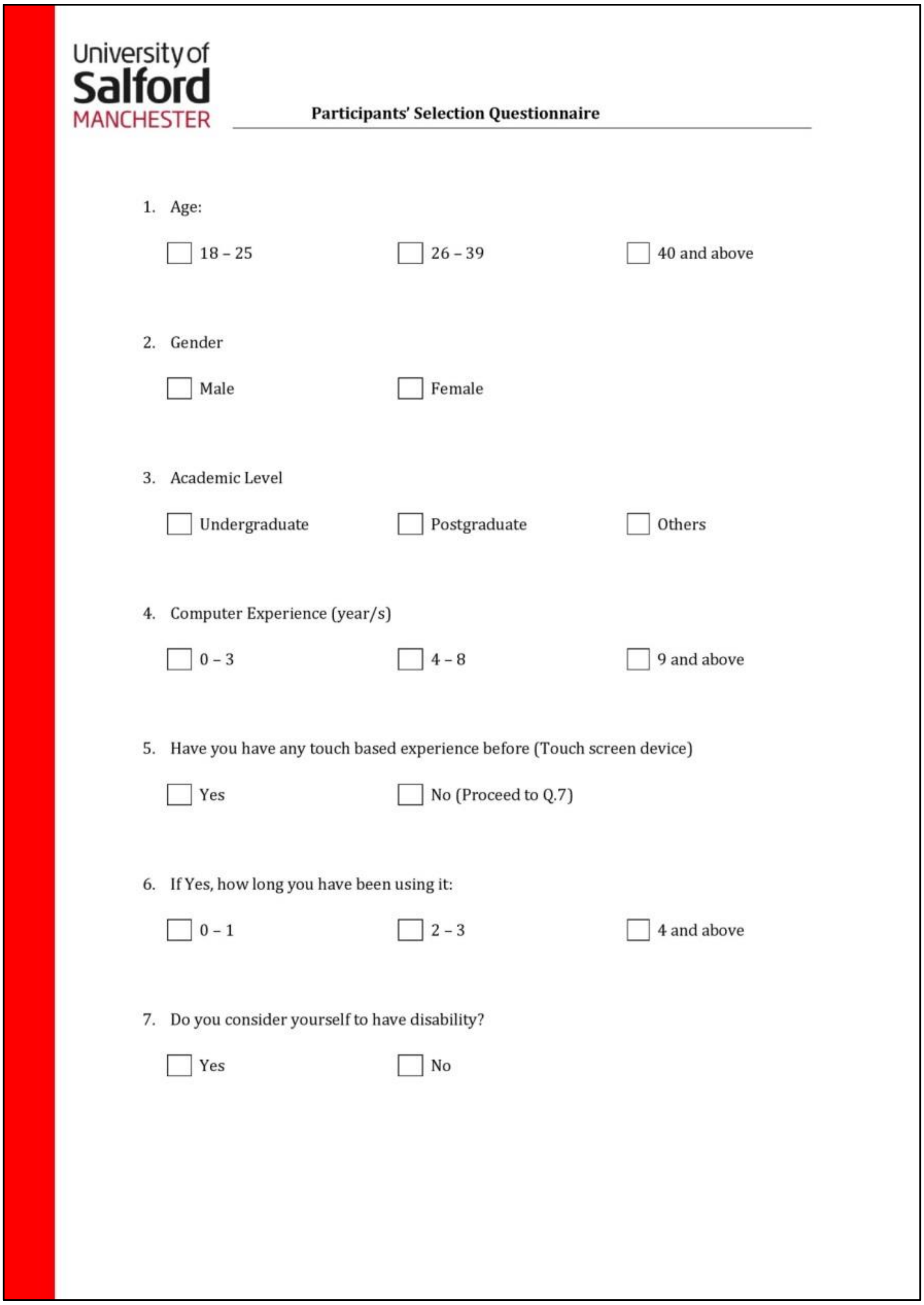
The significant contribution to knowledge that will arise from this research will be the gaining of some evidence to show if touch-based interaction is more effective and preferred by users in real world type tasks/scenarios.

This questionnaire should take approximately 5-10 minutes to complete. The results of the data may be published; however, no names or identifying information will be included in the final document.
May I thank you, in advance, for your valuable co-operation.

Hj Mohd Abul Fadle Hj Maidin
PhD Student,
School of Computing, Science & Engineering, (College of Science & Technology)
The University of Salford
The Crescent, Salford, M5 4WT, United Kingdom
Email: m.a.f.maidin@edu.salford.ac.uk

Telephone Number: +44 7511845478 / +6738895855

A2: Experiment Participants’ Selections Questionnaire



University of
Salford
MANCHESTER

Participants’ Selection Questionnaire

1. Age:

☐ 18 – 25 ☐ 26 – 39 ☐ 40 and above

2. Gender

☐ Male ☐ Female

3. Academic Level

☐ Undergraduate ☐ Postgraduate ☐ Others

4. Computer Experience (year/s)

☐ 0 – 3 ☐ 4 – 8 ☐ 9 and above

5. Have you have any touch based experience before (Touch screen device)

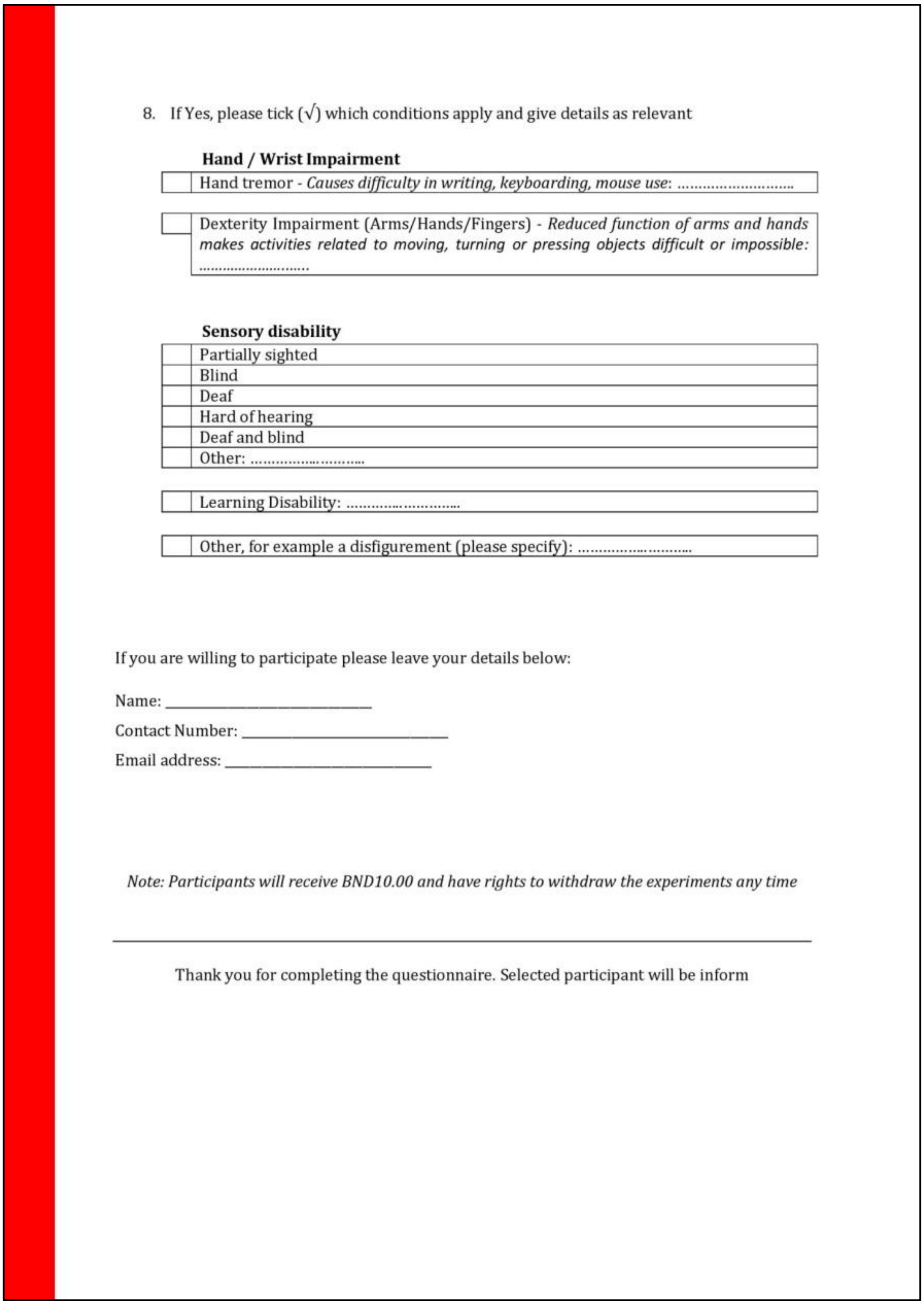
☐ Yes ☐ No (Proceed to Q.7)

6. If Yes, how long you have been using it:

☐ 0 – 1 ☐ 2 – 3 ☐ 4 and above

7. Do you consider yourself to have disability?

☐ Yes ☐ No



8. If Yes, please tick (√) which conditions apply and give details as relevant

Hand / Wrist Impairment

☐ Hand tremor - *Causes difficulty in writing, keyboarding, mouse use*: ……………………….

☐ Dexterity Impairment (Arms/Hands/Fingers) - *Reduced function of arms and hands makes activities related to moving, turning or pressing objects difficult or impossible:* ……………………….

Sensory disability

☐ Partially sighted
☐ Blind
☐ Deaf
☐ Hard of hearing
☐ Deaf and blind
☐ Other: ……………………….

☐ Learning Disability: ……………………….

☐ Other, for example a disfigurement (please specify): ……………………….

If you are willing to participate please leave your details below:

Name: ______________________

Contact Number: ______________________

Email address: ______________________

Note: Participants will receive BND10.00 and have rights to withdraw the experiments any time

---

Thank you for completing the questionnaire. Selected participant will be inform

A3: Experiment Overview (Pilot Experiment)

University of
Salford
MANCHESTER

Experiment Overview

The objective of this experiment is to investigate the effectiveness and user satisfaction of using touch screen interfaces. This includes determining the direct input for selection and dragging operations on a touch-based PC.

The experiment task is designed using a simple tapping-dragging task to determine the speed and accuracy of the movement object. There will be 2 boxes with 3 different sizes (small, medium, large) and the task will be tap plus drag for the primary box and place on to target box, then release. Each size of the boxes will be tested on 6 different target box locations, which means 3 different sized boxes x 6 different target box locations equalling 18 tests. The participant will use Finger tips, Stylus and Mouse in order to perform the test.

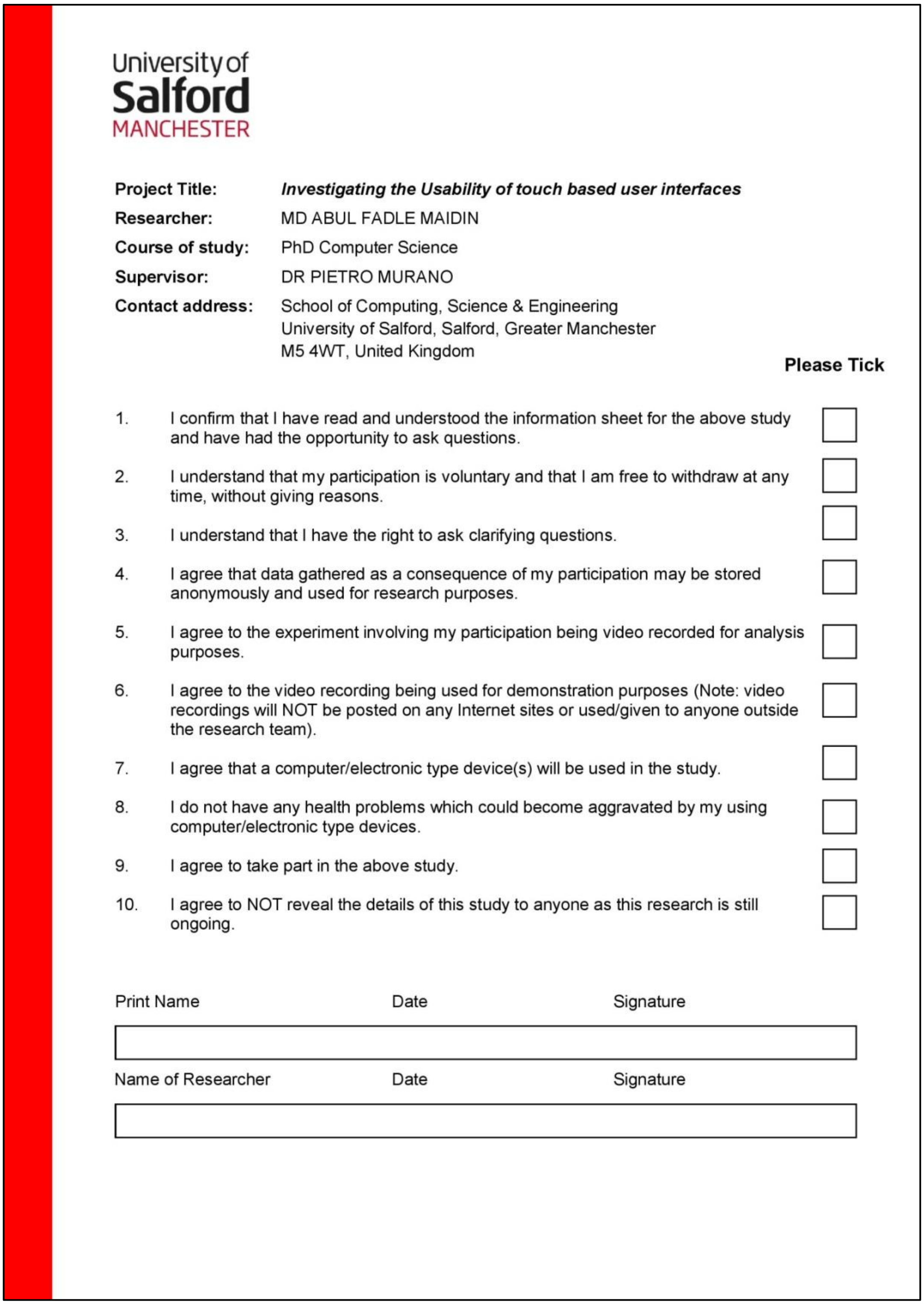
The significant contribution to knowledge that will arise from this research will be the gaining of some evidence to show if touch-based interaction is more effective and preferred by users in real world type tasks/scenarios.

May I thank you, in advance, for your valuable co-operation.

Hj Mohd Abul Fadle Hj Maidin
PhD Student,
School of Computing, Science & Engineering, (College of Science & Technology)
The University of Salford
The Crescent, Salford, M5 4WT, United Kingdom
Email: m.a.f.maidin@edu.salford.ac.uk

Telephone Number: +44 7511845478 / +6738895855

A4: Experiment Consent Form (Pilot Experiment)



University of
Salford
MANCHESTER

Project Title: ***Investigating the Usability of touch based user interfaces***
Researcher: MD ABUL FADLE MAIDIN
Course of study: PhD Computer Science
Supervisor: DR PIETRO MURANO
Contact address: School of Computing, Science & Engineering
University of Salford, Salford, Greater Manchester
M5 4WT, United Kingdom

Please Tick

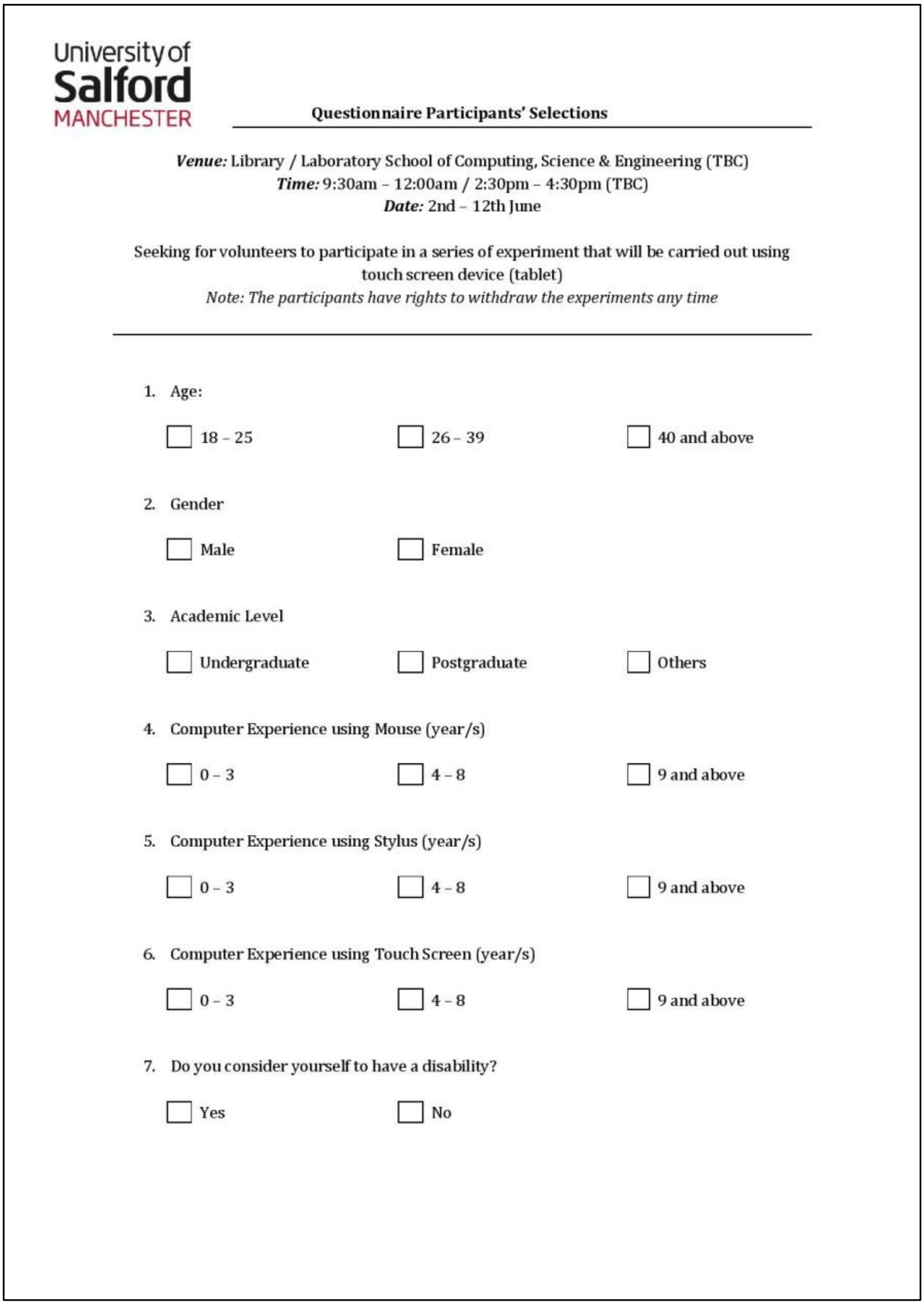
1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reasons. ☐
3. I understand that I have the right to ask clarifying questions. ☐
4. I agree that data gathered as a consequence of my participation may be stored anonymously and used for research purposes. ☐
5. I agree to the experiment involving my participation being video recorded for analysis purposes. ☐
6. I agree to the video recording being used for demonstration purposes (Note: video recordings will NOT be posted on any Internet sites or used/given to anyone outside the research team). ☐
7. I agree that a computer/electronic type device(s) will be used in the study. ☐
8. I do not have any health problems which could become aggravated by my using computer/electronic type devices. ☐
9. I agree to take part in the above study. ☐
10. I agree to NOT reveal the details of this study to anyone as this research is still ongoing. ☐

Print Name	Date	Signature

Name of Researcher	Date	Signature

A5: Experiment Participants’ Selections Questionnaire

Tracing test, Dragging test, One direction tapping test and Multi Directional tapping test (Abstract and Contextual)



University of Salford MANCHESTER

Questionnaire Participants’ Selections

Venue: Library / Laboratory School of Computing, Science & Engineering (TBC)
Time: 9:30am – 12:00am / 2:30pm – 4:30pm (TBC)
Date: 2nd – 12th June

Seeking for volunteers to participate in a series of experiment that will be carried out using touch screen device (tablet)

Note: The participants have rights to withdraw the experiments any time

1. Age:

☐ 18 – 25 ☐ 26 – 39 ☐ 40 and above

2. Gender

☐ Male ☐ Female

3. Academic Level

☐ Undergraduate ☐ Postgraduate ☐ Others

4. Computer Experience using Mouse (year/s)

☐ 0 – 3 ☐ 4 – 8 ☐ 9 and above

5. Computer Experience using Stylus (year/s)

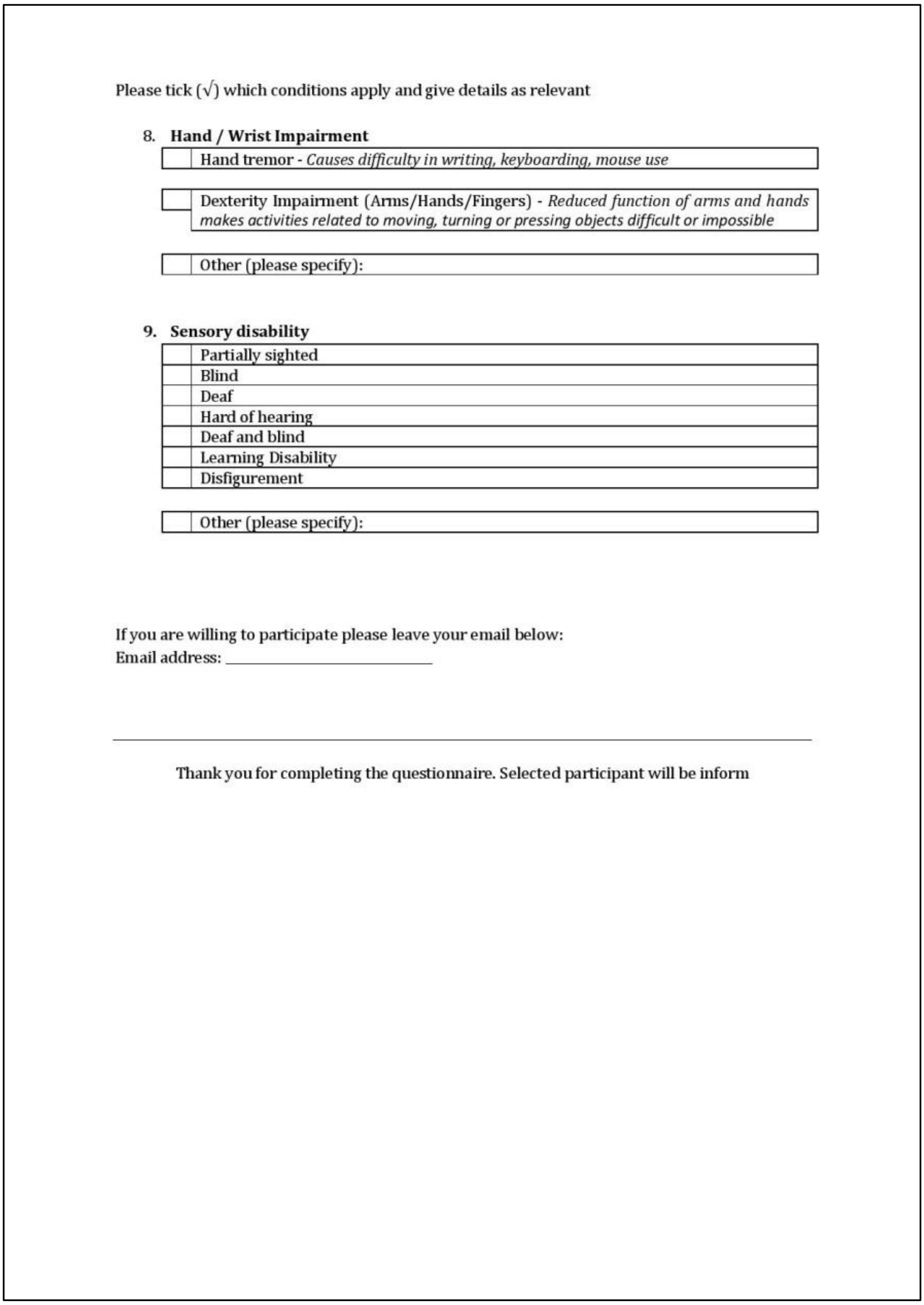
☐ 0 – 3 ☐ 4 – 8 ☐ 9 and above

6. Computer Experience using Touch Screen (year/s)

☐ 0 – 3 ☐ 4 – 8 ☐ 9 and above

7. Do you consider yourself to have a disability?

☐ Yes ☐ No



Please tick (√) which conditions apply and give details as relevant

8. **Hand / Wrist Impairment**

☐ Hand tremor - *Causes difficulty in writing, keyboarding, mouse use*

☐ Dexterity Impairment (Arms/Hands/Fingers) - *Reduced function of arms and hands makes activities related to moving, turning or pressing objects difficult or impossible*

☐ Other (please specify):

9. **Sensory disability**

☐ Partially sighted
☐ Blind
☐ Deaf
☐ Hard of hearing
☐ Deaf and blind
☐ Learning Disability
☐ Disfigurement

☐ Other (please specify):

If you are willing to participate please leave your email below:
Email address: ___________________________

Thank you for completing the questionnaire. Selected participant will be inform

A6: Experiment Overview

Tracing test, Dragging test, One direction tapping test and Multi Directional tapping test (Abstract and Contextual)

EXPERIMENT 2

Experiment Overview

The objective of this experiment is to investigate the effectiveness and user satisfaction of using touch screen interfaces. This includes determining the direct input for selection and dragging operations on a touch-based PC.

The experiment task is designed using a simple tracing and tapping-dragging task to determine the speed, accuracy and error. There will be 2 prototypes in which each prototype consist 4 tasks. The participants have to complete 24 tasks altogether (2 prototype x 4 task x 3 input) which take less than 15 minutes to complete. The participant will use Finger tips, Stylus and Mouse input in order to perform the test.

The significant contribution to knowledge that will arise from this research will be the gaining of some evidence to show if touch-based interaction is more effective and preferred by users in real world type tasks/scenarios.

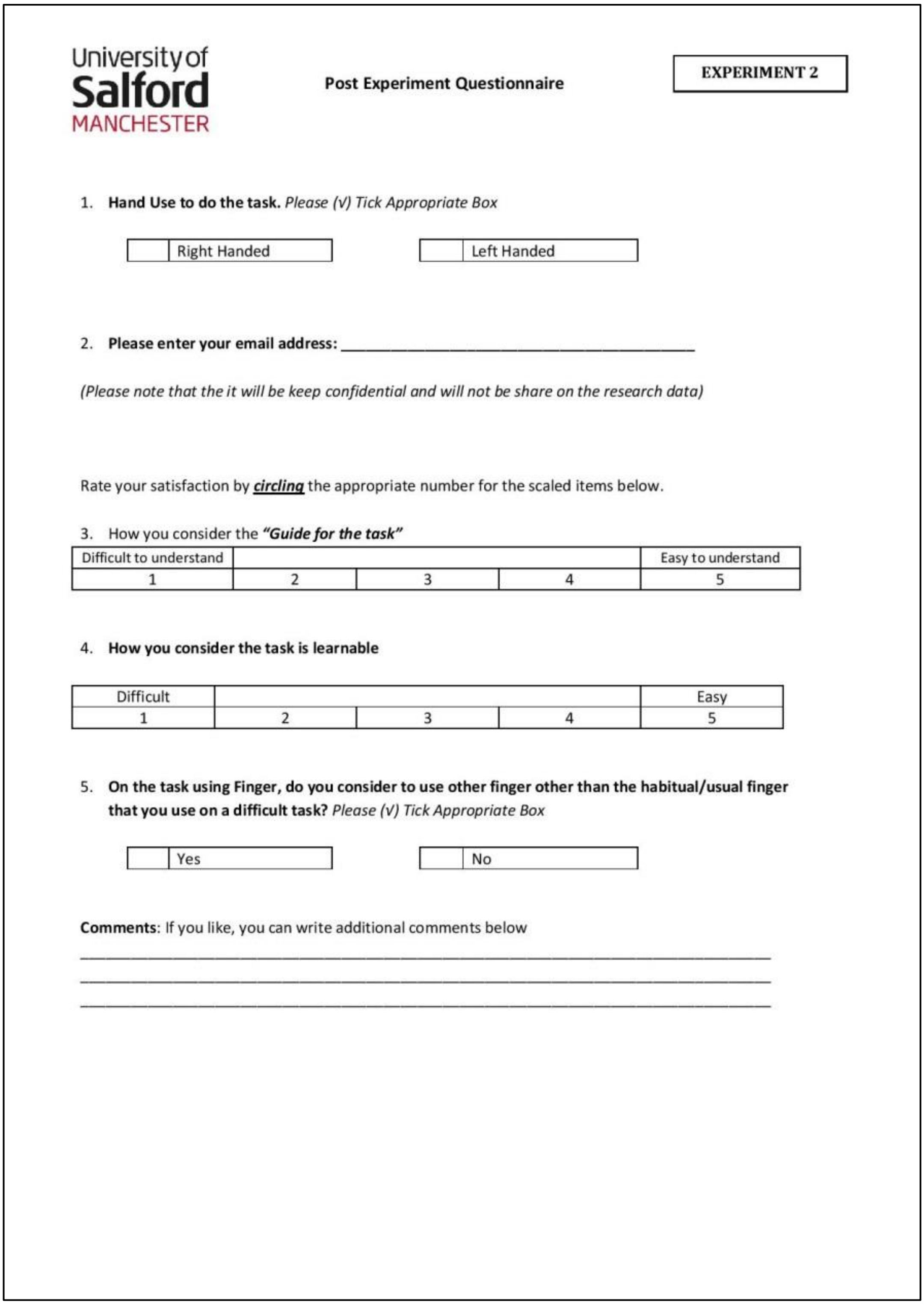
May I thank you, in advance, for your valuable co-operation.

Md Abul Fadle Maidin
PhD Student,
School of Computing, Science & Engineering, (College of Science & Technology)
The University of Salford
The Crescent, Salford, M5 4WT, United Kingdom
Email: m.a.f.maidin@edu.salford.ac.uk

Telephone Number: 07511845478

A7: Post Experiment Questionnaire

Tracing test, Dragging test, One direction tapping test and Multi Directional tapping test (Abstract and Contextual)



University of
Salford
MANCHESTER

Post Experiment Questionnaire

EXPERIMENT 2

1. **Hand Use to do the task.** *Please (√) Tick Appropriate Box*

☐ Right Handed ☐ Left Handed

2. **Please enter your email address:** ____

(Please note that the it will be keep confidential and will not be share on the research data)

Rate your satisfaction by ***circling*** the appropriate number for the scaled items below.

3. How you consider the ***"Guide for the task"***

Difficult to understand				Easy to understand
1	2	3	4	5

4. **How you consider the task is learnable**

Difficult				Easy
1	2	3	4	5

5. **On the task using Finger, do you consider to use other finger other than the habitual/usual finger that you use on a difficult task?** *Please (√) Tick Appropriate Box*

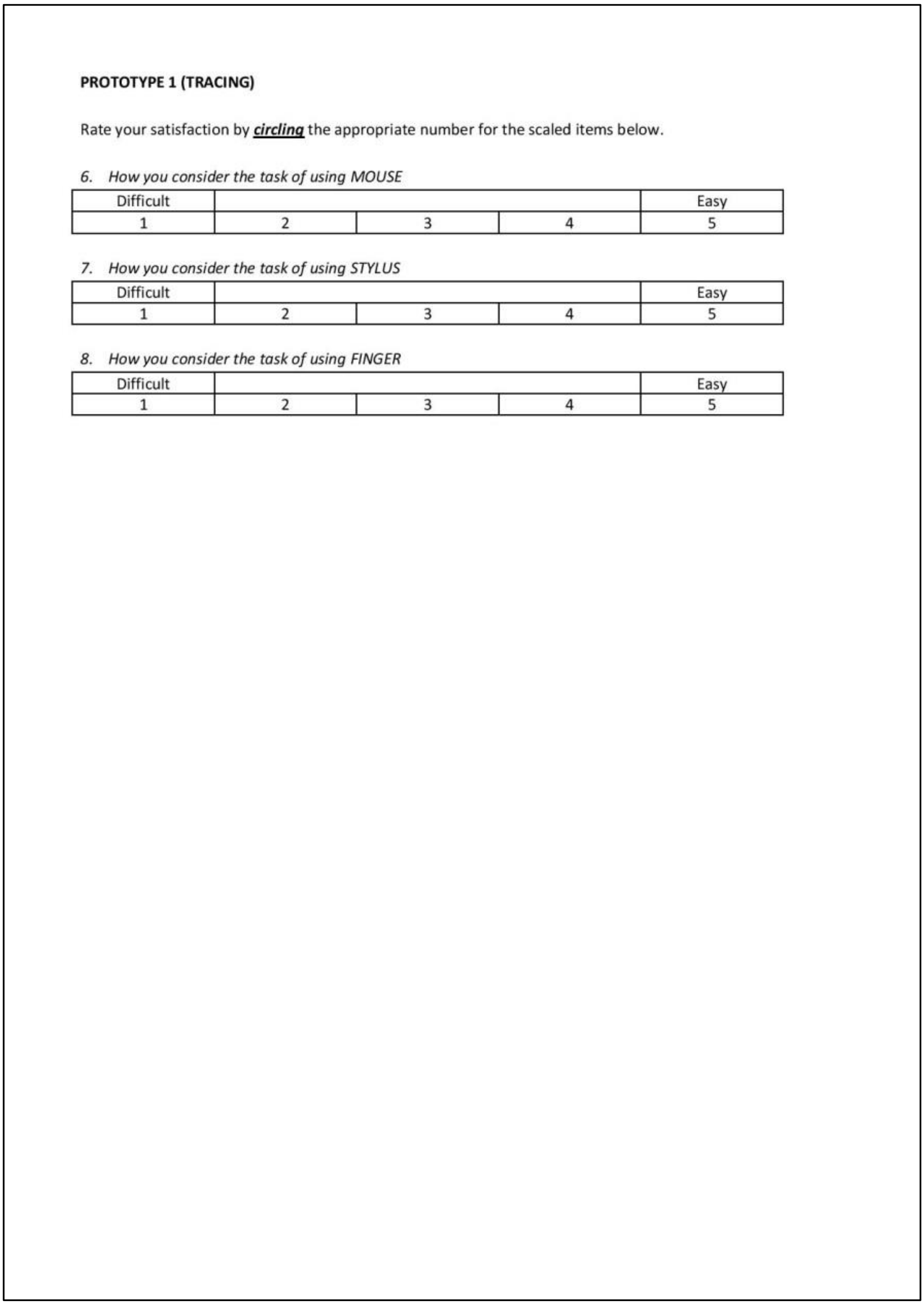
☐ Yes ☐ No

Comments: If you like, you can write additional comments below

____

____

____



PROTOTYPE 1 (TRACING)

Rate your satisfaction by ***circling*** the appropriate number for the scaled items below.

6. *How you consider the task of using MOUSE*

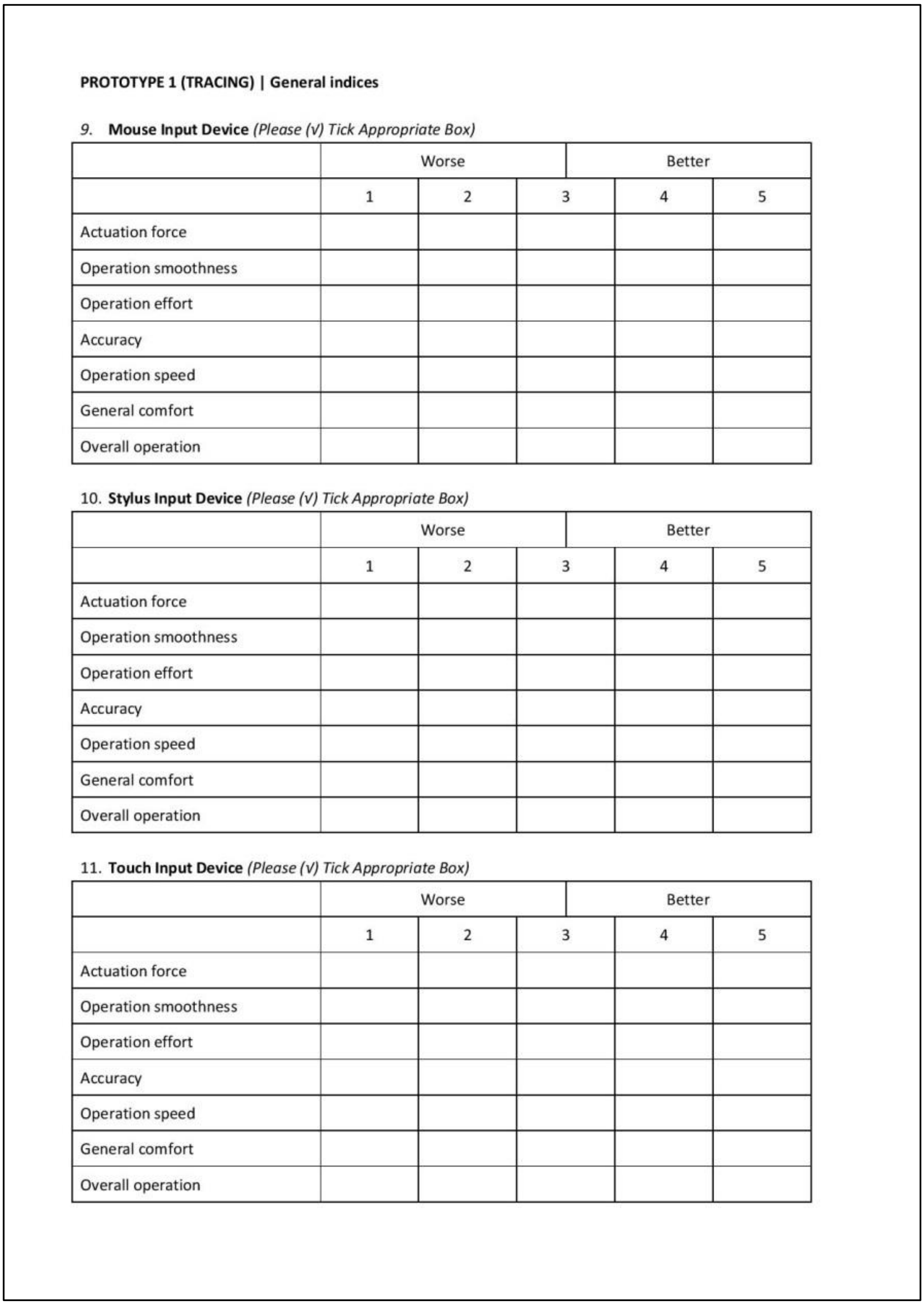
Difficult				Easy
1	2	3	4	5

7. *How you consider the task of using STYLUS*

Difficult				Easy
1	2	3	4	5

8. *How you consider the task of using FINGER*

Difficult				Easy
1	2	3	4	5



PROTOTYPE 1 (TRACING) | General indices

9. **Mouse Input Device** *(Please (√) Tick Appropriate Box)*

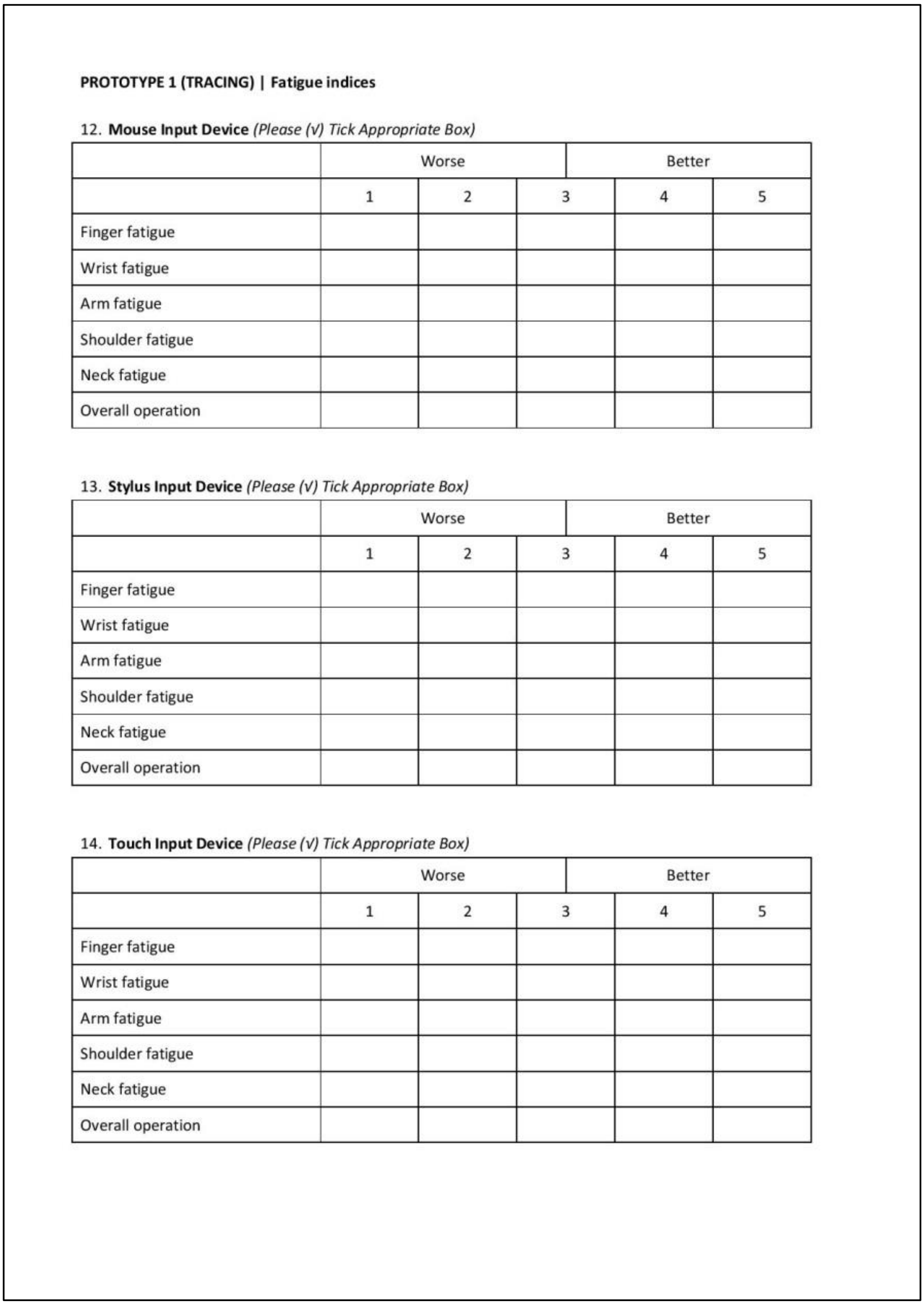
	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

10. **Stylus Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

11. **Touch Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					



PROTOTYPE 1 (TRACING) | Fatigue indices

12. **Mouse Input Device** *(Please (√) Tick Appropriate Box)*

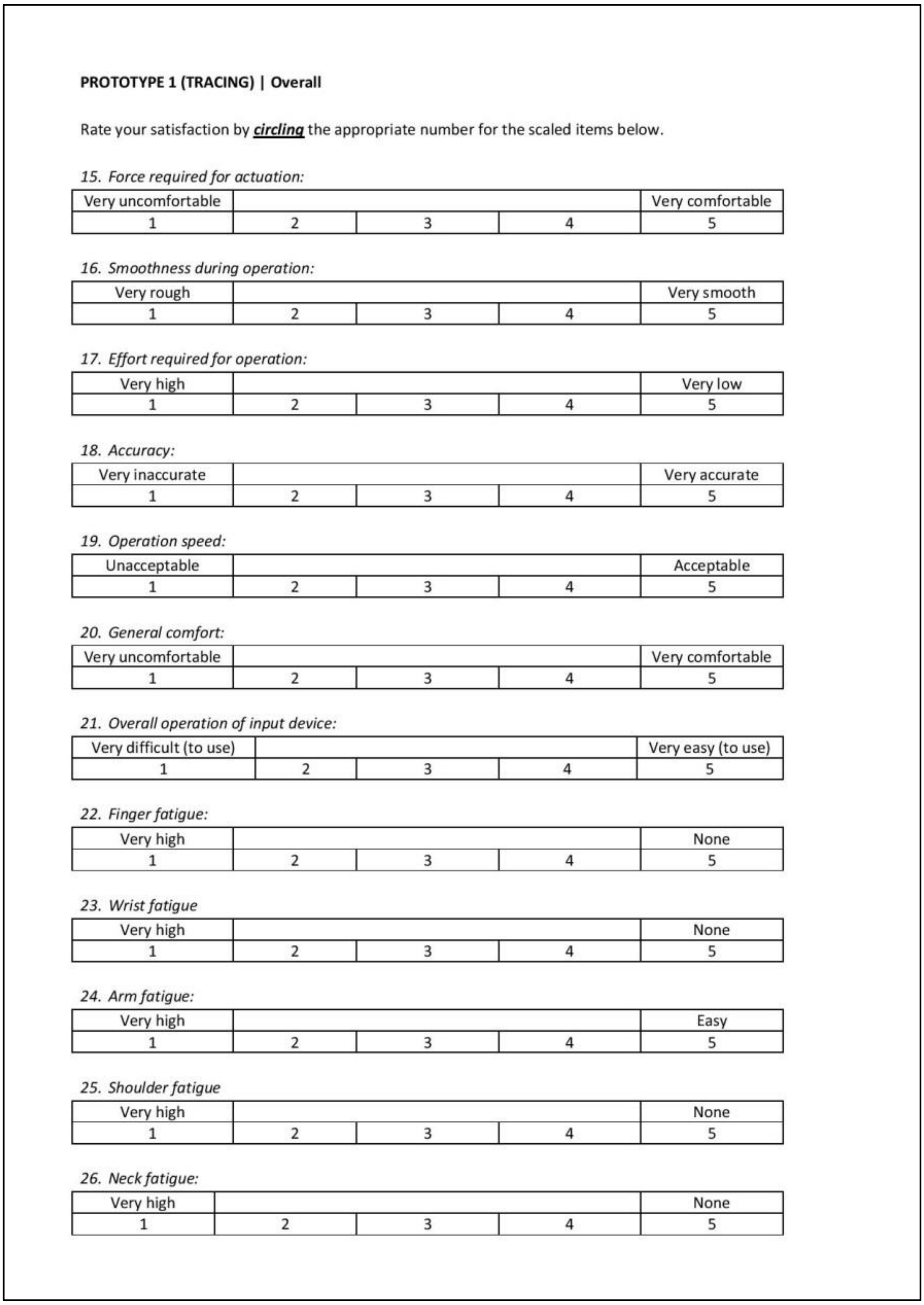
	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

13. **Stylus Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

14. **Touch Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					



PROTOTYPE 1 (TRACING) | Overall

Rate your satisfaction by ***circling*** the appropriate number for the scaled items below.

15. Force required for actuation:

Very uncomfortable				Very comfortable
1	2	3	4	5

16. Smoothness during operation:

Very rough				Very smooth
1	2	3	4	5

17. Effort required for operation:

Very high				Very low
1	2	3	4	5

18. Accuracy:

Very inaccurate				Very accurate
1	2	3	4	5

19. Operation speed:

Unacceptable				Acceptable
1	2	3	4	5

20. General comfort:

Very uncomfortable				Very comfortable
1	2	3	4	5

21. Overall operation of input device:

Very difficult (to use)				Very easy (to use)
1	2	3	4	5

22. Finger fatigue:

Very high				None
1	2	3	4	5

23. Wrist fatigue

Very high				None
1	2	3	4	5

24. Arm fatigue:

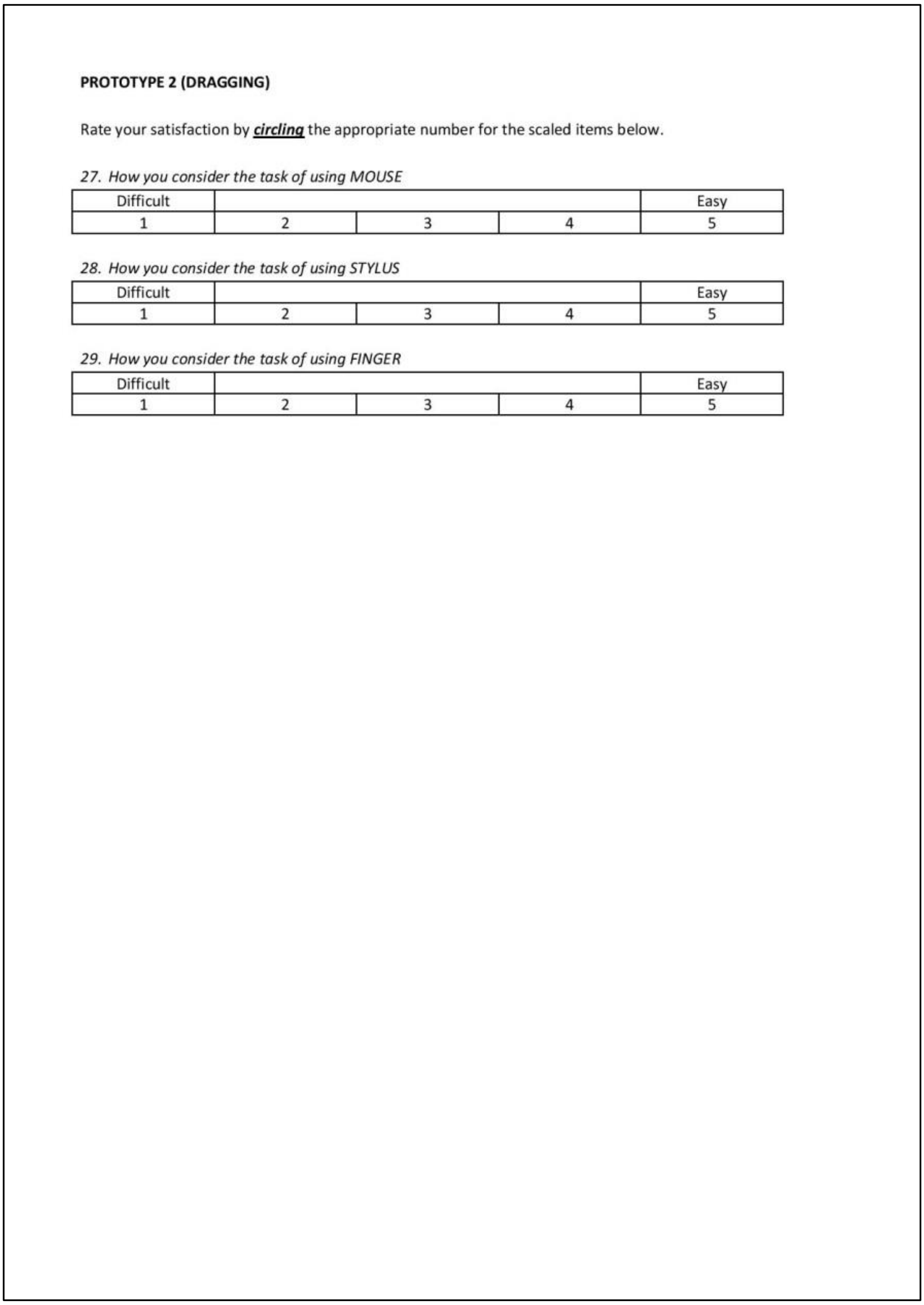
Very high				Easy
1	2	3	4	5

25. Shoulder fatigue

Very high				None
1	2	3	4	5

26. Neck fatigue:

Very high				None
1	2	3	4	5



PROTOTYPE 2 (DRAGGING)

Rate your satisfaction by ***circling*** the appropriate number for the scaled items below.

27. How you consider the task of using MOUSE

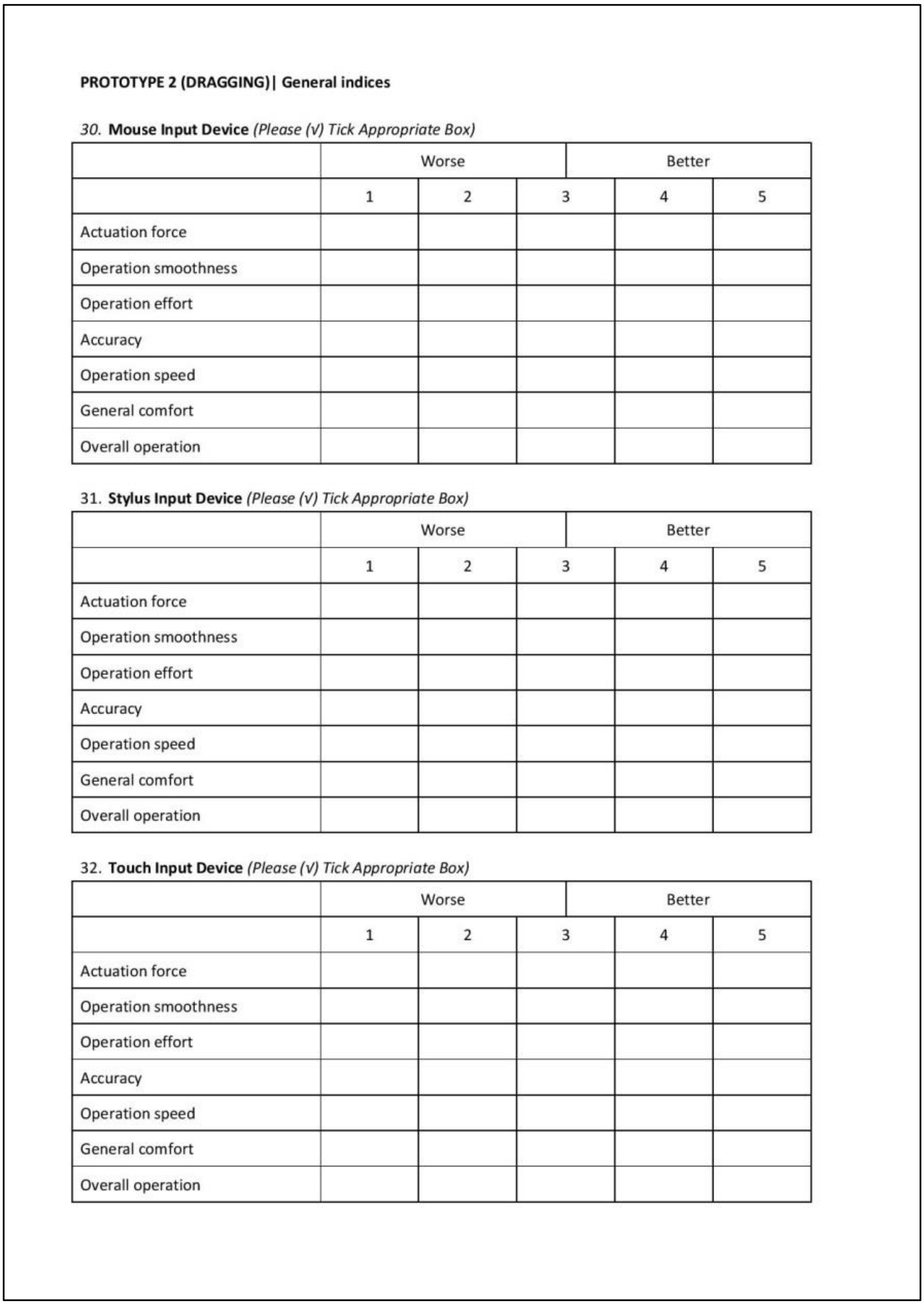
Difficult				Easy
1	2	3	4	5

28. How you consider the task of using STYLUS

Difficult				Easy
1	2	3	4	5

29. How you consider the task of using FINGER

Difficult				Easy
1	2	3	4	5



PROTOTYPE 2 (DRAGGING)| General indices

30. **Mouse Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

31. **Stylus Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

32. **Touch Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Actuation force					
Operation smoothness					
Operation effort					
Accuracy					
Operation speed					
General comfort					
Overall operation					

PROTOTYPE 2 (DRAGGING)| Fatigue indices

33. **Mouse Input Device** *(Please (√) Tick Appropriate Box)*

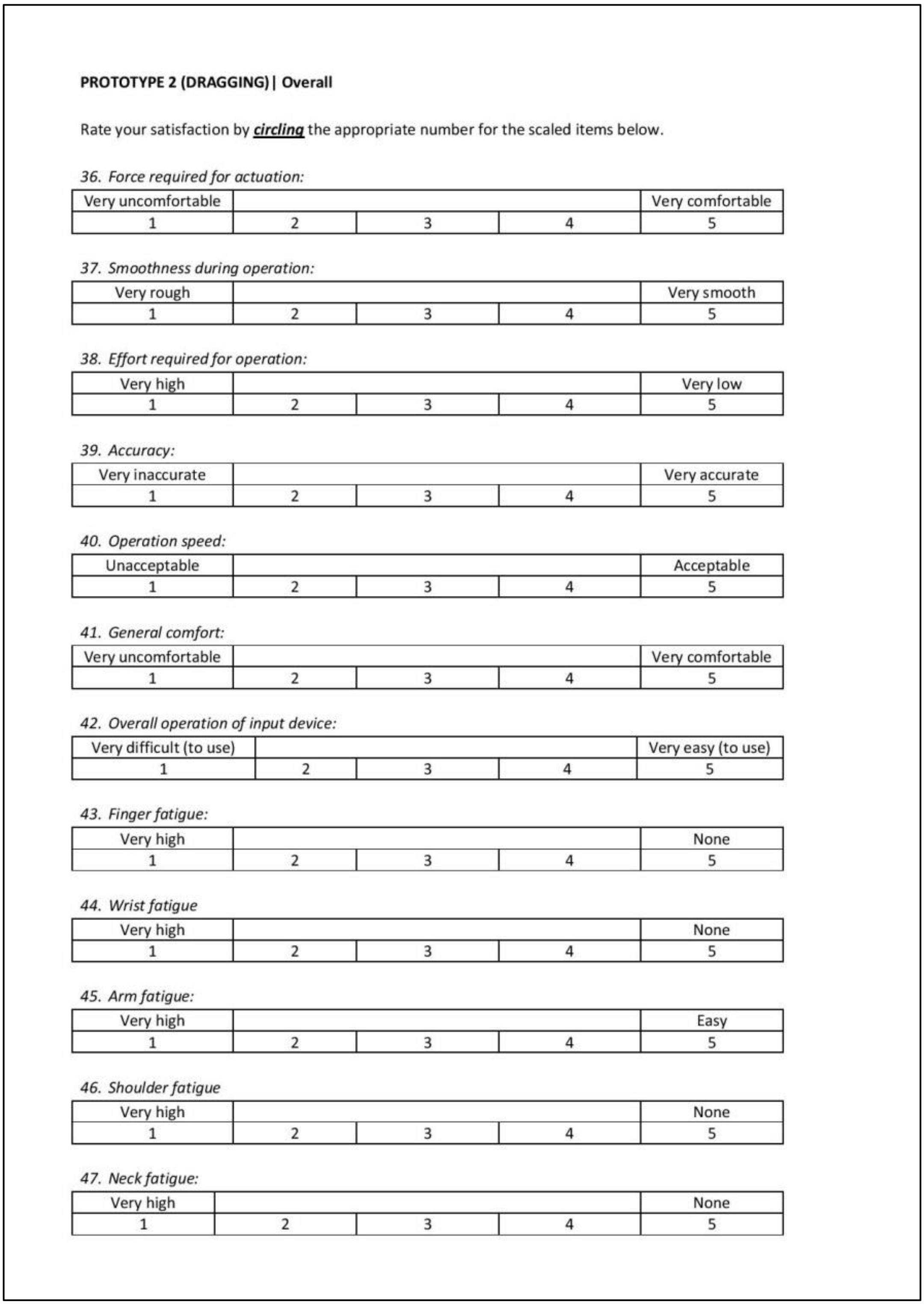
	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

34. **Stylus Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					

35. **Touch Input Device** *(Please (√) Tick Appropriate Box)*

	Worse			Better	
	1	2	3	4	5
Finger fatigue					
Wrist fatigue					
Arm fatigue					
Shoulder fatigue					
Neck fatigue					
Overall operation					



PROTOTYPE 2 (DRAGGING) | Overall

Rate your satisfaction by ***circling*** the appropriate number for the scaled items below.

36. Force required for actuation:

Very uncomfortable				Very comfortable
1	2	3	4	5

37. Smoothness during operation:

Very rough				Very smooth
1	2	3	4	5

38. Effort required for operation:

Very high				Very low
1	2	3	4	5

39. Accuracy:

Very inaccurate				Very accurate
1	2	3	4	5

40. Operation speed:

Unacceptable				Acceptable
1	2	3	4	5

41. General comfort:

Very uncomfortable				Very comfortable
1	2	3	4	5

42. Overall operation of input device:

Very difficult (to use)				Very easy (to use)
1	2	3	4	5

43. Finger fatigue:

Very high				None
1	2	3	4	5

44. Wrist fatigue

Very high				None
1	2	3	4	5

45. Arm fatigue:

Very high				Easy
1	2	3	4	5

46. Shoulder fatigue

Very high				None
1	2	3	4	5

47. Neck fatigue:

Very high				None
1	2	3	4	5

APPENDIX B: DESCRIPTIVES DATA

B1: Tracing Test (Abstract) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		22.20315	1.241069
	95% Confidence Interval for Mean	Lower Bound	19.73792	
		Upper Bound	24.66838	
	5% Trimmed Mean		21.61619	
	Median		18.10100	
	Variance		141.703	
	Std. Deviation		11.903914	
	Minimum		5.631	
	Maximum		53.253	
	Range		47.622	
	Interquartile Range		18.535	
	Skewness		.635	.251
	Kurtosis		-.559	.498
Stylus Task Time	Mean		15.77967	.908743
	95% Confidence Interval for Mean	Lower Bound	13.97457	
		Upper Bound	17.58478	
	5% Trimmed Mean		15.15196	
	Median		13.03350	
	Variance		75.975	
	Std. Deviation		8.716352	
	Minimum		4.187	
	Maximum		40.486	
	Range		36.299	
	Interquartile Range		11.436	
	Skewness		1.116	.251
	Kurtosis		.407	.498
Touch Task Time	Mean		10.08291	.839775
	95% Confidence Interval for Mean	Lower Bound	8.41480	
		Upper Bound	11.75102	
	5% Trimmed Mean		9.05813	
	Median		6.88150	
	Variance		64.880	
	Std. Deviation		8.054838	
	Minimum		2.469	
	Maximum		39.640	
	Range		37.171	
	Interquartile Range		5.411	
	Skewness		2.005	.251
	Kurtosis		3.867	.498

B2: Tracing Test (Abstract) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		201.33696	27.680266
	95% Confidence Interval for Mean	Lower Bound	146.35351	
		Upper Bound	256.32041	
	5% Trimmed Mean		161.59662	
	Median		87.00000	
	Variance		70490.138	
	Std. Deviation		265.499789	
	Minimum		53.000	
	Maximum		1179.000	
	Range		1126.000	
	Interquartile Range		66.000	
	Skewness		2.326	.251
	Kurtosis		4.755	.498
Stylus Task Error	Mean		192.91304	25.465033
	95% Confidence Interval for Mean	Lower Bound	142.32988	
		Upper Bound	243.49620	
	5% Trimmed Mean		159.78261	
	Median		91.50000	
	Variance		59659.047	
	Std. Deviation		244.252016	
	Minimum		49.000	
	Maximum		1146.000	
	Range		1097.000	
	Interquartile Range		56.750	
	Skewness		2.240	.251
	Kurtosis		4.168	.498
Touch Task Error	Mean		317.45652	31.728315
	95% Confidence Interval for Mean	Lower Bound	254.43212	
		Upper Bound	380.48092	
	5% Trimmed Mean		281.07005	
	Median		199.50000	
	Variance		92615.108	
	Std. Deviation		304.327304	
	Minimum		82.000	
	Maximum		1294.000	
	Range		1212.000	
	Interquartile Range		136.250	
	Skewness		1.932	.251
	Kurtosis		2.465	.498

B3: Dragging Test (Abstract) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		33.64700	1.158924
	95% Confidence Interval for Mean	Lower Bound	31.34494	
		Upper Bound	35.94906	
	5% Trimmed Mean		32.55750	
	Median		31.17200	
	Variance		123.566	
	Std. Deviation		11.116008	
	Minimum		18.233	
	Maximum		94.102	
	Range		75.869	
	Interquartile Range		9.866	
	Skewness		2.318	.251
	Kurtosis		9.422	.498
Stylus Task Time	Mean		32.24916	1.506429
	95% Confidence Interval for Mean	Lower Bound	29.25683	
		Upper Bound	35.24150	
	5% Trimmed Mean		30.33770	
	Median		29.07750	
	Variance		208.778	
	Std. Deviation		14.449158	
	Minimum		15.931	
	Maximum		123.656	
	Range		107.725	
	Interquartile Range		10.429	
	Skewness		3.664	.251
	Kurtosis		18.928	.498
Touch Task Time	Mean		36.55923	1.872779
	95% Confidence Interval for Mean	Lower Bound	32.83918	
		Upper Bound	40.27927	
	5% Trimmed Mean		34.94902	
	Median		30.92350	
	Variance		322.672	
	Std. Deviation		17.963067	
	Minimum		15.586	
	Maximum		96.551	
	Range		80.965	
	Interquartile Range		24.326	
	Skewness		1.281	.251
	Kurtosis		1.217	.498

B4: Dragging Test (Abstract) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		17.89130	.793014
	95% Confidence Interval for Mean	Lower Bound	16.31608	
		Upper Bound	19.46653	
	5% Trimmed Mean		17.41787	
	Median		16.50000	
	Variance		57.856	
	Std. Deviation		7.606325	
	Minimum		7.000	
	Maximum		41.000	
	Range		34.000	
	Interquartile Range		11.000	
	Skewness		.909	.251
	Kurtosis		.195	.498
Stylus Task Error	Mean		25.19565	1.136864
	95% Confidence Interval for Mean	Lower Bound	22.93741	
		Upper Bound	27.45389	
	5% Trimmed Mean		24.55072	
	Median		23.00000	
	Variance		118.906	
	Std. Deviation		10.904419	
	Minimum		8.000	
	Maximum		63.000	
	Range		55.000	
	Interquartile Range		14.000	
	Skewness		.884	.251
	Kurtosis		.833	.498
Touch Task Error	Mean		53.39130	1.619279
	95% Confidence Interval for Mean	Lower Bound	50.17481	
		Upper Bound	56.60780	
	5% Trimmed Mean		53.35024	
	Median		54.00000	
	Variance		241.230	
	Std. Deviation		15.531575	
	Minimum		18.000	
	Maximum		99.000	
	Range		81.000	
	Interquartile Range		22.500	
	Skewness		.045	.251
	Kurtosis		-.032	.498

B5: One Direction Tapping Test (Abstract) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		55.09058	1.141113
	95% Confidence Interval for Mean	Lower Bound	52.82249	
		Upper Bound	57.35866	
	5% Trimmed Mean		55.00259	
	Median		54.51750	
	Variance		114.588	
	Std. Deviation		10.704587	
	Minimum		20.898	
	Maximum		84.382	
	Range		63.484	
	Interquartile Range		13.108	
	Skewness		.137	.257
	Kurtosis		1.284	.508
Stylus Task Time	Mean		56.98274	1.876863
	95% Confidence Interval for Mean	Lower Bound	53.25227	
		Upper Bound	60.71321	
	5% Trimmed Mean		55.22440	
	Median		53.83050	
	Variance		309.990	
	Std. Deviation		17.606540	
	Minimum		29.111	
	Maximum		163.348	
	Range		134.237	
	Interquartile Range		15.022	
	Skewness		2.940	.257
	Kurtosis		14.864	.508
Touch Task Time	Mean		46.28066	1.085458
	95% Confidence Interval for Mean	Lower Bound	44.12319	
		Upper Bound	48.43813	
	5% Trimmed Mean		45.75662	
	Median		44.78600	
	Variance		103.683	
	Std. Deviation		10.182502	
	Minimum		24.146	
	Maximum		82.065	
	Range		57.919	
	Interquartile Range		13.260	
	Skewness		.914	.257
	Kurtosis		1.497	.508

B6: One Direction Tapping Test (Abstract) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		4.09091	1.091268
	95% Confidence Interval for Mean	Lower Bound	1.92190	
		Upper Bound	6.25992	
	5% Trimmed Mean		2.14141	
	Median		1.00000	
	Variance		104.796	
	Std. Deviation		10.237003	
	Minimum		.000	
	Maximum		55.000	
	Range		55.000	
	Interquartile Range		2.000	
	Skewness		3.677	.257
	Kurtosis		13.674	.508
Stylus Task Error	Mean		3.76136	.940496
	95% Confidence Interval for Mean	Lower Bound	1.89203	
		Upper Bound	5.63070	
	5% Trimmed Mean		2.04040	
	Median		1.00000	
	Variance		77.839	
	Std. Deviation		8.822638	
	Minimum		.000	
	Maximum		51.000	
	Range		51.000	
	Interquartile Range		3.000	
	Skewness		3.868	.257
	Kurtosis		15.635	.508
Touch Task Error	Mean		6.14773	1.436894
	95% Confidence Interval for Mean	Lower Bound	3.29174	
		Upper Bound	9.00371	
	5% Trimmed Mean		3.59596	
	Median		2.00000	
	Variance		181.691	
	Std. Deviation		13.479264	
	Minimum		.000	
	Maximum		88.000	
	Range		88.000	
	Interquartile Range		6.000	
	Skewness		4.347	.257
	Kurtosis		20.944	.508

B7: Multi Directional Tapping Test (Abstract) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		25.83445	.396322
	95% Confidence Interval for Mean	Lower Bound	25.04294	
		Upper Bound	26.62596	
	5% Trimmed Mean		25.81386	
	Median		26.03400	
	Variance		10.367	
	Std. Deviation		3.219737	
	Minimum		19.267	
	Maximum		33.277	
	Range		14.010	
	Interquartile Range		3.999	
	Skewness		.035	.295
	Kurtosis		-.383	.582
Stylus Task Time	Mean		22.57780	.441684
	95% Confidence Interval for Mean	Lower Bound	21.69570	
		Upper Bound	23.45991	
	5% Trimmed Mean		22.45892	
	Median		22.46250	
	Variance		12.876	
	Std. Deviation		3.588255	
	Minimum		15.596	
	Maximum		35.114	
	Range		19.518	
	Interquartile Range		3.905	
	Skewness		.664	.295
	Kurtosis		1.470	.582
Touch Task Time	Mean		21.53488	.475062
	95% Confidence Interval for Mean	Lower Bound	20.58611	
		Upper Bound	22.48364	
	5% Trimmed Mean		21.49625	
	Median		20.92300	
	Variance		14.895	
	Std. Deviation		3.859425	
	Minimum		12.142	
	Maximum		31.236	
	Range		19.094	
	Interquartile Range		4.578	
	Skewness		.251	.295
	Kurtosis		.120	.582

B8: Multi Directional Tapping Test (Abstract) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		.40909	.116229
	95% Confidence Interval for Mean	Lower Bound	.17697	
		Upper Bound	.64122	
	5% Trimmed Mean		.27104	
	Median		.00000	
	Variance		.892	
	Std. Deviation		.944250	
	Minimum		.000	
	Maximum		4.000	
	Range		4.000	
	Interquartile Range		.000	
	Skewness		2.362	.295
	Kurtosis		4.663	.582
Stylus Task Error	Mean		.69697	.131283
	95% Confidence Interval for Mean	Lower Bound	.43478	
		Upper Bound	.95916	
	5% Trimmed Mean		.55724	
	Median		.00000	
	Variance		1.138	
	Std. Deviation		1.066550	
	Minimum		.000	
	Maximum		5.000	
	Range		5.000	
	Interquartile Range		1.000	
	Skewness		1.897	.295
	Kurtosis		4.035	.582
Touch Task Error	Mean		2.90909	.379993
	95% Confidence Interval for Mean	Lower Bound	2.15019	
		Upper Bound	3.66799	
	5% Trimmed Mean		2.51347	
	Median		2.00000	
	Variance		9.530	
	Std. Deviation		3.087081	
	Minimum		.000	
	Maximum		16.000	
	Range		16.000	
	Interquartile Range		3.000	
	Skewness		2.337	.295
	Kurtosis		7.119	.582

B9: Tracing Test (Context) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		40.95326	1.999485
	95% Confidence Interval for Mean	Lower Bound	36.98378	
		Upper Bound	44.92274	
	5% Trimmed Mean		39.02362	
	Median		37.38200	
	Variance		383.802	
	Std. Deviation		19.590867	
	Minimum		15.191	
	Maximum		112.359	
	Range		97.168	
	Interquartile Range		23.875	
	Skewness		1.456	.246
	Kurtosis		2.358	.488
Stylus Task Time	Mean		25.83124	1.080514
	95% Confidence Interval for Mean	Lower Bound	23.68615	
		Upper Bound	27.97633	
	5% Trimmed Mean		24.80535	
	Median		22.35950	
	Variance		112.081	
	Std. Deviation		10.586833	
	Minimum		12.939	
	Maximum		61.187	
	Range		48.248	
	Interquartile Range		12.635	
	Skewness		1.436	.246
	Kurtosis		1.905	.488
Touch Task Time	Mean		22.63377	1.540305
	95% Confidence Interval for Mean	Lower Bound	19.57588	
		Upper Bound	25.69166	
	5% Trimmed Mean		20.22160	
	Median		18.78750	
	Variance		227.764	
	Std. Deviation		15.091847	
	Minimum		9.424	
	Maximum		96.520	
	Range		87.096	
	Interquartile Range		8.939	
	Skewness		3.175	.246
	Kurtosis		11.305	.488

B10: Tracing Test (Context) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		158.19792	7.086969
	95% Confidence Interval for Mean	Lower Bound	144.12850	
		Upper Bound	172.26733	
	5% Trimmed Mean		152.65046	
	Median		147.50000	
	Variance		4821.613	
	Std. Deviation		69.437836	
	Minimum		42.000	
	Maximum		448.000	
	Range		406.000	
	Interquartile Range		78.500	
	Skewness		1.638	.246
	Kurtosis		4.657	.488
Stylus Task Error	Mean		108.04167	3.641468
	95% Confidence Interval for Mean	Lower Bound	100.81244	
		Upper Bound	115.27089	
	5% Trimmed Mean		106.25463	
	Median		99.00000	
	Variance		1272.988	
	Std. Deviation		35.678953	
	Minimum		50.000	
	Maximum		204.000	
	Range		154.000	
	Interquartile Range		44.750	
	Skewness		.788	.246
	Kurtosis		.059	.488
Touch Task Error	Mean		230.31250	7.417726
	95% Confidence Interval for Mean	Lower Bound	215.58645	
		Upper Bound	245.03855	
	5% Trimmed Mean		229.43056	
	Median		226.50000	
	Variance		5282.175	
	Std. Deviation		72.678573	
	Minimum		70.000	
	Maximum		404.000	
	Range		334.000	
	Interquartile Range		91.500	
	Skewness		.260	.246
	Kurtosis		.046	.488

B11: Dragging Test (Context) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		26.63068	1.476991
	95% Confidence Interval for Mean	Lower Bound	23.69848	
		Upper Bound	29.56287	
	5% Trimmed Mean		24.80894	
	Median		23.83400	
	Variance		209.424	
	Std. Deviation		14.471497	
	Minimum		11.287	
	Maximum		113.636	
	Range		102.349	
	Interquartile Range		14.608	
	Skewness		3.210	.246
	Kurtosis		15.278	.488
Stylus Task Time	Mean		21.81398	.878989
	95% Confidence Interval for Mean	Lower Bound	20.06897	
		Upper Bound	23.55899	
	5% Trimmed Mean		21.33843	
	Median		20.83850	
	Variance		74.172	
	Std. Deviation		8.612300	
	Minimum		9.245	
	Maximum		51.531	
	Range		42.286	
	Interquartile Range		13.461	
	Skewness		.664	.246
	Kurtosis		.405	.488
Touch Task Time	Mean		27.19170	1.410557
	95% Confidence Interval for Mean	Lower Bound	24.39140	
		Upper Bound	29.99201	
	5% Trimmed Mean		26.09303	
	Median		23.15100	
	Variance		191.008	
	Std. Deviation		13.820576	
	Minimum		8.558	
	Maximum		79.178	
	Range		70.620	
	Interquartile Range		13.935	
	Skewness		1.339	.246
	Kurtosis		1.729	.488

B12: One Direction Tapping Test (Context) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		52.07335	.958073
	95% Confidence Interval for Mean	Lower Bound	50.17025	
		Upper Bound	53.97644	
	5% Trimmed Mean		51.57083	
	Median		51.45150	
	Variance		84.447	
	Std. Deviation		9.189511	
	Minimum		35.879	
	Maximum		81.230	
	Range		45.351	
	Interquartile Range		12.797	
	Skewness		.715	.251
	Kurtosis		.390	.498
Stylus Task Time	Mean		48.67397	.755985
	95% Confidence Interval for Mean	Lower Bound	47.17230	
		Upper Bound	50.17564	
	5% Trimmed Mean		48.63258	
	Median		47.92500	
	Variance		52.579	
	Std. Deviation		7.251154	
	Minimum		34.763	
	Maximum		64.442	
	Range		29.679	
	Interquartile Range		10.468	
	Skewness		.213	.251
	Kurtosis		-.653	.498
Touch Task Time	Mean		38.41068	.718625
	95% Confidence Interval for Mean	Lower Bound	36.98322	
		Upper Bound	39.83814	
	5% Trimmed Mean		38.29326	
	Median		37.74950	
	Variance		47.511	
	Std. Deviation		6.892808	
	Minimum		5.211	
	Maximum		56.270	
	Range		51.059	
	Interquartile Range		7.611	
	Skewness		-.505	.251
	Kurtosis		5.517	.498

B13: One Direction Tapping Test (Context) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		.85870	.118299
	95% Confidence Interval for Mean	Lower Bound	.62371	
		Upper Bound	1.09368	
	5% Trimmed Mean		.72705	
	Median		1.00000	
	Variance		1.288	
	Std. Deviation		1.134683	
	Minimum		.000	
	Maximum		6.000	
	Range		6.000	
	Interquartile Range		1.000	
	Skewness		1.899	.251
	Kurtosis		5.003	.498
Stylus Task Error	Mean		.97826	.221825
	95% Confidence Interval for Mean	Lower Bound	.53763	
		Upper Bound	1.41889	
	5% Trimmed Mean		.67874	
	Median		.00000	
	Variance		4.527	
	Std. Deviation		2.127674	
	Minimum		.000	
	Maximum		18.000	
	Range		18.000	
	Interquartile Range		1.000	
	Skewness		5.970	.251
	Kurtosis		45.522	.498
Touch Task Error	Mean		1.31522	.245810
	95% Confidence Interval for Mean	Lower Bound	.82695	
		Upper Bound	1.80349	
	5% Trimmed Mean		.97343	
	Median		1.00000	
	Variance		5.559	
	Std. Deviation		2.357729	
	Minimum		.000	
	Maximum		16.000	
	Range		16.000	
	Interquartile Range		2.000	
	Skewness		4.138	.251
	Kurtosis		21.862	.498

B14: Multi Directional Tapping Test (Context) – Movement Time

			Statistic	Std. Error
Mouse Task Time	Mean		24.09000	.758838
	95% Confidence Interval for Mean	Lower Bound	22.57450	
		Upper Bound	25.60550	
	5% Trimmed Mean		23.22953	
	Median		22.51150	
	Variance		38.005	
	Std. Deviation		6.164830	
	Minimum		19.023	
	Maximum		60.650	
	Range		41.627	
	Interquartile Range		3.739	
	Skewness		3.816	.295
	Kurtosis		19.047	.582
Stylus Task Time	Mean		22.92038	.470348
	95% Confidence Interval for Mean	Lower Bound	21.98103	
		Upper Bound	23.85973	
	5% Trimmed Mean		22.42471	
	Median		21.80550	
	Variance		14.601	
	Std. Deviation		3.821124	
	Minimum		18.151	
	Maximum		42.136	
	Range		23.985	
	Interquartile Range		3.145	
	Skewness		2.775	.295
	Kurtosis		10.373	.582
Touch Task Time	Mean		20.30665	.221768
	95% Confidence Interval for Mean	Lower Bound	19.86375	
		Upper Bound	20.74955	
	5% Trimmed Mean		20.22218	
	Median		20.07550	
	Variance		3.246	
	Std. Deviation		1.801652	
	Minimum		16.987	
	Maximum		25.207	
	Range		8.220	
	Interquartile Range		1.962	
	Skewness		.773	.295
	Kurtosis		.202	.582

B15: Multi Directional Tapping Test (Context) – Error

			Statistic	Std. Error
Mouse Task Error	Mean		.36232	.094547
	95% Confidence Interval for Mean	Lower Bound	.17365	
		Upper Bound	.55098	
	5% Trimmed Mean		.22705	
	Median		.00000	
	Variance		.617	
	Std. Deviation		.785363	
	Minimum		.000	
	Maximum		4.000	
	Range		4.000	
	Interquartile Range		.500	
	Skewness		2.818	.289
	Kurtosis		8.776	.570
Stylus Task Error	Mean		.24638	.063319
	95% Confidence Interval for Mean	Lower Bound	.12003	
		Upper Bound	.37273	
	5% Trimmed Mean		.16989	
	Median		.00000	
	Variance		.277	
	Std. Deviation		.525967	
	Minimum		.000	
	Maximum		2.000	
	Range		2.000	
	Interquartile Range		.000	
	Skewness		2.088	.289
	Kurtosis		3.605	.570
Touch Task Error	Mean		.42029	.090950
	95% Confidence Interval for Mean	Lower Bound	.23880	
		Upper Bound	.60178	
	5% Trimmed Mean		.32367	
	Median		.00000	
	Variance		.571	
	Std. Deviation		.755486	
	Minimum		.000	
	Maximum		4.000	
	Range		4.000	
	Interquartile Range		1.000	
	Skewness		2.281	.289
	Kurtosis		6.652	.570

REFERENCES

Accot, J. and Zhai, S. (1999). Performance evaluation of input devices in trajectory-based tasks: an application of the steering law. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 466-472). ACM.

Ahlström, B., Lenman, S. and Marmolin, T. (1992). Overcoming touchscreen user fatigue by workplace design. *Proceedings of ACM Conference on Human Factors in Computing Systems CHI 92:* 101-102

Ahsanullah, M. A. K. B. and Sulaiman, S. (2010). Investigation of fingertip blobs on optical multi-touch screen, *Information Technology (ITSim), 2010 International Symposium,* vol.1, no., pp.1-6, 15-17

Albinsson, P. and Zhai, S. (2003). High precision touch screen interaction. *ACM CHI Conference on Human Factors in Computing Systems.* p. 105-112.

Arnott, N. (2013). *iPhone Touchscreen Accuracy – A lesson in understanding test requirements and goals.* Retrieved October 25, 2014 from: http://www.neglectedpotential.com/2013/10/iphone-touchscreen-accuracy/

Astala, A. & Mansikkaniemi, T. (2003). *U.S. Patent No. 6,590,568*. Washington, DC: U.S. Patent and Trademark Office.

Atkinson, S., Woods, V., Haslam, R. A. and Buckle, P. (2004). Using non-keyboard input devices: interviews with users in the workplace. *International Journal of Industrial Ergonomics*, *33*(6), 571-579

Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher*, *32*(1), 21-24.

Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. (2003). Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch and pen operated systems. *Proceedings of Interact.* p. 57-64.

Benko, H., Wilson, A. and Baudisch, P. (2006). Precise selection techniques for multi-touch screens. *Proc. of CHI' 06*, 1263-1272.

Betts, P., Brown, C. J., Lynott, J. J. and Martin, H. F. (1965). Light Beam Matrix Input Terminal, *IBM Technical Disclosure Journal*, 9(5), 493-494

Bevan, N. & Curson, I. (1999, May). Planning and implementing user-centred design. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems* (pp. 137-138). ACM.

Bevan, N. & Macleod, M. (1994). Usability measurement in context, *Behaviour & Information Technology,* 13:1-2, 132-145.

Bevan, N. (2001). International standards for HCI and usability. *International journal of human computer studies*, *55*(4), 533-552.

Bevan, N. (2009). International standards for usability should be more widely used. *Journal of Usability Studies*, *4*(3), 106-113.

Bhalla, M. R. and Bhalla, A. V. (2010). Comparative study of various touchscreen technologies. *International Journal of Computer Applications*, 6(8), 12-18.

Bolton, W. (2009). *Programmable logic controllers* (5th ed.). Newnes.

Burkhard, M. & Koch, M. (2012). Evaluating Touchscreen Interfaces of Tablet Computers for Elderly People. In *Mensch & Computer Workshopband* (pp. 53-59).

Buxton, B. (2010). A Touching Story: A Personal Perspective on the History of Touch Interfaces Past and Future. *Society for Information Display (SID) Symposium Digest of Technical Papers,* May 2010, Volume 41(1), Session 31, 444-448.

Carmody, T. (2010). *Why 'Gorilla Arm Syndrome' Rules Out Multitouch Notebook Displays*. Wired. Retrieved December 30, 2015 from: http://www.wired.com/2010/10/gorilla-arm-multitouch/

Carr, M. & Verner, J. (1997). *Prototyping and Software Development Approaches*. Working Paper. Department of Information Systems. City University of Hong Kong.

Clark, C. and Harper, P. (2002). Emerging Technologies. *Discussion paper, Australian Association of the Deaf, Sydney*. Retrieved October 17, 2011 from: http://www.aad.org.au/download/EmergingTechnology.pdf

Clark, J. (2012). *New Rule: Every Desktop Design Has To Go Finger-Friendly*. Global Moxie. Retrieved December 30, 2015 from: http://globalmoxie.com/blog/desktop-touch-design.shtml

Collomb, M., Hascoët, M., Baudisch, P. & Lee, B. (2005, May). Improving drag-and-drop on wall-size displays. In *Proceedings of Graphics interface 2005* (pp. 25-32). Canadian Human-Computer Communications Society.

Davis, A. M. (1995). Tracing: A simple necessity neglected. *IEEE Software,* (5), 6-7.

Dix, A., Finlay, J., Abowd, G. D., & Beal, R. (2004). Interaction design basics. *Ch. 5 in human computer interaction,* 189-224.

Douglas, S. A., Kirkpatrick, A. E. & MacKenzie, I. S. (1999). Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 215-222). ACM.

Ebeling, F.A., Goldhor, R.S. and Johnson, R.L. (1972). A Scanned Infrared Light Beam Touch Entry System, *Proceedings of Society of Information Displays*, 134-135

Ezor, J. (2010). *Palm Pre screen tests: color great, touch accuracy not so much.* Retrieved October 24, 2014 from: http://www.webosnation.com/palm-pre-screen-tests-color-great-touch-accuracy-not-so-much.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, *47*(6), 381.

Floyd, C. (1984). A Systematic Look at Prototyping, in: Budde, R., Kuhlenkamp, K., Mathiassen, L. and Zullighoven, H. (Eds.) *Approaches to Prototyping*, Springer-Verlag: Heidelberg, 1-17.

Fuglerud, K. S. and Rossvoll, T. H. (2010). Previous and Related Research on Usability and Accessibility Issues of Personal Identification Management Systems. *Norwegian Computing Center, Oslo (Norway), Tech. Rep. DART/10/10.*

Golshani, F. (2007). TUI or GUI–it's a matter of Somatics. *Multimedia, IEEE*, vol. 14, no. 1.

Gorilla Arm. (1996). In *New Hacker's Dictionary.* The MIT Press.

Gregor, P., Newell, A. F. & Zajicek, M. (2002, July). Designing for dynamic diversity: interfaces for older people. In *Proceedings of the fifth international ACM conference on Assistive technologies* (pp. 151-156). ACM.

Hansen, T. E., Hourcade, J. P., Virbel, M., Patali, S. & Serra, T. (2009, November). PyMT: a post-WIMP multi-touch user interface toolkit. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (pp. 17-24). ACM.

Helander, M. G., Landauer, T. K. & Prabhu, P. V. (Eds.). (1997). *Handbook of human-computer interaction*. Elsevier.

Hlady, A. M. (1969). A Touch Sensitive X-Y Position Encoder for Computer Input, *Proceedings of the AFIPS Fall Joint Computer Conference*, Vol. 35, 545

Holzinger, A. and Miesenberger, K. (Eds.). (2009). *HCI and Usability for e-Inclusion: 5th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2009, Linz, Austria, November 9-10, 2009, Proceedings* (Vol. 5889). Springer.

Horsley, M., Eliot, M., Knight, B. A. & Reilly, R. (Eds.). (2014). *Current Trends in Eye Tracking Research*. Springer.

ISO (1997). *ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs).* Part 11 - Guidelines for specifying and measuring usability. International Standards Organisation. Also available from the British Standards Institute, London.

ISO (1998). *ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs).* Part 11: Guidance on usability. International Standards Organisation. Also available from the British Standards Institute, London.

ISO (1999). *ISO 13407: Human-centred Design Processes for Interactive Systems.* Geneva: International Standards Organisation. Also available from the British Standards Institute, London.

ISO (2000). *9241-9: Ergonomic requirements for office work with visual display terminals (VDTs).* Part 9: Requirements for non-keyboard input devices. International Standards Organisation. Also available from the British Standards Institute, London.

ISO (2011). *9241- 420: Ergonomics of human-system interaction* — Part 420: Selection of physical input devices. Also available from the British Standards Institute, London.

ISO (2008). *9241-410: Ergonomics of human-system interaction*. Part 410: Design criteria for physical input devices. Also available from the British Standards Institute, London.

Jin, Z. X., Plocher, T. and Kiff, L. (2007). Touch screen user interfaces for older adults: button size and spacing. In *Universal acess in human computer interaction. coping with diversity* (pp. 933-941). Springer Berlin Heidelberg.

Jota, R., Forlines, C., Leigh, D., Sanders, S. & Wigdor, D. (2014). Towards Zero-Latency User Experiences. *Tactual Labs Co.*

Kim, H. K., Lee, S. and Yun, K. S. (2011). Capacitive tactile sensor array for touch screen application. *Sensors and Actuators A: Physical*, *165*(1), 2-7.

Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M. & Ifukube, T. (2011). Elderly user evaluation of mobile touchscreen interactions. In *Human-Computer Interaction INTERACT 2011* (pp. 83-99). Springer Berlin Heidelberg.

Langston, W. (2014). Langston, *Research Methods*, *Notes 10 -- Within Participants Designs*, Middle Tennessee State University. Retrieved December 27, 2015 from: http://capone.mtsu.edu/wlangsto/RMNotes10.html

Lazar, J. (Ed.). (2007). *Universal usability: designing computer interfaces for diverse user populations*. John Wiley & Sons.

Lazar, J., Feng, J. H. and Hochheiser, H. (2010). *Research methods in human-computer interaction*. John Wiley & Sons.

Lee, S. K., Buxton, W. and Smith, K. C. (1985). A multi-touch three dimensional touch-sensitive tablet, *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI'85)*, 21-27

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*(140), 1-55.

Lund, A. & Lund, M. (2013). *One-way repeated measures ANOVA in SPSS Statistics*. Retrieved December 31, 2015, from: https://statistics.laerd.com/premium/rma/repeated-measures-anova-in-spss.php

Lyle, K. B., Hanaver-Torrez, S. D., Hackländer, R. P. & Edlin, J. M. (2012). Consistency of handedness, regardless of direction, predicts baseline memory accuracy and potential for memory enhancement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(1), 187.

Macleod, M. & Rengger, R. (1993). The development of DRUM: A software tool for video-assisted usability evaluation. *People and Computers,* 293-293.

Macleod, M., Bowden, R., Bevan, N. and Curson, I. (1997). *The MUSiC Performance Measurement Method*. Behaviour and Information Technology, 16

Maguire, M. (2001b). Methods to support human-centred design. *International Journal of Human-Computer Studies,* 55, 4 587-634.

Maguire, M. C. (2001a). Context of use within usability activities. *International Journal of Human-Computer Studies*, 55, 453}483.

McGookin, D., Brewster, S. & Jiang, W. (2008, October). Investigating touchscreen accessibility for people with visual impairments. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (pp. 298-307). ACM.

Moin (2010). *Touch Test: Touch Screen Test App For Your Android Smartphone*. Retrieved October 24, 2014 from: from http://www.makeuseof.com/tag/touch-test-touch-screen-test-app/

Motti, L. G., Vigouroux, N. & Gorce, P. (2013, September). Interaction techniques for older = adults using touchscreen devices: a literature review. In *25ème conférence francophone sur l'Interaction Homme-Machine, IHM'13*.

Muller, M. J. (2003). Participatory design: the third space in HCI. *Human-computer interaction: Development process, 4235*.

Nauman, J. D. & Jenkins, M. (1982). Prototyping: The New Paradigm for Systems Development, *MIS Quarterly*, 6, 3, 29-44.

Nicolau, H., Guerreiro, T., Jorge, J., & Gonçalves, D. (2014). Mobile touchscreen user interfaces: bridging the gap between motor-impaired and able-bodied users. *Universal Access in the Information Society*, *13*(3), 303-313.

Nielsen, J. & Landauer, T. K. (1993, May). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). ACM.

O'Neill, E., Kaenampornpan, M., Kostakos, V., Warr, A., & Woodgate, D. (2006). Can we do without GUIs? Gesture and speech interaction with a patient information system. *Personal and Ubiquitous Computing*, *10*(5), 269-283.

Okada, H. & Akiba, T. (2010). *Fitts' Law Index of Difficulty Evaluated and Extended for Screen Size Variations*. INTECH Open Access Publisher.

Olsen, D. (2009). *Building interactive systems: principles for human-computer interaction*. Cengage Learning.

Osborne, J. W. and Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, 9(6). Retrieved December 26, 2015 from: http://PAREonline.net/getvn.asp?v=9&n=6

Ostroff, D. and Shneiderman, B. (1988). Selection devices for users of an electronic encyclopedia: An empirical comparison of four possibilities, *Information Processing & Management*, 24(6), 665-680.

Padre, J. (2014). *Understanding touch responsiveness – Touchscreen technology series 2*. Retrieved October 27, 2014 from: http://developer.sonymobile.com/2014/07/02/understanding-touch-responsiveness touchscreen-technology-series-2/

Peters, M. & Murphy, K. (1992). Cluster analysis reveals at least three, and possibly five, distinct handedness groups. *Neuropsychologia, 30,* 373–380.

Pickering, J. A. (1986). Touch-sensitive screens: the technologies and their application, *International Journal Man-Machine Studies,* 25 249-269.

Potter, R., Shneiderman, B. and Weldon, L. (1988). Improving the accuracy of touch screens: an experimental evaluation of three strategies. *Proceedings of CHI'88,* 27-32.

Preece, J. (2000). *Online communities: Designing usability and supporting socialbilty*. John Wiley & Sons, Inc.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. (1994). *Human-computer interaction*. Addison-Wesley Longman Ltd..

Reiss, E. (2012). *Usable usability: simple steps for making stuff better*. John Wiley & Sons.

Rogers, Y. (2012). HCI theory: classical, modern, and contemporary. *Synthesis Lectures on Human-Centered Informatics, 5*(2), 1-129

Schöning, J., Brandl, P., Daiber, F., Echtler, F., Hilliges, O., Hook, J. & Roth, T. (2008). Multi-touch surfaces: A technical guide. *IEEE Tabletops and Interactive Surfaces, 2,* 11.

Sears, A. (1991). Improving Touchscreen Keyboards: Design issues and a comparison with other devices. *Interacting with Computers, 3,* 3, 252-269.

Seow, S. C. (2005). Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. *Human-Computer Interaction*, *20*(3), 315-352.

Shneiderman, B. (2003, November). Promoting universal usability with multi-layer interface design. In *ACM SIGCAPH Computers and the Physically Handicapped* (No. 73-74, pp. 1-8). ACM.

Soloway, E., Guzdial, M. & Hay, K. E. (1994). Learner-centered design: The challenge for HCI in the 21st century. *Interactions*, *1*(2), 36-48.

Swann, M. (2006). *Ergonomics of Touch Screens.* Ergonomic Solutions International. Retrieved December 31, 2015 from: http://www.ergonomic-solutions.net/wp content/uploads/2015/04/Touch_Screen_Ergonomics_06.pdf

The Economist (2008). Case History, Touching the future, Technology Quarterly: Q3 2008. Retrieved October 4, 2011 from: http://www.economist.com/node/11999181?story_id=11999181 (Accessed 4th October 2011)

Tode, J., Wilson, D. & Wynn, B. (2001). *U.S. Patent No. 6,201,532*. Washington, DC: U.S. Patent and Trademark Office.

U.S. Environmental Protection Agency (2007). *Guidance for Preparing Standard Operating Procedures* (QA/G-6). EPA/600/B-07/001. April 2007. Retrieved December 28, 2015 from: http://www.epa.gov/sites/production/files/2015-06/documents/g6-final.pdf

University of Strathclyde (2014). Design Strategies in Quantitative Research - Between and Within Participant Designs. Retrieved October 4, 2014 from: http://www.strath.ac.uk/aer/materials/2designstrategiesineducationalresearch/unit2/betweenandwithinparticipantdesigns/

Vogel, D. and Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 657-666). ACM.

Wang, E. M. Y. and Shih, S. S. Y. (2009). A study on thumb and index finger operated interface for personal mobile devices: Mobile phone keypad and joystick. In *Industrial Engineering and Engineering Management, 2009. IE&EM'09. 16th International Conference on* (pp. 766-770). IEEE.

Wang, F. and Ren, X. (2009). Empirical evaluation for finger input properties in multi-touch interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1063-1072). ACM

Wigdor, D. and Wixon, D. (2011). *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier.

Wu, M. and Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology* (pp. 193-202). ACM.

Zabramski, S., Shrestha, S. and Stuerzlinger, W. (2013). Easy vs. Tricky: The Shape Effect in Tracing, Selecting, and Steering With Mouse, Stylus, and Touch. In *Proceedings of International Conference on Making Sense of Converging Media* (p. 99). ACM.

Zhai, S., Kong, J. & Ren, X. (2004). Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International journal of human-computer studies, 61*(6), 823-856.

Zhang, P., Carey, J., Te'eni, D. and Tremaine, M. (2005). Integrating human-computer interaction development into the systems development life cycle: a methodology. *Communications of the Association for Information Systems (Volume 15, 2005), 512*(543), 543

Zhang, X. & MacKenzie, I. S. (2007). Evaluating eye tracking with ISO 9241-part 9. In *Human Computer Interaction. HCI Intelligent Multimodal Interaction Environments* (pp. 779 788). Springer Berlin Heidelberg.

Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R. and Baudisch, P. (2007). Earpod: eyes-free menu selection using touch input and reactive audio feedback. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1395-1404). ACM.

Zhou, X. and Ren, X. (2010). An investigation of subjective operational biases in steering tasks evaluation. *Behaviour & Information Technology*, *29*(2), 125-135.