

**A Usability Study of Multimodal Field Data Entry
Techniques for the Construction Industry**

Version 2.5

By
Nathan Langton

University of New Brunswick
Faculty of Computer Science

National Research Council of Canada
Institute for Information Technology
Fredericton, New Brunswick



March 29, 2006

ABSTRACT

Tablet PCs, laptops and handheld devices, such as cell phones and PDAs, are becoming increasingly more powerful, with faster chips and more memory, opening the doors for a wide variety of potentially more practical mobile applications. Handheld devices are becoming increasingly smaller in an effort to make them more mobile, which results in restrictions and limitations to the user interface. Multiple modes of interaction are quickly becoming a necessity. One of the significant areas of development today, in terms of mobile interaction, is voice technology which allows the use of speech as a mode of input and output.

As part of an ongoing investigation into multimodal technology and input techniques at the National Research Council of Canada, I helped create the Multimodal Field Data Entry (MFDE) prototype. The purpose of the MFDE system is to enable a technician/worker located on a construction site to record data using the MFDE application on his or her portable mobile device, such as a laptop, tablet or a Pocket PC (PPC). Since this application is multimodal, the technician can choose to enter the data using voice, stylus or a keyboard. This is especially convenient if their hands are dirty or occupied with machinery, equipment or instruments.

This paper focuses specifically on conducting a usability study for the MFDE application. We conducted a lab-based usability study of this prototype application in order to (a) determine and compare the effectiveness and usability of the two different input options (speech input v. stylus-based input) and (b) determine which of the two input options is preferred by users. This study also explored the viability of multimodal data entry techniques being used in work environments, such as the construction industry. Key construction site elements were reflected in the lab-setup which included three construction noise levels, mobility of the participants, and various distractions such as heavy machinery.

This paper will also report results and conclusions pertaining to speech accuracy in environments with various levels of industrial noise, how users interacted with the application while presented with distractions, and users' subjective preferences.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who made it possible for me to complete this thesis. This study could not have been completed without their help.

Thanks to my supervisors Professor Natalie Webber (Faculty of Computer Science), Dr. Joanna Lumsden (National Research Council), and Dr. Irina Kondratova (National Research Council) who helped by giving valuable advice, guidance and encouragement throughout my honors thesis.

I am deeply indebted to Dr. Irina Kondratova from the National Research Council - Institute for Information Technology (e-business) who gave me employment for three work terms as a co-op student enabling me to continue my collaboration with NRC for my honors thesis. I count myself fortunate to be given the opportunity to learn from excellent researchers and gain experience by working at NRC.

My appreciation and thanks to the National Research Council, who provided me with access to the many resources that were needed throughout the study including the use of the HCI and Voice Laboratories. Thanks to the employees who helped by sharing their time and knowledge. A special mention to Danny Damours at NRC for helping me with all types of technical problems that I encountered, Bob Longworth who helped by creating a server and application for running the projectors, and Scott Durling who assisted me with various problems and helped with the experimental setup.

I would also like to express my thanks to Dr. Jeff Rankin (UNB Civil Engineering Department) for giving me material that was useful in this study and Dr. Maureen Tingley (Department of Mathematics and Statistics) for her statistical expertise.

Finally, I would like to thank my fiancée Mari-Kaye and my family for their love, encouragement and support throughout this university year.

Table of Contents

ABSTRACT	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	IV
1 INTRODUCTION	6
2 BACKGROUND	7
2.1 Mobile Device Limitations	7
2.2 Voice Enabled Systems	8
2.3 Multimodality	9
2.4 XHTML+VoiceXML Multimodal Markup language	10
3 MULTIMODAL FIELD DATA ENTRY PROTOTYPE	12
3.1 Background	12
3.2 Purpose	12
FIGURE 1:	13
3.3 Functionality	14
3.4 Usage Scenarios	16
4 EVALUATION DESIGN AND PROCEDURE	17
4.1 Evaluation Environment	17
4.1-1 Level of Mobility	18
4.1-2 Construction Noise Levels and Safety Precautions	18
4.1-3 Heavy Equipment	21
4.2 Experimental Forms	23
4.3 Experimental Process	26
4.3-1 Data Collected and Recorded	27
4.3-2 Experimental Procedure	28
4.3-3 Experimental Hypotheses	31

5	RESULTS AND DISCUSSIONS	32
5.1	Participant Background Data.....	32
5.2	Task Completion Time	33
5.3	Average Percentage of Distractions Identified.....	34
5.4	Average Entry Precision Rates	36
5.5	Data Type Precision Rates	39
5.5-1	Number and Date – Entry Precision Rates.....	40
5.5-2	Drop-Down Lists – Entry Precision Rates.....	41
5.5-3	Decimal – Entry Precision Rates	42
5.5-4	Accuracy Rates among Data Types.....	43
5.5-5	Accuracy Rates across Input-Styles and Data Types	43
5.6	Questionnaire Results.....	44
5.7	Preferences Results	45
5.8	Experimental Limitations	47
6	CONCLUSIONS AND FUTURE WORK	47
	BIBLIOGRAPHY	50
	APPENDIX A: REB ETHICS APPLICATION	52
	APPENDIX B: EXPERIMENT INSTRUCTIONS	64
	APPENDIX C: EXPERIMENT RESULTS	84
	APPENDIX D: EXPERIMENT STATISTICS	91
	APPENDIX E: RESEARCH PAPER ACCEPTED	106
	APPENDIX F: HONOURS THESIS SCHEDULE/TIMETABLE	107

1 Introduction

Tablet PCs, laptops and handheld devices such as cell phones and PDAs are becoming increasingly more powerful with faster chips and more memory, opening the doors for a wide variety of potentially more practical mobile applications. Handheld devices are becoming increasingly smaller in an effort to make them more mobile which results in restrictions and limitations to the user interface. Multiple modes of interaction are quickly becoming a necessity. One of the significant areas of development today, in terms of mobile interaction, is Voice Technology which allows the use of speech as a mode of input and output.

While working at the National Research Council of Canada, I helped develop the Multimodal Field Data Entry System (MFDE). This multimodal prototype allows workers to enter field information using a mobile device such as a wireless pocket PC, a laptop or a tablet PC. The worker has the option to enter data either using speech or a stylus. MFDE was designed specifically to allow concrete technicians or engineers conducting tests in the field to have quick and convenient access to forms which they can fill out using speech. This is especially convenient if their hands are dirty or occupied with machinery, equipment or instruments.

Before this prototype can be proven and made commercially available, extensive testing needs to be conducted. As a part of my honours thesis, I have conducted a usability study of the MFDE application. The research that is presented in this paper is part of an ongoing investigation into the feasibility of data entry techniques using different modalities with mobile devices. This paper explores the viability of multimodal data entry techniques for use in noisy and dangerous work environments such as the construction industry. It discusses the development, execution and results of an experiment to evaluate the multimodal MFDE design, and from this draws conclusions and recommendations to help enhance the prototype application.

2 Background

As mobile technology advances, it becomes possible to develop more practical mobile applications which allow mobile devices such as tablet PCs, laptops and PDAs to be used more effectively in work environments. Mobile applications can be created to suit the needs of various industries such as engineering, forestry, mining, and health services. This new form of e-Business/e-Commerce brings new ideas and technologies into the real world where people are constantly on the move.

2.1 Mobile Device Limitations

As mobile devices become increasingly smaller, this imposes limitations and restrictions on the user interface. One such limitation is the small screen size on most mobile devices which can create an inconvenience if an application's primary mode of input requires the use of a pen or stylus to enter data and commands. Smaller devices usually have smaller buttons and keyboards which can also make stylus data input difficult.

Physical work environments may not be suitable for using the traditional data entry techniques such as a pen or stylus with mobile technology. For example, a worker in the field may be operating equipment where both hands are in use. Interacting with a mobile device in this situation may not, therefore, be convenient and, more significantly, could be potentially be dangerous by distracting the user and making him unaware of his external surroundings.

Multiple modes of interaction are quickly becoming a necessity given the user interface restrictions mentioned above. One advance in mobile interaction techniques is Voice Technology that allows the use of speech as a mode of input and output.

2.2 Voice Enabled Systems

Voice Enabled Systems have an interface that utilizes a speech recognition engine (input) and a speech synthesis engine (output) to enable a user to interact with a system using spoken phrases.

Speech recognition is the technology that “*enables computers or other electronic systems to identify the sound of a human voice, separate that sound from noise in the environment, and accept the message from the voice as input for controlling the system*” [1, p. 98]. This is also known as Automatic Speech Recognition (ASR).

Speech synthesis refers to the technologies that enable computers or other electronic systems to output simulated human speech known as text-to-speech (TTS) [1].

These technologies allow the user to “speak” to the system, and the system to “speak” to the user. Voice Interfaces provide a natural way for humans to interact with devices, just as we interact with humans.

Voice Interfaces provide an easy-to-use environment because everyone knows how to speak, but not everyone knows how to interact with a computer or mobile device using a keyboard, stylus or a mouse. Voice Interfaces can also provide a beneficial way to interact for users with disabilities, like the visually impaired or those with physical disabilities.

Voice Interfaces also normally support hands-free interaction. It can be very useful in situations where a user needs the use of his hands for his primary task, such as driving a car or operating other equipment and instruments. “*Speech processing is one of the key technologies to simplifying the use of handheld devices for mobile workers*” [2, 3, 4 p. 345]

Voice Technology creates many possibilities for mobile solutions to address the challenges of data entry in work environments where the traditional methods would be either impossible or impractical. *Multimodal* applications allow for multiple ways of interaction providing flexible and efficient ways of entering data beneficial to many work environments.

2.3 Multimodality

A multimodal application allows the end-users to interact with the devices interface in a flexible manner using a combination of keyboard, touch screen, stylus, telephone keys, gestures and voice. Each of these modes can be used independently or concurrently while interacting with the application.

The most common multimodal interfaces integrate visual modality (e.g. a display, keyboard, stylus, and mouse) with voice modality. Embedded voice modality requires a speech recognition engine (i.e. ASR) for input and speech synthesis (i.e. TTS) and/or recorded audio for output. With multimodal technologies you can speech-enable the visible elements of a visual interface, which has particular importance for mobile users [5].

Humans are multimodal by nature, as we move dynamically between visual and audio communications, often without being aware of the shifts. [6] The choice of multiple modalities for interaction and presentation enables the user to receive and input information efficiently using a mobile device, which might have limited input, output and display capabilities. The ability to switch between interaction modes allows for an eyes-free, hands-free interaction using audio-only creating new opportunities to collect data in situations where the traditional technique of stylus/keyboard input are difficult or cannot be used.

Multimodal applications can be implemented and run using a fat client model, which employs embedded speech recognition on the mobile device and allows conducting speech processing locally [7]. Opera and NetFront are two Multimodal browsers which employ embedded IBM speech recognition and text-to-speech technology, allowing users to view and interact with multimodal applications built using markup languages such as SALT and XHTML+VoiceXML (X+V).

2.4 XHTML+VoiceXML Multimodal Markup language

Both SALT and X+V markup languages use W3C standards for grammar and speech synthesis, but only X+V is based entirely on standardized languages [7]. X+V is based on a modular architecture allowing applications to be developed in parts, where experts in voice programming can develop the voice elements and experts in visual programming can develop the visual elements. For these simple reasons we have chosen to make use of the X+V (XML tag based) as the markup language of choice for our prototype implementation.

X+V is a markup language for developing multimodal web pages that combines XHTML and a subset of VoiceXML. The XHTML part dictates the pages visual flow and typically involves laying out elements such as graphics, input fields, text prompts, and checkboxes. More sophisticated interfaces could also include some type of scripting such as JavaScript.

VoiceXML is a dialog-oriented language (W3C standard) that dictates the page prompts and listening flow (such as prompting the user for a desired field within a form). It uses speech and telephone touchtone recognition for input and pre-recorded audio and text-to-speech synthesis (TTS) for output. VoiceXML has been carefully designed to give authors full control over the spoken dialog between the user and the application. The VoiceXML code utilizes an external grammar to define the possible utterances that an end-user may make.

A grammar is essentially a set of utterances (words and phrases) that constitute the acceptable user response to a given prompt [8]. All the words you want the speech recognition engine to recognize when the users respond to your application must be included in the grammar.

In an X+V application, each field or form can specify one or more speech recognition grammars. There are several grammar formats supported by the X+V platform including XML Form of the SRGS (W3C Speech Recognition Grammar Specification), the Augmented BNF (ABNF) Form of the SRGS, and the Java Speech API Grammar Format (JSGF) [9]. We have chosen to use the Java Speech API Grammar Format (JSGF) format as it is platform independent allowing us to implement the prototype on various devices with different platforms.

X+V utilizes a speaker-independent speech recognition engine as opposed to a speaker-dependent speech recognition engine which requires a process of training the computer to recognize an individuals voice. A prototype created using X+V allows for any potential user to pick up the devices and interact immediately with the application. The next section discusses one such prototype created as part of an ongoing investigation into multimodal technology and input techniques.

3 Multimodal Field Data Entry Prototype

The Multimodal Field Data Entry (MFDE) System is a multimodal application prototype that was developed at the NRC Institute for Information Technology (IIT), by Ruairi Meagher and Nathan Langton under the supervision of Dr. Irina Kondratova.

3.1 Background

This project was a joint project between UNB Civil Engineering and NRC IIT. Collaborating with UNB Civil Engineering, we were able to create a generic multimodal application for recording concrete field inspection data.

With new emerging multimodal technology, it seemed that applications for the construction industry might potentially benefit from the combined use of voice and graphical input and output; users could potentially navigate and enter data using a variety of modes such as speech, stylus, pen or keyboard. To examine these potential benefits the MFDE prototype was developed.

Using voice commands with MFDE, a caller can input information to and retrieve information from a concrete placing testing form. This allows a concrete technician to record testing information in the field using a mobile device.

3.2 Purpose

Presently when a field technician or engineer test concrete in the field, they record data using a paper-based system and either transcribe the information into a database when they get back to the office, or phone it into the headquarters to be entered into the corporate system.

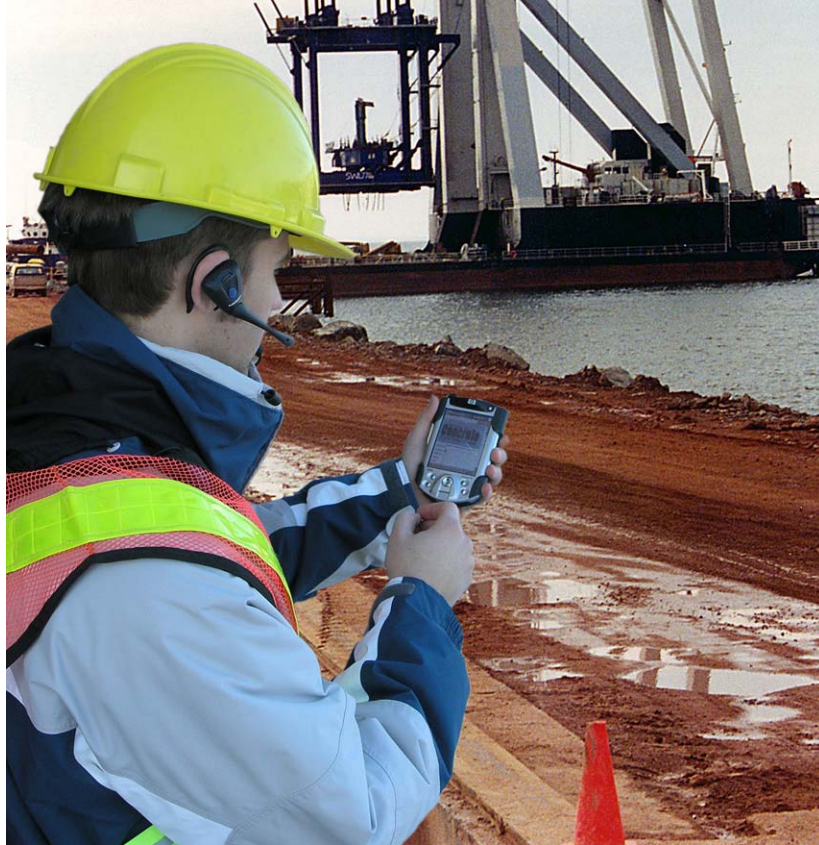


Figure 2: *A field worker using speech to enter data in a multimodal application*

3.3 Functionality

As a prototype, only a limited number of functions have been implemented in the MFDE application. At present, only the “concrete placing report” form has been implemented (see Figure 3(b)). This has three sections: General Information, Placing Information and Testing Information. Each section contains a variety of fields, including text fields (for numbers or dates) and drop-down lists.

The user can perform various actions on the concrete placing reports such as: create a new form, search, edit and delete an existing form, and save a form (see Figure 3(a)).

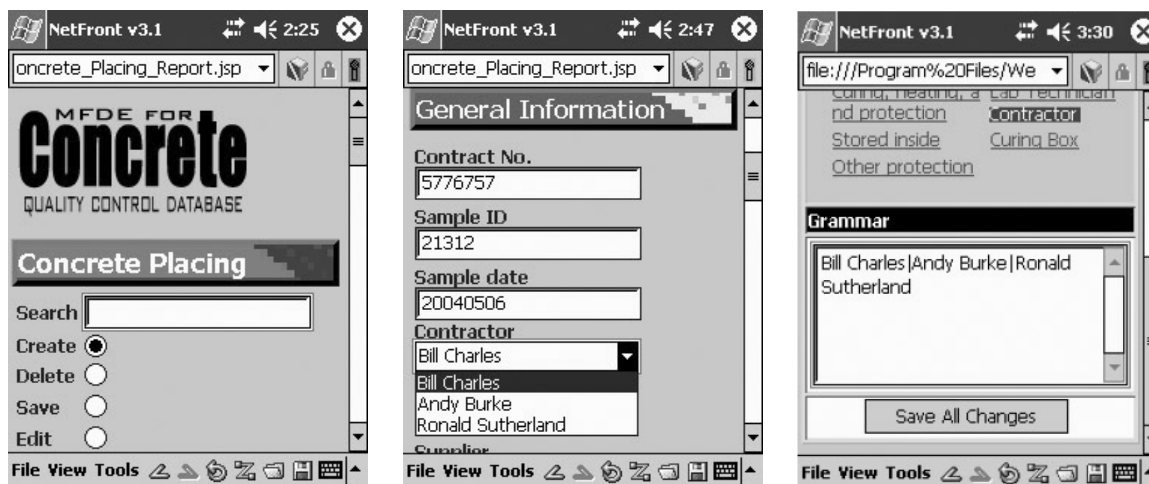


Figure 3 - MFDE application showing: (a) actions that can be performed on a form; (b) an example of a form; and (c) an example of the grammar administration page where grammar and dropdowns can be edited without accessing the source code.

Located under the Administration Menu is an additional feature called “Manage Drop-Down Menus” (see Figure 3(c)). This page allows one to edit the grammar and values displayed in the drop-down menus that can be seen in the concrete placing reports. This additional feature allows the user to modify the options in the grammar and drop-downs, without having to edit source code for the application. The grammar for digits, numbers, dates and various commands was created using the Java Speech Grammar Format (JSGF).

The MFDE application can be deployed and run on a desktop, laptop, tablet, or a Pocket PC (PPC). A Multimodal browser such as NetFront or Opera is needed to interpret the multimodal code.

3.4 Usage Scenarios

A technician/worker in the field (i.e., construction site) would have access to a mobile device with the MFDE application installed. There are two usage scenarios that will be possible with the current prototypic version of the MFDE system (the infrastructure necessary to support the scenarios is shown in Figure 4).

Scenario 1: Imagine the worker is at a location that allows wireless connection; he or she will be able to run the application, accessing and updating content directly from a database through the synchronization server which could be located at the company headquarters.

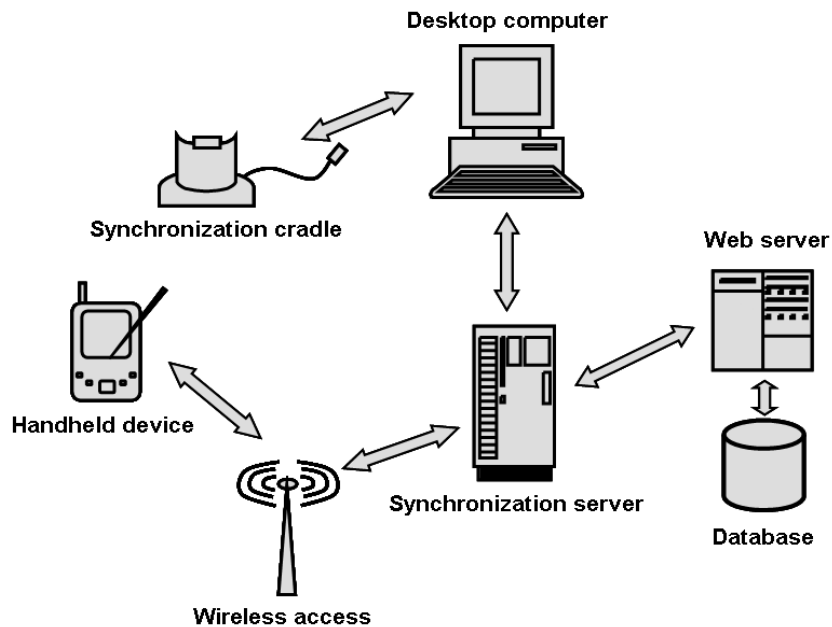


Figure 4: Multimodal Field Data Entry infrastructure [7]

Scenario 2: In the situation where the worker is not in a “wireless zone”, he or she could use the MFDE application as a stand-alone application on the mobile device to create and alter records in an embedded database. This embedded database could then be synchronized with the main corporate database once the worker has entered a wireless zone or connects the mobile device to a synchronization cradle or desktop computer.

4 Evaluation Design and Procedure

To determine and compare the effectiveness and usability of the two multimodal input techniques used in the MFDE application – speech input versus stylus-based input – we developed a usability experiment. Aside from measuring the usability of the two input techniques, the study was also designed to determine which of the two input options the user preferred and to determine the users’ ability to remain aware and cognitive of their surroundings while using the MFDE application.

4.1 Evaluation Environment

For our usability experiment to return meaningful results, we had to simulate the real world environment in which the MFDE application will ultimately be used – that is, a construction site environment. It had to reflect this physical environment as well as be mindful of the fact that the technology is designed to be employed by users who are mobile within that environment.

One might argue that effective evaluations are only achievable in the field, but a recent investigation has shown that there is little or no benefit to undertaking evaluations in the field as opposed to in the lab [10]. The usability experiment took place in a controlled environment to ensure the safety of the study participants, allow for a valid replication of a construction site, and support the recording of necessary study data. This was made possible as a result of NRC’s custom-built Mobile HCI Lab (8.65m x 17.3).

In order to “bring the construction site into the lab” we identified three key environmental factors that would potentially influence a worker’s/technician’s ability to use one or both of the input techniques.

The three key environmental factors that we replicated and represented in the lab were: 1) the level of mobility of the user while interacting with the application; 2)

environmental noise levels that are present on a construction site; and 3) the presence of heavy equipment which can pose potential risk to workers.

4.1-1 Level of Mobility

We required the experimental participants to be mobile in a manner representative of the actual anticipated level of mobility on a construction site. A concrete technician on a construction site will typically approach the concrete pour, pause to conduct the tests and record the relevant data, then proceed to the next relevant location. Typically the technician is not mobile while recording the actual data. To reflect this in our experimental design, we placed tables at various locations in the lab, each containing data which the participant was asked to enter into a form in the MFDE application. During the experimental sessions participants were required to move to each location (table), pause to enter data, and then proceed to the next appropriate location in the lab.

4.1-2 Construction Noise Levels and Safety Precautions

We hypothesised that noise levels on a construction site will detrimentally affect the effectiveness of the speech-based input on the MFDE application, since speech recognition engines often struggle to recognize speech input if background noise is present. To recreate the ambient noise levels inherent on a typical construction site we used a 7.1 surround sound system to deliver construction site sounds within the lab while the participants were interacting with the MFDE application.

Video was obtained from the UNB Faculty of Civil Engineering containing footage of various projects involving construction sites. Audio was extracted from the video footage and spliced together, with the final cut being a conglomerate of various construction sounds ranging from pounding jackhammers to large machinery such as concrete trucks.

In order to evaluate usability of the MFDE application in a construction environment we needed to replicate noise at the same levels found on a construction worksite. The environmental noise on a construction site typically ranges between 70dB and 100dB [11]. The table below shows various pieces of equipment used on a construction site and the level noise they typically generate.

Table 1. Noise levels generated by selected construction equipment (noise level measured at 15 m; used equipment >5 years old).

Equipment	Noise level (dBA)		Source ^a
	New equipment	Used equipment	
Compressor	73	76–80	3, 4
Backhoe	75–80	83–88	1, 3, 5, 6
Concrete mixer	75	85–87	3, 4–6
Concrete pump	75	82–85	1, 3, 4–6
Concrete saw	90	na	1
Concrete vibrator	75	76–80	1, 3
Crane	75	85–90	3, 5
Bulldozer	75	83–88	1–6
Dump truck	84	85–88	1, 3
Excavator	83	85–87	1–4
Grader	72–75	85–90	1, 3–6
Jackhammer	75–80	85–90	1, 3, 5, 6
Loader	80–85	92–95	2, 3
Paver	80–85	95–101	1, 3–5
Rock drill	80–85	95–98	1, 3, 5, 6
Roller	80	88–90	3, 4
Scrapers	78–83	96–98	4–6
Tractor	75	85–90	1, 5, 6
Trencher	83–88	na	4
Vibratory pile driver	95	na	1

Note: na, noise range not available.

^a1, Commonwealth of Massachusetts, Section 721-560; 2, City of Hamilton, Ont., Canada, By-law 79-292; 3, Schexnayder and Ernzen (1999); 4, Eaton (2000); 5, Thumann and Miller (1976); 6, Wilson (1989).

Table 1 – Noise produced by construction equipment [12].

To provide a sense of perspective, a whisper is about 30 dB, conversational speech is about 60 dB, a noisy restaurant or highway traffic is about 70 dB, an alarm clock or hairdryer is about 80 dB, and city traffic is about 90 dB [13].

The experiment was conducted over three ranges of noise levels to enable us to compare the effects that each level of noise might have on speech recognition and the

various data entry techniques. The three groups were: 70-80 dB, 80-90 dB and 90-100 dB. A decibel sound meter was used to measure the sound levels to ensure that each experiment was conducted accurately.

Individuals using power tools often are exposed to noise levels exceeding 85 dB. Exposure to sounds exceeding 85 dB for long periods of time (beyond 8 hours) may cause hearing damage [13, 14]. The Canadian Centre for Occupational Health and Safety (CCOHS) recommends that people wear hearing protection when exposed to noise levels at or above 85 dB [14].

New Brunswick regulations state that a person should not be exposed to 85 dB sound levels for more than eight hours in any one twenty-four hour period. To comply with standard safety procedures, we provided participants with hearing protectors (model: Intruder EM7202 Head Band Earmuff) the use of which was mandatory during the experiment. Not only did this mitigate the risk to our participants, but it also represented a construction site where the worker/technician would be required to wear hearing protection. This model of ear protectors has a noise reduction rate (NRR) of 28 dB. The formula [14] to calculate the estimated noise exposure during our experiment is:

$$\text{Estimated noise exposure} = \text{workplace noise level in dB} - (\text{NRR} - 7)$$

Noise Level	Using Ear Protections
Lowest Noise Level Exposure (70 dB)	70 dB – (28 dB -7) = 49 dB
Highest Noise Level Exposure (100 dB)	100 dB – (28 dB -7) = 79 dB

Table 2 – Noise exposure during the experiment.

By wearing the provided hearing protectors, participants were exposed to 49 dB at the lowest end of our noise levels (70dB) and at the highest noise level they were exposed to 79 dB, which is equivalent to an alarm clock or hair dryer [13].

As can be seen, at no point were participants exposed to noise beyond the 85 dB threshold, and no participant was exposed to more than thirty minutes of noise during the course of the experiment.

Since all the necessary precautions were calculated and taken, there was no more than minimal risk to participants who chose to take part in this study. Furthermore, we obtained informed consent from all participants prior to taking part in this study. If a participant became uncomfortable as a result of the noise, they were free to immediately withdraw from the study; this did not prove necessary for any participant.

This study was conducted with the approval of the UNB Research Ethics Board (see the REB Ethics Application in Appendix A).

4.1-3 Heavy Equipment

Heavy equipment in a construction environment is a safety hazard that can result in fatal accidents if not given due cognizance. Workers must always be cognitive and aware of the dangers that are present in this environment.

This key factor was added to the experiment to determine whether using the MFDE application impedes or otherwise users' ability to remain aware and cognizant of their surroundings, such as heavy equipment in their vicinity.

In order to incorporate this activity into our experiment, without putting participants in harms way, we utilized the six ceiling mounted projectors inside the HCI lab to project, at random time intervals, photographic images of heavy equipment and construction sites around the four walls of the lab space.

There were two kinds of images displayed on the walls: a series of “safe” images and a unique “danger” image. The series of “safe” construction site photographs are considered safe because they contain no heavy equipment that might pose a threat to an individual working in that environment. There was one “danger” image which contains a cement truck as shown in Figure 5(b).

The images were displayed in a pre-set sequence with a random location and duration such that all the projectors were simultaneously projecting photographs but only one occurrence of the “danger” photograph was displayed at any point in time. No “danger” picture was displayed sequentially on the same projector. Figure 5 shows an example of a “safe” photograph and the “danger” photograph.



Figure 5 – *Visual Distractions Used – (a) An Example of a ‘Safe’ Visual Distraction and (b) the ‘Danger’ distraction.*

Whilst using the MFDE prototype to enter data, the participants were required to be cognitive of their surroundings; they were asked to keep a mental tally of the number of “danger” images they believe they saw during the experiment and report this number back to the coordinator once the experiment was completed. Using this information, we are able to determine if the two modalities, levels of noise or a combination of both play a key factor in the participants’ ability to remain aware of their surroundings. The results are discussed in section 5.

4.2 Experimental Forms

The MFDE application includes four types of data entry fields: numbers, decimal numbers, dates and a series of drop-down lists for selection of pre-entered data. The current MFDE application contains many fields, navigation options, prompts and other commands and features, making it too complex and cumbersome for use in an experiment. For the purpose of this experimental study, a smaller abridged version of the MFDE data entry form was used. This form was loaded using the NetFront multimodal browser on a HP 4700 series iPAQ for the experiment.

A subset of seven actual fields in the MFDE application was used which included one or more occurrences of each of the four different field types. The MFDE fields were based on extensive research by a Civil Engineering graduate student and represent an accurate reflection of the types of entries a worker/technician would enter during concrete testing. For this reason, the number of occurrences of each field type in our experimental form was determined by the frequency of their use in the full MFDE application.

For example, the date field is only used once in each MFDE form, therefore we included only one date field in the experiment. Decimal numbers and drop-down lists are the most frequently occurring data fields and as such they were included several times in the experiment form.

As will be discussed in the section 4.3, the experiment included two sessions to allow us to compare the two input techniques (speech-based input and stylus-based input) for each data entry type. Two forms (see Figure 6) were created for the experiment, one form for each session. The forms used in each of the two sessions contained the same data fields, but the order in which the fields were presented and the precise data elements to be entered were different to mitigate against learning effects.

Figure 6 (a) - Example of Session 1 Form **Figure 6 (b)** - Example of Session 2 Form

For both sessions the precise data elements to be entered were balanced in complexity to eliminate any potential bias that may have arisen due to some data elements being perceived as easier than others. The precise data was carefully selected to reflect real examples of entries and formats used on construction site paper-based forms. Table 3 shows the seven field names, the data type and precise data entered for each session.

<u>Field Name</u>	<u>Data Type</u>	<u>Session 1</u>	<u>Session 2</u>
Contract Number	Number	258391	007452
Date	Date	2005-11-22	2006-06-02
Material Used	Drop-Down List	Steel	Concrete
Contractor	Drop-Down List	Alex Boyd	Sally Wilson
Location	Drop-Down List	Lower Fredericton	Plaster Rock
Shipped Concrete	Decimal Number	30.8	9.058
Placed Concrete	Decimal Number	82.09	14.37

Table 3 – Data types, field names and precise data used for both sessions

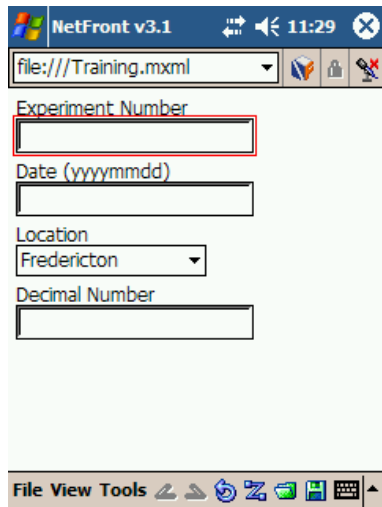
The precise data entries for “number” fields were made up of six digits each, containing a combination of odd and even digits. Session two contained a series of digits with leading zeros, which follows the format frequently used to reference a particular form in a large database pool of information.

In drop-down lists, each pre-entered data item that was chosen for selection was placed in a pre-determined position in the list, to balance the searching through lists in each session. For example, each session contains one drop-down entry found at the top of a field list, one near the middle and one found at the bottom of the list.

Like the “number” fields, the “decimal” fields were also chosen to include a balanced combination of even and odd numbers. Several formats were chosen, with one decimal number containing two digits followed by a period and one more digit (i.e. 30.8) whereas another decimal number contained one digit followed by a period and three digits (i.e. 9.058). Two digits before and after the period (i.e. 82.09) were also used.

Before each session (speech-based or stylus-based), participants were trained in the use of the input technique that they were about to use. As part of this training the participants were given a sample four-field form, giving them the opportunity to practice with the input technique on all four data type fields. This training was conducted in a similar manner to the experimental tests, with participants walking between four training tables, reading the instructions, and entering the supplied data into the form.

Figure 7 shows the training form which includes all four data types. See the Experiment Instructions in Appendix B for further information on how participants were trained before the experimental session.



The image shows a screenshot of a web browser window titled "NetFront v3.1". The address bar displays "file:///Training.mxml". Below the address bar, there is a form with the following fields:

- Experiment Number: A text input field with a red border.
- Date (yyyyymmdd): A text input field.
- Location: A dropdown menu with "Fredericton" selected.
- Decimal Number: A text input field.

At the bottom of the browser window, there is a toolbar with icons for "File", "View", and "Tools".

Figure 7 - *The Training Form*

4.3 Experimental Process

As shown in Figure 7, the laboratory set-up contained a 7.1 surround sound to deliver construction site sounds and six ceiling mounted projectors which displayed photographic images of heavy equipment and construction sites around the four walls of the lab space. From the start/end table, participants were asked to walk between seven tables in the lab space (see Figure 7), to enter data into the experimental form provided for that session before returning back to the start/end table.

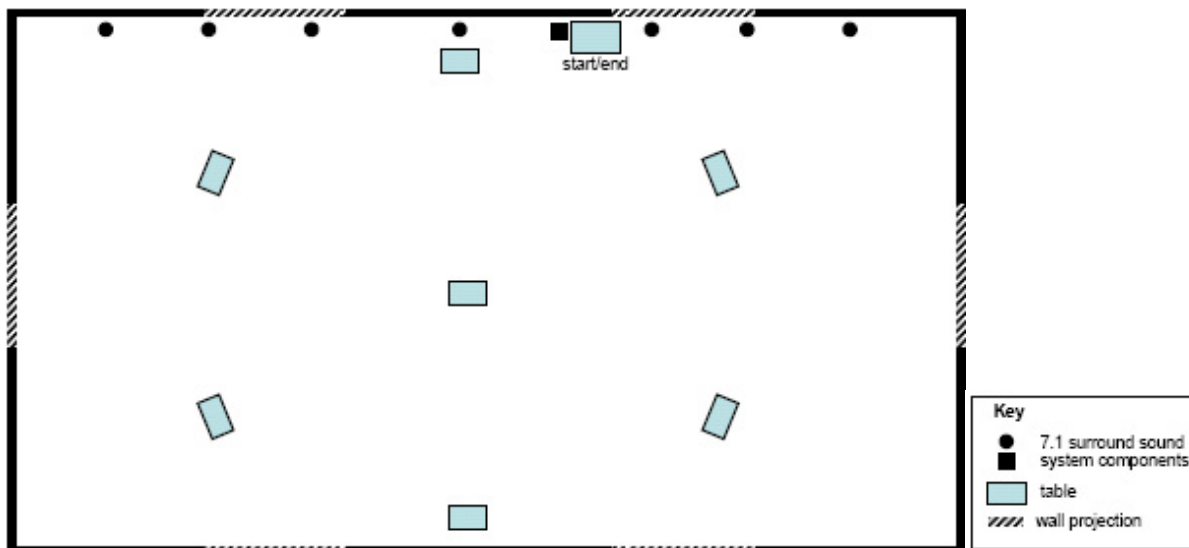


Figure 7 – *HCI Lab Set-up (8.65m x 17.3m)* [14]

Each table in the lab was labeled with a name corresponding to one of the seven field names on the experimental form (field names were shown in Table 3). The participant visited the tables in as the order their corresponding field appeared on the MFDE form. As discussed earlier, two forms were created - one for each session. Although both contained the same fields, they were presented in a different order, which means that participants did not follow the same path during each session, thus further mitigating against the learning effect. Each table in the lab also displayed instructions regarding the data to be entered into the corresponding data field in their form. The data to be entered into the forms also differed between the two sessions.

4.3-1 Data Collected and Recorded

Data obtained from the participants during the experiment was collected and recorded for post experimental evaluation and interpretation. Specifically, the data collected was designed to help us:

- determine and compare the effectiveness and usability of the two different input options (Speech-based input versus Stylus-based input); and
- determine which of the two options is preferred by users in a construction site environment.

During this experiment a range of measures was taken to assess the effectiveness and usability of the two input techniques. We measured the length of time the participants took to complete their MFDE form, the actual data they entered during the experiment, and the number of danger images that they believed had been projected during the experimental session.

Participants were not allowed to correct any erroneous or omitted fields in order to simplify our measure of accuracy. This was done to help us identify which data entry techniques are prone to more errors under various noise conditions and which data types performed better at various noise conditions.

For each modality, participants completed two subjective workload tests using the NASA Task-Load Index (TLX) which is designed to determine participants' perception of workload as they performed the data entry tasks [15, 16]. Individual factors contributing to the measure of workload include: frustration levels, performance levels achieved, mental demand, physical demand and temporal demand. The TLX workload tests used in this experiment are shown in Appendix B – Experiment Instructions.

Lastly we also recorded the participants' overall preferences in terms of data entry technique, and their individual preferences for using a specific data entry technique for each of the four data types.

4.3-2 Experimental Procedure

Participants were divided into three groups, one for each of the three noise levels: Group A at 70-80 dB, Group B at 80-90 dB, and Group C 90-100 dB. The allocation of participants into groups was done according to their order of participation in the experiment; therefore each participant had an equal chance of being allocated to each group. Members in each group participated in the experiment in exactly the same way, the only difference being the noise levels under which they worked.

A total of 18 participants were divided equally among the three groups. Within each group, the 6 participants were required to complete the experiment twice, once using speech and once using stylus for data input. This was done in counterbalanced order to further mitigate against learning effects. For example, in each group three participants used speech first, while the remaining three participants used the stylus first. Table 4 shows the structure and sequence of the entire experiment process.



	# Participants		Session 1 Input Style 1 		Session 2 Input Style 2 		
<u>Group A</u> 70 dB – 80 dB	6	Background Data Collection	Train & Test	TLX Workload Test	Train & Test	TLX Workload Test	Preference Ratings
<u>Group B</u> 80 dB – 90 dB	6		Train & Test		Train & Test		
<u>Group C</u> 90 dB – 100 dB	6		Train & Test		Train & Test		
			<i>Input styles delivered in counterbalanced order per group</i>				

Table 4 – *Experimental Procedure, showing the three noise groups and the counterbalancing of conditions during the two sessions. [15]*

Prior to beginning the study, participants was asked to provide some demographic information concerning their age, gender, and experience with the technologies involved.

As already mentioned, prior to each session, each participant was trained in the use of the input technique that he/she was about to use. This training was conducted under similar conditions to the actual experimental sessions. After each session, the participants were required to fill out the NASA Task-Load Index (TLX) questionnaire, and once they had completed the entire experiment, participants were asked to give their overall preference for input technique as well as their preferences for data entry techniques according to individual data types.

Further information regarding the experimental process and the experimental instructions are found in Appendices A and B. The following pictures (Figures 8(a) and

(b)) show a participant interacting with the mobile device during the experimental procedure in our lab space. Notice each table in appropriately labeled and data entry instructions are placed in a binder on the top of each table.



Figure 8 (a) – *Example of Experiment in Progress*



Figure 8 (b) – *Example of Experiment in Progress*

4.3-3 Experimental Hypotheses

Before conducting the experiment we identified a number of hypotheses. As mentioned previously we hypothesized that, due to the high noise levels, the entry precision rate returned would be higher for stylus and keyboard compared to speech-based entry input.

As a result of the small screen size of the mobile device, which means the keyboard buttons are displayed at a small size, we hypothesized that participants would be slower when using stylus compared to completing their tasks using speech.

We also hypothesized that by using speech-based input, users would be more aware of their surroundings due to the fact that they are not required to view the screen constantly for data entry (as is the case when using stylus and keyboard).

With the combination of small screen size and the mobility of the participants during the experiment, we hypothesized that the workload levels for stylus-based input would be higher than that of speech-based input.

5 Results and Discussions

To analyze the results of our experiment we typically used a combination of a two-factor ANOVA test and post hoc “Honestly Significantly Different” tests (or Tukey HSD tests) to identify those results that are statistically significant.

The two-factor ANOVA analysis of variance consists of three significance tests: a test of each of the two main effects and a test of the interaction of the two main effects. The analysis of variance is the most widely-used method of statistical inference for the analysis of experimental data and is a method for testing hypotheses about means [17]. The Tukey HSD tests test all pairwise comparisons among means for all the sources of variation such as Noise Levels, Modalities and the interaction between Noise Level and Modality [18].

The following sections are devoted to showing all the statistically significant results that we found during our analysis of the experimental results.

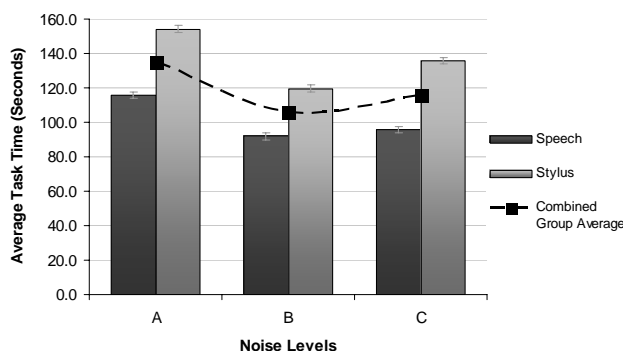
5.1 Participant Background Data

Prior to beginning of the study, each participant was asked to provide us with some background information. The experiment included 18 participants – 10 males and 8 females, ranging in age from 18 to 45 years. The participants had a wide range of backgrounds, with one third from civil engineering and the remainder coming from nursing, business administration, computer science, forestry, project managers and sales clerks. Eleven participants had previously used a handheld device, but most of whom had only using this technology on a small number of occasions. Only two participants had used a speech-based interface before; again, this use was limited to a small number of occasions.

5.2 Task Completion Time

During the experiment we kept a record of each participant's task completion times for speech-based and for stylus-based data entry. A two factor ANOVA showed that both noise level ($F_{2,30} = 10.42$, $p < 0.0001$) and modality ($F_{1,30} = 44.74$, $p < 0.0001$) had a significant effect on participants' average task completion time. We identified no significant affect of the combination of noise level and modality.

Tukey HSD tests showed us that there is a significant difference between Group A and Group B ($p=0.0003$) and between Group A and Group C ($p=0.0156$) where Group A, B and C refer to the Noise Levels 70-80dB, 80-90dB and 90-100dB respectively.



Noise Level	Speech	Stylus	
A (70-80 dB)	115.7	154.2	134.9
B (80-90 dB)	92.0	119.7	105.8
C (90-100 dB)	95.7	135.7	115.7
	101.1	136.5	

Table 5 – Avg. Task Time in seconds

Figure 9 – Average Task Time according to Noise Levels for both modalities

Across both input techniques, participants in Group A (70-80dB) took an average of 134.9 seconds to complete their data entry tasks which was significantly longer than participants in Group B (80-90dB) who had an average of 105.8 seconds and was also significantly longer than Group C (90-100dB) who took 115.7 seconds to complete the MFDE application.

Figure 9 demonstrates the difference in task completion times across the noise levels. We can see that Group A is significantly different to the other two noise levels. We were surprised that there was no significant difference between Groups B and C as we hypothesized that the higher noise levels would have a detrimental effect on participants

using speech which in turn would impact (increasing) the overall average task time with each increase in noise.

The high task completion time for the lowest noise level was also surprising but we anticipate that perhaps participants did not feel as pressured because of feeling more comfortable and relaxed (i.e. less stressed) at this lower noise level which resulted in them taking their time to complete the task. Further investigation will be required to conclusively determine why this surprising result occurred.

Tukey HSD tests also showed that, across all participants, there is a significant difference ($p < 0.0001$) between the average task completion times when using speech-based (101.1 seconds) input verses using the stylus-based input (136.5 seconds). Figure 9 shows that across all Noise Levels, the task completion times returned by participants when using stylus is higher than their task completion time when using speech. This is only significant for Group A ($p = 0.0028$) and Group C ($p = 0.0018$). This seems to follow our hypotheses that task times for speech-based input is faster than stylus-based input.

5.3 Average Percentage of Distractions Identified

No significant differences were found when comparing the percentage of distractions identified by participants according to either noise level or input modality. This was surprising since we hypothesized that due to the nature of using speech-based input, users are would be more aware of their surroundings when using speech than when using stylus-based input.

Having observed the participants during the experiment, we noticed that participants where inclined to be more observant during their second session. They appeared to have settled into the experimental process and remembered that they had to watch out for the “danger” photographs. As a result we decided to isolate the second session’s data.

A two factor ANOVA showed that input modality significantly affected participants' ability to recognize distractions ($F_{1,12} = 6.27, p=0.028$). According to this isolated study, the noise levels had no affect on participants' ability to recognize distractions in their environment.

Noise Level	Speech	Stylus	
A (70-80 dB)	120%	90%	100%
B (80-90 dB)	130%	90%	110%
C (90-100 dB)	110%	90%	100%
	120%	90%	

Table 6 – Avg. % Distraction (Isolating Session 2 only)

Table 6 shows that participants using stylus-based input reported seeing 90% of the “danger” photographs projected compared to 120% reported by participants using speech-based input.

Participants using speech-based input report a percentage rate higher than 100% which essentially means that participants reported seeing more pictures than were actually displayed. Participants commented that, when using speech-based input, they were able to look up from the device more often than when using stylus-based input. Also speech-based input allowed participants to complete their tasks faster. This higher percentage rate for distraction awareness can be therefore perhaps be attributed to the fact that participants recognized a danger picture, then quickly finished their task, only to look up one more time and count the same picture they had recognized before as a new distraction.

In this controlled experiment, users had not been advised that “danger” images would not appear sequentially on the same projector. Although this might be considered a flaw in the experimental design, the finding would at least suggest that speech-based interaction did permit a greater awareness of the physical environment than the stylus-based interaction. Furthermore, the 10% of distractions that the participants using stylus-based input missed could potentially represent the hazard that proves to be fatal or near-fatal (i.e. in this environment, it is better to be over cognizant than to miss something).

Based on the comments made by participants, we noticed that even participants who recorded an incorrect number of distractions still thought that speech enhanced their ability to recognize distractions and their environment. We suspect that further investigations of this distraction element will support our hypothesis that speech-based input allows users to be more cognitive of their surroundings.

5.4 Average Entry Precision Rates

The entry precision rates achieved by participants using the two different input techniques was calculated by analyzing each of the seven data entry fields and assigning a score of 1 if the data entered was entirely correct and a score of zero if it was in any way incorrect. The ratio of data fields that were completed correctly relative to the total number of fields was then used as a measure of average entry precision rates.

Again using a two factor ANOVA, comparing the average entry precision rate according to noise level and modality, we found significant differences between the various noise levels ($F_{2,30} = 6.06$, $p = 0.006$) and between the two modalities ($F_{1,30} = 86.27$, $p < 0.0001$). Furthermore the combination of noise level and modality was also found to significantly affect the precision rates ($F_{2,30} = 8.53$, $p = 0.001$).

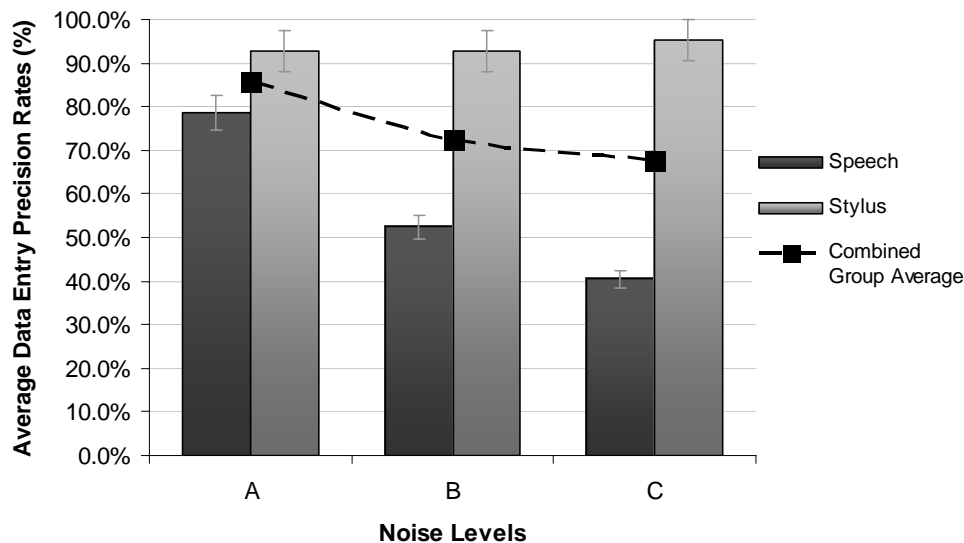


Figure 10 – Average Entry Precision Rates

There is a significant difference in the average precision rates when using speech compared to stylus ($p < 0.0001$) across all participants. The precision rates for speech were 57.1% compared to the stylus precision rate of 93.7%. We can, therefore, conclude that using speech returns a lower accuracy rate than using stylus.

Noise Level	Speech	Stylus	
A (70-80 dB)	78.6%	92.9%	85.7%
B (80-90 dB)	52.4%	92.9%	72.6%
C (90-100 dB)	40.5%	95.2%	67.9%
	57.1%	93.7%	

Table 7 – Average Entry Precision Rates

Tukey HSD tests revealed that, when comparing the modalities within each noise level, there was no significant difference in Group A. There was, however, a significant difference in the average precision rates between speech and stylus in Groups B and C ($p < 0.0001$ in both cases). Looking in more detail (see Table 7) we can see that within Group B, participants achieved an average precision rate of 52.4% for speech and 92.9% for stylus where as for Group C participants achieved 40.5% for speech and 95.2% for stylus.

As can be seen in Figure 10 and Table 7, there is a significant difference ($p=0.0454$) between the average data entry precision rate achieved by participants in Group A (85.7%) and that achieved by participants within Group B (72.6%). Tukey HSD tests also show that there is a significant difference ($p=0.0061$) in terms of data entry precision achieved by Group A and Group C (67.9%). Although the average entry precision rate achieved by Group C is lower than Group B, it is not statistically significant. These results suggest that data entry precision rates decrease with increasing noise levels when speech input is included in the equation.

Tukey HSD tests also confirmed that the average data entry precision rates achieved by Group A participants for speech-based input (78.6%) compared to the rates achieved by Group B participants (54.4%) was significant ($p=0.0082$). There was also a significant difference ($p=0.0002$) between Group A (78.6%) and Group C (40.5%) when using speech. This tells us that speech-based data input precision achieved in the lowest noise level is significantly better than the precision achieved at higher noise levels. There was no significant difference between the average precision rates in Groups B and C when using speech.

There was no significant difference between the stylus-based data entry precision rates achieved across the different noise levels. The stylus-based precision rates achieved by Group A were significantly better than the speech-based results for Groups B ($p<0.0001$) and C ($p<0.0001$). Similarly the stylus-based precision rates achieved by Group B were significantly better than the speech-based results for Group C ($p<0.0001$).

As we can see from Figure 10, the average precision rates when using speech are below the combined group average, for all three groups, with stylus performing above average. This supports our hypothesis that the presence of noise would reduce the accuracy rates achieved for data input using speech. Stylus-based input was generally more accurate than speech based input and, with increasing noise levels, stylus-based input increasingly exceeded that of speech-based input.

Precision rates achieved under the lowest noise level were significantly better than the two higher noise levels for speech-based input. This is a very interesting observation and as discussed in the next sections, even though the precision rates are low with speech, most users commented on preferring using speech over stylus.

5.5 Data Type Precision Rates

Now that we have examined the overall data entry precision rates, consider how each input technique affected the accuracy rates achievable for individual data entry fields. Using the two factor ANOVA analysis of variance and Tukey HSD tests, to analyze the data entry precision rates for each data type (numbers, dates, decimals, drop down lists) versus the three noise level groups (70-80 dB, 80-90 dB, 90-100 dB) and two modalities (Speech and Stylus) we found the following significant results.

Data Type	Speech	Stylus	
Drop-Down Lists	81.5%	98.1%	89.8%
Decimals	47.2%	97.2%	72.2%
Dates	33.3%	94.4%	63.9%
Numbers	27.8%	72.2%	50.0%
	47.5%	90.5%	

Table 8 – *Average Data Precision Rates*

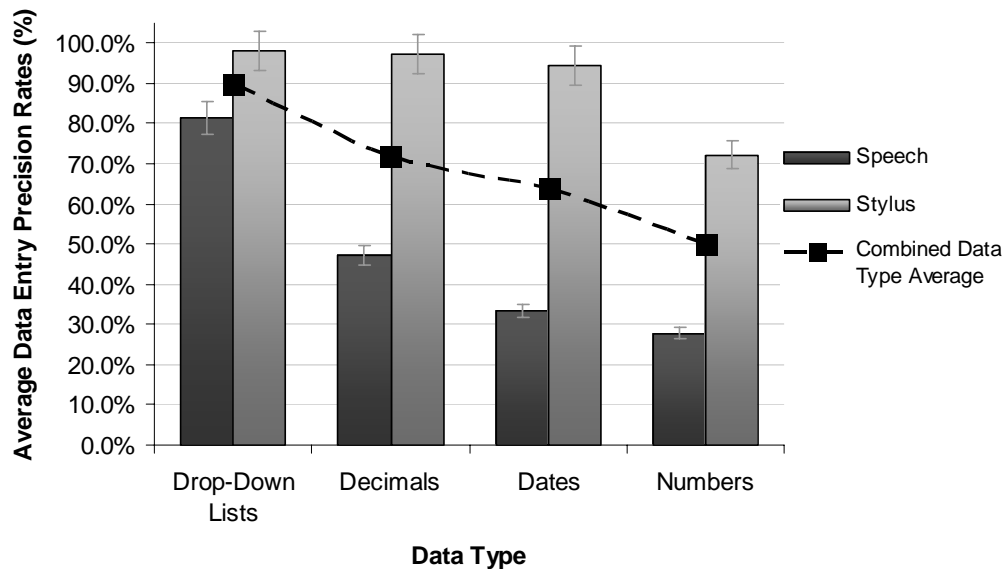


Figure 11 – *Average Data Precision Rates*

A two factor ANOVA revealed that input technique (modality) significantly affected participants' precision rates for all data types: for numbers ($F_{1,30} = 8.42$, $p = 0.007$), for dates ($F_{1,30} = 22.41$, $p < 0.0001$), for decimals ($F_{1,30} = 24.55$, $p < 0.0001$), and for drop-down lists ($F_{1,30} = 12.25$, $p = 0.001$). As shown in Table 8 and Figure 11, significantly higher data entry precision rates were achieved using stylus-based input for each data type. An important point is that drop-down lists recorded the best data precision rates over all other data types when using speech-based input.

5.5-1 Number and Date – Entry Precision Rates

Although the data entry precision rate among the number fields appears to decline at higher noise levels, there were no significant differences to make any substantial conclusions. As mentioned above, the data entry precision rate for number entry did however decline substantially with the use of Speech (27.8%) compared to using Stylus (72.2%).

As with the number fields, there were no significant differences between noise levels in terms of date entry precision rates but the data entry precision rate did substantially decline with the use of Speech (33.3%) compared to using Stylus (94.4%) as shown in Figure 11. This is not surprising since the date data type is a more complex data entry.

Further investigation will need to be conducted in terms of the data entry precision rate for dates and numbers. A limitation of our experiment was that we only included one date and one number field which could potentially have affected the results; future research should focus on the inclusion of more fields of those data types.

5.5-2 Drop-Down Lists – Entry Precision Rates

As show in Table 9 and Figure 11, the data entry precision rate for drop-down lists was almost perfect when using a stylus (98.3%) while using speech (82.6%) it was less accurate. This is surprising given the complexity of selecting an option in a drop-down list, which requires precision with the stylus and scrolling, whereas with speech the participant only had to say the item that was to be selected. The reason why the speech was less accurate is likely because of the environmental noise which had an impact accuracy of the speech input.

Noise Level	Speech	Stylus	
A (70-80 dB)	94.4%	94.4%	94.4%
B (80-90 dB)	83.3%	100.0%	91.7%
C (90-100 dB)	66.7%	100.0%	83.3%
	82.6%	98.3%	

Table 9 – *Average Drop Down Lists Precision Rate*

Tukey HSD tests also revealed significant differences in data entry precision rates for drop-down lists as a result of the combination of noise levels and modality ($F_{2,30} = 4.23$, $p=0.024$). Although it appears that increasing noise levels lower the precision rate of speech, there is only a significant difference when comparing Group A (94.4%)

against Group C (66.7%) noise levels ($p=0.0184$). Another significant difference found is between using stylus-based input (100%) and speech-based input (66.7%) at the highest noise level of 90-100 dB.

5.5-3 Decimal – Entry Precision Rates

As shown in Table 10 and Figure 11 the data entry precision rates for decimals varied with the stylus being almost perfect (97.2%) compared to speech where the accuracy was significantly less (47.2%).

Noise Level	Speech	Stylus	
A (70-80 dba)	83.3%	91.7%	87.5%
B (80-90 dba)	33.3%	100.0%	66.7%
C (90-100 dba)	25.0%	100.0%	62.5%
	47.2%	97.2%	

Table 10 – Average Decimal Precision Rate

Tukey HSD tests also revealed a significant difference was the found when comparing decimals with combining noise levels and modalities ($F_{2,30} = 4.32$, $p=0.022$). Although it appears that the higher the noise levels the lower the precision rate of Speech, there are only significant differences ($p=0.0250$) between Group C (25.0%) and Group A (83.3%) when using speech.

Surprisingly there are no significant differences when using speech between the Group A and Group B or between Group B and Group C. There was no significant difference between the modalities at the lowest noise level, however Groups B and C demonstrated a significant difference ($p=0.0076$ and $p=0.0022$ respectively) between modalities, which shows the impact of increasing noise levels on participants' ability to achieve accurate input using speech.

There is no difference in precision when entering decimal numbers using stylus and speech at the lowest noise level. With speech there is greater precision accuracy for decimal numbers at the lower noise levels (83.3%) than compared to the two higher levels (33.3% and 25.0% respectively). This tells us that with speech-based input the higher noise levels had a detrimental effect on the accuracy rates for decimals.

5.5-4 Accuracy Rates among Data Types

Data types had a significant impact on the achievable precision rates ($F_{3,132} = 5.95$, $p = 0.001$) with drop-down lists being more accurate than dates ($p = 0.0374$) and numbers ($p = 0.0004$). The accuracy rates when using drop-down lists (89.8%) were high compared to numbers (50.0%) and dates (63.9%), but there was no significant difference when compared to decimals which is surprising, since our initial preconceptions was that participants would achieve a significantly higher accuracy rate between drop-down lists and decimals especially since it seemed that participants had troubles with decimals while using speech during the experiment.

5.5-5 Accuracy Rates across Input-Styles and Data Types

There are also significant differences when comparing the accuracy rates against the two modalities ($F_{1,136} = 53.01$, $p < 0.0001$) as we can see from Table 10 with stylus (90.5%) being significantly more accurate than using speech (47.5%).

There is also a significant difference when comparing the combinations of modalities and data types ($F_{3,136} = 2.73$, $p = 0.046$) as we can see from Figure 11, the accuracy when using speech for dates (33.3%) compared with speech for drop down lists (81.5%).

When using speech for decimals there are significant differences in accuracy compared to dates, decimals and drop-down lists when using stylus ($p = 0.0024$, $p = 0.0010$, $p = 0.0007$ respectively). When using speech for drop-down lists there is a significant

difference ($p=0.0002$) in accuracy that using speech for numbers. There is also a significant difference when comparing using speech for numbers to using stylus for all the four data types.

Drop-down lists and decimals performed better using speech at the lower noise levels (70-80 dB) compared to the highest noise level (90-100dB). All these results support our hypothesis that stylus is more reliable than speech at higher noise levels, with drop-down lists performing the best of all data types when using speech-based input at higher noise levels.

5.6 Questionnaire Results

Data from the TLX workload tests were analyzed and scored using a twenty point scale. The averaged results can be seen in Appendix D. A two-factor ANOVA did not find a significant difference when comparing the overall workload reported according to Noise Level or Modality. This was surprising as we hypothesized that participants would return higher overall workload levels for stylus-based input than speech-based input as with stylus users have to locate small buttons for numbers or items within the drop-down lists, whereas using speech as participant would have to just say the items or numbers.

When we looked at the individual factors contributing to workload we found a significant difference with the performance level achieved when compared among modalities ($F_{1,30} = 5.62, p=0.024$). Participants reported an average rating for speech-based input of 12.3 (out of a possible 20) and for stylus-based input (15.7). This difference suggests that participants felt that they performed better using stylus which is not surprising since the overall average precision rates are better using stylus giving the participant a better feeling of success.

The fact that participants did not report worse levels of effort and frustration for speech is likely because the participants were not permitted to correct erroneous or

omitted entries. If users were asked only to continue once their entries were correct, we foresee participants becoming very frustrated as they might have to re-enter fields multiple times. In this case, the task completion time for speech-based input would become longer. The testing of frustration levels is an issue that can be investigated in future experiments.

5.7 Preferences Results

As Figure 12 below shows, there is no significant difference in the overall preference of using speech or stylus. Overall 8 participants stated overall preference for the stylus-based data entry technique whereas 10 participants stated overall preference for the speech-based data entry technique. The average preference ratings were identical for both types of data input (14.0 out of 20).

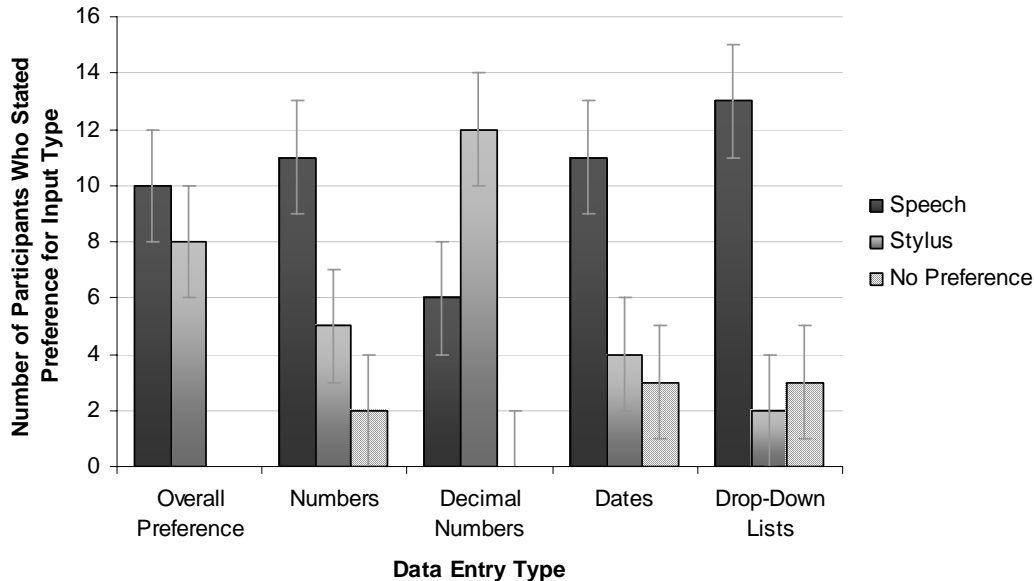


Figure 12 – Subjective preference for input modality according to data entry type.

There are not enough differences to make any significant conclusions on preferences; however as Figure 12 shows, speech was slightly preferred in our limited experiment for all data types with the exception of decimal numbers where stylus-based entry was preferred.

This slight preference for speech contradicts the results from the average performance level achieved which states that users felt more successful using stylus and the results of the data precision rates. Although the slight preference for speech contradicts these results, they do reflect the comments that participants made which included:

“I think this method (speech) of data entry would be useful on a job site.”

“The voice controlled entry was easier to use by far.”

“The speech was preferred because I didn’t have to try to locate small buttons.”

“I like the speech-based entry because you are able to focus more on your surroundings.”

“I felt more distracted while using the keyboard – I couldn’t watch what was going on around me.”

In general participants commented on the lack of accuracy using stylus, but they still preferred speech or said that they would prefer speech once they have become more accustomed to the device and the speech entry technique. Most participants also commented that speech-based input allowed them to be more aware of their surroundings than when using stylus-based input. There were a few comments about speech-based input requiring less effort than stylus-based input, although this was not verified by our results, it would be valuable to conduct future tests to measure effort while using both input techniques.

5.8 Experimental Limitations

This experiment was limited to a small population sample size of 18 participants mainly chosen due to the time involved in completing this evaluation study. Although this number is big enough to give meaningful results, future testing with more participants would be valuable returning a greater meaning and accuracy of results.

One of the biggest limitations to our experiment was the quality of microphones we used during the testing procedure, more specifically the quality of noise reduction built into the microphone. Prior to the experiment, five different microphones had been tested to be sure that we were using the model most capable of functioning well in our experimental environment.

After much testing we resorted to using the IPAQ built-in microphone as it returned the best pre-experimental data precision results. The microphone can have a detrimental effect on our results for this experiment and as a result more extensive testing in the future could be performed to choose a better microphone which potentially leads to more favorable results for speech-based data entry.

6 Conclusions and Future Work

Using the HCI labs at NRC we were able to conduct a meaningful evaluation of the MFDE application to compare two potential data input techniques, speech-based and stylus-based.

Stylus-based data entry was significantly more accurate than speech under our experimental conditions. Participants performed tasks significantly faster when using speech-based input although, as was previously mentioned, if participants were required to fix erroneous fields before moving onto the next field, this could prolong the time and also increase users' workload ratings such as effort and frustration levels. Further

investigations which would include participants having to fix erroneous fields during the experiment could be conducted in the future.

Participants were significantly more aware of their surroundings while using speech-based input although this was only found by isolating session two's data and further studies will need to be conducted to see if this result is more generally applicable.

Stylus-based input performed better than speech-based input for all data types. Precision rates achieved under the lowest noise level were significantly better than higher noise levels; more specifically the average accuracy rates were lower when using speech at the highest noise levels.

The data entry precision rates were significantly higher for drop-down lists for speech compared to the other data types as a result of their restrictive vocabulary. The two-highest noise levels (80-100dB) did affect the speech accuracy rates compared to the 70-80 dB noise level. Although 80dB is below the average noise levels on a construction site, our results suggest that if the interface was redesigned to rely more heavily on drop-down lists it may be possible to increase the accuracy rates of the whole system at higher noise levels making the application more reliable and usable under those extreme conditions. Future investigation will need to be conducted in this area to determine whether this is indeed possible.

While aware of the inaccuracy using speech, participants still seemed to favour this technique slightly. Since no significant conclusions were found regarding participants' preferences, further investigations could include an extra session (a third session) in the experiment to allow a participant to choose the input technique of their preference for each data type in that session. This would help us obtain a good measure of participants' preferences for data entry.

Further work could be done to enable us to record all data in real-time, with time stamps for each field, to obtain an accurate measure of task completion time differences

among modalities and data types. Users commented on how they felt secure when using the stylus because they saw the results in real-time and knew they were correct. Feedback mechanisms for speech will need to be investigated to allow users to check and change erroneous field entries.

As mentioned previously microphones can be a limitation and have an effect on the average precision rates if they do not employ a good quality noise reduction feature. One possibility for future research in this area is to investigate the use of a bone-transfer microphone as a means to reduce the impact of external noise. A further investigation is needed to identify the particular types of microphones that could benefit working in noisy environments such as construction sites.

This experiment was helpful for our future development of the MFDE application as well as beneficial to mobile and multimodal interface developers in general. We made a significant contribution to science by showing the possibilities of reflecting key environmental elements inside of a laboratory for usability studies. We essentially brought a construction site into a laboratory space where we could conduct a meaningful evaluation of two input techniques for use with a mobile data collection application.

We successfully established that speech is not an effective input technique at higher noise levels above 80 dB, but with future advancements of speech recognition one day speech-based input could be a viable and effective option that is preferred over the traditional stylus based input when used in high noise levels such as a construction environment.

Bibliography

1. Weinschenk, S., and Barker, D., (2000) "*Designing Effective Speech Interfaces*," Wiley, New York
2. Burkhardt J., Henn. H., Hepper S., Rindtorff K. and Schack T. (2002). *Pervasive Computing. "Technology and architecture of mobile Internet applications"*, Addison-Wesley, New York.
3. IDC Viewpoint. (2002). "*Five segments will lead software out of the complexity crisis*", by A.C. Picardi, December 2002, Doc #VWP000148.
4. Kondratova, I. (2004). "*Voice and Multimodal Technology for the Mobile Worker*", *Electronic Journal of Information Technology in Construction*, Vol. 9, Special Issue Mobile Computing in Construction. September 2004. p. 345-353
5. IBM Pervasive Computing, (July 2004) "*Multimodal Tools V4.1.2 / 4.1.2.2, Frequently Asked Questions*," [Online]. Available at ftp://ftp.software.ibm.com/software/pervasive/info/multimodal/multimodal_faq.pdf [Accessed on March 2006]
6. Sajor, M., Planet PDA Magazine, (July 10, 2002) "*Multimodality Will Drive Wireless Application Adoption*," [Online]. Available at <http://www.planetpdamag.com/content/071002ki.htm> [Accessed on March 2006]
7. Kondratova, I. (2004). "*Mobile Field Data Entry for Concrete Quality Control Information*", in *Proceedings of the European Conference on Products and Processes Modeling (ECPPM 2004)*. Istanbul, Turkey. September 8-11, 2004.
8. IBM Pervasive Computing, (February 2004) "*XHTML + Voice Programmer's Guide*," [Online]. Available at ftp://ftp.software.ibm.com/software/pervasive/info/multimodal/XHTML_voice_programmers_guide.pdf [Accessed on March 2006]
9. IBM Pervasive Computing, (February 2003) "*X+V and Speech Considerations*," [Online]. Available at ftp://ftp.software.ibm.com/software/pervasive/info/XV_and_Speech_Considerations.pdf [Accessed on March 2006]
10. Kjeldskov, J., Skov, M.B., Als, B.S., and Høegh, R.T., "*Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field*", in *Proceedings of the 6th International Symposium on Mobile Human-Computer Interaction (MobileHCI'04)*, Brewster, S., and Dunlop, M., (Eds), Glasgow, Scotland, Sept. 13-16, 61-73.
11. Gilchrist, A., Allouche, E.N., and Cowan, C., (2003). "*Prediction and Mitigation of Construction Noise in the Urban Environment*", in the *Canadian Journal of Civil Engineering*, 30, 659-672.
12. Eaton, S. (2000). "*Construction noise*," Engineering section report. Workers Compensation Board of B.C., Vancouver, B.C.

-
13. Miyara, F., “*Sound Levels*,” [Online]. Available at <http://www.wiw.fcia.unr.edu.ar/~acustica/comite/soundlev.htm> [Accessed on January 2006]
 14. Canadian Centre for Occupational Health & Safety. “*Hearing Protectors*,” [Online]. Available at http://www.ccohs.ca/oshanswers/prevention/ppe/ear_prot.html [Accessed on March 2006]
 15. Kondratova, I., Lumsden, J. and Langton N. (2006). “*Multimodal Field Data Entry: Performance and Usability Issues*” to appear in Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, Canada, June 14-16
 16. Hart, S.G., and Wickens, C., (1990). “*Workload Assessment and Prediction*” in MANPRINT: An Approach to Systems Integration, Booher, H.R., (Ed), Van Nostrand Reinhold: New York, 257-296.
 17. Rice Virtual Lab in Statistics “*2-factor ANOVA*,” [Online]. Available at http://www.ruf.rice.edu/~lane/stat_sim/two_way/index.html [Accessed on March 2006]
 18. Land, D., HyperStat Online Textbook, “*Tukey’s HSD procedure*,” [Online]. Available at <http://davidmlane.com/hyperstat/B95118.html> [Accessed on March 2006]

Appendix A: REB Ethics Application

According to NRC and UNB procedures for experiments containing human participants, an Ethics Application was submitted for approval before starting the experiment phase. *Appendix A* contains the Ethics Application submitted to the Research Ethics Board. This application included as an appendix the experimental instructions to be given to participants. These experimental instructions are shown in *Appendix B*.

Below is a small selection of my approval email sent by the Research Ethics Board:

Dear Mr. Langton:

As Chair of the Research Ethics Board (REB), I have reviewed your application (*A Usability Study of Multimodal Field Data Entry Techniques for the Construction Industry*--REB #2006-013) for its compliance with Tri-Council Policy (TCP) and with UNB Policy (UNBP). On the basis of the review, I consider your project to be eligible for expedited review since any risk to participants that might exist appears not to exceed the "minimal risk" outlined in TCP.

...

Best wishes for the successful completion of your research project.
Peter Kepros, Ph.D.
Professor and Dean Emeritus and
Chair UNB Research Ethics Board

FOR REB USE ONLY: File No.: _____
 Date Complete App'n. Rec'd.: _____
 Approved/Approved with.: _____
 Modification/Rejected.: _____
 Date.: _____

.....

**University of New Brunswick
 Application for Review of Research Involving Humans**

Principal Investigator(s): Name(s); Academic Status (Faculty, Undergraduate Student or Graduate Student); Academic Unit, e-mail Address, Office Telephone, Home Telephone:

**Nathan S. Langton, Faculty of Computer Science Undergraduate Honors Thesis Student
 UNB Faculty of Computer Science
 e-Mail: nathan.langton@unb.ca / Office Tel: (506) 444 0486 / Home Tel: (506) 474 0150**

Title of Proposed Research:

“A Usability Study of Multimodal Field Data Entry Techniques for the Construction Industry”

Commencement Date: **Sept. 2005** Completion Date: **Mar. 2006**

Co-Investigator(s): Academic Unit, e-mail address, Office Telephone

N/A

Supervisor(s) (if Principal Investigator is a student); Academic Unit, e-mail address, Office Tel:

**Dr. Irina Kondratova, UNB Dept. of Civil Engineering (Adjunct Professor)
 e-Mail: irina.kondratova@nrc.gc.ca / Tel: (506) 444 0489**

**Dr. Joanna Lumsden, UNB Faculty of Computer Science (Adjunct Professor)
 e-Mail: jo.lumsden@nrc.gc.ca / Tel: (506) 444 0382**

**Prof. Natalie Webber, UNB Faculty of Computer Science
 e-Mail: nwebber@gmail.com / Tel: (506) 452 6328**

The undersigned parties certify that they have read, and undertake to comply fully with, the Tri-Council Policy Statement “Ethical Conduct for Research Involving Humans.”

Principal Investigator(s): _____

Co-Investigator(s): _____

Supervisor(s): _____

The undersigned certifies that the proposed research has been reviewed by, and is acceptable in all respects to, the academic unit(s) responsible.

Dean/Director/Chair(s) Typed Name Dr. Virendra C. Bhavsar

Signature _____

Date submitted to the REB: January 30, 2006

1. Summary: Provide here, in approximately 300 words, a summary of the proposed research, indicating clearly the role of the research subjects and any procedures to which they will be subjected.

The Multimodal Field Data Entry (MFDE) System is a prototype mobile application designed (by NRC IIT) to allow concrete technicians (or engineers) to have a quick and convenient means by which to record concrete test data when conducting tests on concrete on a building/construction site. Using a mobile device such as a handheld PC or a tablet PC, an engineer uses speech or a stylus to enter test data into a simple electronic form. The research proposed here aims to conduct a lab-based usability study of this prototype application in order to (a) determine and compare the effectiveness and usability of the two different input options (speech input v. stylus-based input) and (b) determine which of the two input options is preferred by users.

Unlike standard desktop PC-based software, software applications – such as the MFDE prototype – which are designed for mobile technologies are intended to be used when in motion. In order, therefore, for usability studies of such technologies to return meaningful results, these studies (a) must reflect the physical environment in which the technology is to be used and (b) require that the technology be used by users who are *mobile* within that environment (or a representation of the environment). To both control the study and ensure the safety of study participants, usability studies of mobile technologies are most often performed in labs where the relevant attributes of the physical environment are replicated/represented in a manner that can be controlled by the evaluator. Such a set-up – utilizing the purpose-built Mobile HCI Lab in the NRC building on campus – will be used for this study.

Relevant to the MFDE prototype, we have identified two key environmental factors which we will *represent* in the lab: (a) the environmental noise levels inherent on a construction site; and (b) the presence of heavy equipment. The noise levels on a typical construction site will potentially influence the effectiveness of the speech-based input option on the MFDE system, since speech recognition engines often struggle to recognize speech input if background noise is sufficiently intrusive. To investigate this in our study, we will use a surround sound system to deliver construction site sounds within the lab space whilst the participants are interacting with the MFDE application (please see attached documentation for the precise details of this aspect of the study set-up and the safety precautions in place to ensure no more than minimal risk to participants). With respect to the presence of heavy equipment, the key factor of relevance to this study is that using the MFDE system does not impede a user's ability to be cognizant/aware of heavy equipment in their immediate vicinity. To incorporate this into our study without putting participants in positions of danger we will project, at random times, images of heavy equipment around the walls of the lab; whilst using the MFDE prototype, the participants will be asked to keep a mental tally of the number of images of this nature they believe they saw and relay this to us when they are finished.

Finally, we require that the users are mobile – in a manner representative of the actual intended use on a construction site – when using the MFDE prototype during the study.

(contd. over)

To this end, we will ask each participant to complete a series of data entry tasks using each of the input options in turn; the participants will have to move between points in the lab to obtain the data they are required to input (see attached documentation). As indicated previously, the participants will be surrounded by representative environmental noise as well as images of heavy equipment while completing the data entry tasks. We will record the length of time the participants take to complete their data entry tasks, the details of the data they entered, and the number of heavy equipment images they believed were projected. Additionally, we will ask each participant to indicate their experience of workload when completing their data entry tasks as well as which of the two input options they preferred. Prior to beginning the practical aspects of the study, each participant will be asked to provide some demographic information concerning their age, gender, and experience with the technologies involved such that we can look for influence of these relevant factors on the results obtained. Participation should take approximately 1 hour in total, 30 minutes of which will involve exposure to the images and noise.

Attached to this application is a copy of all information that will be distributed to the participants during the course of each study session.

2. Risk: In your opinion, does this research pose more than minimal risk (Tri-Council Policy, Section 1.C1) to participating subjects? *Yes* *No* ✓

If yes, provide here a statement that describes in detail the aspects of the research procedure that pose a risk to subjects, and provide your assessment of the risk of harm (probability and severity). Note that not only physical injury but also anxiety or embarrassments are included in the concept of harm. Describe means adopted to minimize risk, and means (such as provision of counseling) to deal with harms, which subjects may experience. Describe as well the potential benefit, which will result from this research, which justifies the above risk of harm.

Neither the completion of the data entry tasks nor the measure of awareness of heavy equipment (see section 1 and additional attached documentation) pose more than minimal risk to participants. We also do not consider the manner in which we are introducing construction site level environmental noise to the study to pose more than minimal risk to participants. For the purpose of clarity, we explain this assessment below.

As noted in the attached documentation, we intend to deliver environmental noise in the range between 70 dB and 100 dB to reflect the typical range of construction site noise [1]. To provide a frame of reference, a whisper is about 30 dB, conversational speech is about 60 dB, a noisy restaurant or highway traffic or a busy office is about 70 dB, an alarm clock or hairdryer is about 80 dB, and city traffic is about 90 dB [2]. The Canadian Centre for Occupational Health and Safety (CCOHS) recommends that people wear hearing protection when exposed to noise levels at or above 85 dB [3]. We will provide, and insist that our participants wear, hearing protectors (model: Intruder EM7202 Head Band Earmuff) irrespective of whether they are exposed to noise levels above or below this threshold. At the lowest end of our noise range, wearing these ear protectors will reduce the level of noise heard by participants to 49 dB and at the highest end to 79 dB.

(contd. over)

As can be seen, this essentially means that at no point will participants be exposed to noise beyond the 85 dB threshold – indeed, they will not be exposed to noise levels above those which they would typically encounter in everyday life.

New Brunswick regulations state that a person should not be exposed to 85 dB sound levels for more than 8 hours in any one 24 hour period. Added to the fact that we are lowering the level of actual exposure experienced by participants to below this threshold level (via the use of hearing protection) we will not be exposing any one participant to any of the construction noise for more than 30 minutes in total during the course of an experimental session; we therefore do not consider our participants to be at more than minimal risk when participating in this experiment.

Furthermore, we are seeking informed consent from participants prior to taking part in this study; as can be seen from the attached consent form, we explain the issues of the noise levels and the measures in place to ensure exposure rates are acceptable. Participants will be free to decide not to take part at this stage. Additionally, participants will be free to withdraw from the study at any point without penalty and so, once started, should a participant become uncomfortable as a result of the noise, he/she can immediately withdraw and the associated exposure will cease.

We feel we have been duly diligent in both investigating the issue of environmental noise and have taken more than adequate precautions to ensure that we are not exposing participants to more risk than they would encounter in their everyday lives. If we are unable to utilize the noise factor in this study, our results will lack relevance and meaning in relation to the environment in which the software application has been designed to be used. By designing the study as we have done, we feel that we are enabling realistic testing of the technologies involved without subjecting participants to more than minimal risk.

[1] Gilchrist, A., Allouche, E.N., & Cowan, C., (2003), *Prediction and Mitigation of Construction Noise in an Urban Environment*, in Canadian Journal of Civil Engineering, **30**, pg. 659 – 672

[2] <http://www.wiw.fcia.unr.edu.ar/~acustica/comite/soundlev.htm>

[3] http://www.ccohs.ca/oshanswers/prevention/pppe/ear_prot.html

3. Deception: Does this research involve deception or partial disclosure? *Yes* *No* ✓

If yes, refer to the Tri-Council Policy, Section 2, specifically Article 2.1(c) and subsequent commentary, and provide here an explanation of how you plan to comply with the requirements of that Section for debriefing. Describe as well the potential benefit, which will result from this research, which justifies waiving the normal requirements for full disclosure.

4. Funding: Has funding been received for this research? *Yes* ✓ *No*

If yes, *from* what agency and for what period?

Participant honoraria is being provided from Dr. Irina Kondratova's NSERC Discovery Grant.

If no, has funding been requested for this research? *Yes* *No*

If yes, *from* what agency and for what period?

5. Research Subjects:

5.1 Number of Subjects: How many subjects will participate in this research? **18**

5.2 Recruitment: How will they be recruited, and from what population?

Study participants will be recruited from the student, faculty, and staff population on the UNB Fredericton Campus. Calls for Participant (see attached documentation) will be posted on notice boards in appropriate faculty/departments (e.g., Faculty of Computer Science and Faculty of Engineering). Additionally, the Call for Participation will be included in faculty-wide and campus-wide email circulations (e.g. the e-Daily). Respondents will be scheduled on a first-come-first-served basis and will receive a \$10 honorarium for their participation.

6. Informed Consent:

6.1 Informing Subjects: How will the nature of the research be explained to potential subjects, in compliance with Section 2D of the Tri-Council Policy? Attach a copy of any document(s), such as an explanatory letter, to be used for this purpose.

Please see the Call For Participation and the Consent Form documentation (attached). Additionally, please also see the experimental documentation (attached) that will be used to explain to participants what they are required to do at each stage of the experiment.

6.2 Consent: If written evidence of informed consent will be obtained, attach a copy of the consent form. (See Requirements for Informed Consent Forms.) **If written evidence of informed consent will not be used, explain here, in detail, how you intend to comply with the requirements of Section 2A of the Tri-Council Policy:** see particularly Article 2.1(b) and subsequent commentary.

Please see the attached consent form which will be used to obtain written consent from each participant. Please also note that at every stage of the experimental procedure, participants will be reminded that their participation remains optional and that they can withdraw from the study without penalty.

6.3 Children as Research Subjects: If the proposed research involves children as subjects, provide here a statement indicating how compliance with Section 2E, and specifically with Articles 2.5, 2.6 and 2.7 of the Tri-Council Policy, will be achieved.

Not Applicable.

6.4 Incompetent Adults as Research Subjects: If the research involves adults of diminished competence as subjects, provide a statement indicating how compliance with Section 2E, and specifically with Articles 2.5, 2.6 and 2.7 of the Tri-Council Policy, will be achieved.

Not Applicable.

7. Inducements: Will any inducements (money, grade points, etc.) be offered to encourage participation? *Yes* ✓ *No*

If yes, indicate here how compliance with Section 2B of the Tri-Council Policy (concerning voluntariness) will be achieved. If academic rewards are to be used, give details of alternative means of achieving equivalent rewards.

All participants will be informed prior to the study that their participation is voluntary. They will similarly be informed both prior to and during the study that they are free to withdraw from the study at any time without penalty – that is, they will receive their \$10 honorarium irrespective of whether they withdraw once started or whether they complete the study. A Call for Participation (see attached documentation) will be issued (see Section 5.2) and participants will be scheduled on a voluntary basis, each receiving \$10 if they consent to take part in the study (irrespective of whether they complete it).

8. Private Information: Does the proposed research involve accessing identifiable personal information about subjects by means of surveys, questionnaires, etc.? *Yes* *No* ✓

If yes, indicate here, in detail, how you propose to meet the requirements of the Tri-Council Policy, Section 3, specifically Article 3.2. A copy of any questionnaire, survey document or interview schedule to be used should be attached as well.

Note: although “questionnaires” are used to obtain some technology-related background information about participants and to assess participant reaction to certain aspects of the prototype and their interaction with it, all data is anonymous and no personal identifiable information is being collected in this study; no individual will be identifiable from the data they contribute to the study.

9. Feedback: Describe the measures, which you propose for providing feedback to research subjects concerning the outcome of the research.

On the consent form (see attached documentation) we ask participants to indicate whether they would like to receive information about the results of this study; if so, they are asked to provide contact details that will allow us to convey this information to them. We will provide an executive summary of the results to all participants who indicate interest in this way.

10. Data Security: Describe the measures, which you propose for ensuring the security of any identifiable personal data, which will be retained after completion of the research.

No identifiable personal data will be obtained therefore none will be retained.

11. Continuing Review: All research requires brief annual reports and a brief report upon completion of the research. Suitable report forms are included at the end of this file. **Research involving more than minimal risk may require additional measures for continuing review.** If your research involves more than minimal risk, describe here the measures you propose for facilitating continuing review of this research, in compliance with Article 1.13 of the Tri-Council Policy.



Not applicable (not more than minimal risk).


12. Additional Information: Please feel free to append any additional information, which you feel may be helpful to the REB in evaluating this application.

Please find attached to this application, the following information:

- the recruitment Call for Participation;
- experimental handouts for participants including a short questionnaire gathering anonymous background information used to clarify the participants experience with mobile devices and speech applications etc;
- the consent form.

The following table shows the structure and sequence of the experimental conditions:

	# Participants		Input Style 1 		Input Style 2 		
Group A 70 dB – 80 dB	6	Background Data Collection	Train & Test	TLX Workload Test	Train & Test	TLX Workload Test	Preference Ratings
Group B 80 dB – 90 dB	6		Train & Test		Train & Test		
Group C 90 dB – 100 dB	6		Train & Test		Train & Test		
			<i>Input styles delivered in counterbalanced order per group</i>				

Participants will be assigned to one of three groups A, B, or C; allocation will be done according to order of participation in the experiment and so every participant has an equal chance of being allocated to each group. Within each group, the participants will be required to complete, in counterbalanced order, the experimental session for one input style and then the experimental session for the other input style. At each stage, clear written instructions will be provided and the researcher will be on hand to personally answer any queries or deal with any participant concerns. The use of hearing protection will be compulsory when construction noise is being fed into the lab – this is indicated in the above table by a . Noise will not be used at any other stage.

Please note that all explanation will be provided in written form to prevent researcher bias influencing the results of the experiment.

Checklist for Attachments to Application for Review of Research Involving Humans

For items that are attached, indicate X ; for items that are not applicable, indicate N/A .

Provide the following attachments where applicable:

X or N/A	
N/A	Where the academic unit responsible for the research has a process of formal ethics review , a copy of the approval notice from that process, together with any substantive comments provided by the reviewers.
N/A	If external funding has been sought or obtained for this research, one copy only of the complete application form as well as two copies of any reviewers' comments which have been received.
X	For student research , one copy of the full research proposal if one has been submitted to the relevant academic unit.
X	A copy of any proposed information letter and/or informed consent form. (Do not duplicate if already included in above material.)
X	A copy of any questionnaire(s), survey documents or interview schedules to be used in the research. (Do not duplicate if already included in above material.)
N/A	A copy of any debriefing material to be provided to subjects.
N/A	For research under the jurisdiction of more than one institution , an indication of which other REBs will review the research, and the results of such review if available (see Tri-Council Policy Section 1G).
N/A	For all research (including student research) that exceeds minimal risk, which has not been approved by a sanctioned peer review process , the applicant must recommend two reviewers competent to undertake a "scholarly review" of the proposed research. "Scholarly review" in this context refers to the process of determining whether the design of the research project is capable of addressing the questions being asked in the research.
X	In all cases , a full description of the proposed research, if this is not already contained in the material listed above.
X	In all cases , an electronic text version of the Summary (Item 1 on the application form), via e-mail to ethics@unb.ca on the UNBF campus or to flagel@unbsj.ca on the UNBSJ campus.

Please append this checklist to the application form.

CONSENT FORM

for

participation in the research study entitled:

**A USABILITY STUDY OF MULTIMODAL FIELD DATA ENTRY TECHNIQUES
FOR THE CONSTRUCTION INDUSTRY**

(REB #: 2006-013)

Please read this consent form carefully and ask all the questions you might have before deciding to participate or not in this research study. You are free to ask questions before, during, or after your participation in this research.

Dear participant,

The purpose of this consent form is to seek your free and informed consent to participate in a research study entitled “A Usability Study of Multimodal Field Data Entry Techniques for the Construction Industry” (REB #: 2006-013). This research is essentially a usability study of a mobile application that has been designed to allow concrete engineers to enter concrete test results into an electronic form on a mobile device such as a handheld PC or a tablet PC. This study is designed to (a) determine and compare the effectiveness and usability of two different input options (speech input v. stylus-based input) for use with this mobile application and (b) determine which of the two input options is preferred by users. The results of this study will help improve the design of the application as well as provide some valuable information about the two input techniques. The study is being conducted on campus by Nathan Langton as part of his undergraduate honours thesis research with the UNB Faculty of Computer Science.

If you agree to take part in this study, you will be asked to provide some background information regarding your experience with the technologies being evaluated and then complete a series of data entry tasks using both speech-based input and stylus-based input on a mobile device. Essentially, you will be asked to complete a set of data entry tasks using one of the two input options. You will then be asked to repeat this process using the other of the two input options. After each session, you will be asked some questions about your opinion of the input option you have used. After you have completed both sessions, you will be asked to rate your preference for the two input options. Each individual data entry task will require you to walk to a specific table in the lab, read the details of the data, and enter the data using the appropriate input technique into a form on a mobile device. We will record the data you enter as well as the time it takes you to complete the tasks.

Construction sites are busy, noisy places. So that the results from this study are applicable to the environment in which the technology will ultimately be used (i.e. a construction site), you will be asked to complete the above noted data entry tasks in the context of two environmental conditions. Firstly, at the same time as you are entering data, you will be asked to pay attention to projected images around the walls of the lab. These images represent hazards that occur on construction sites and of which a concrete engineer would have to be aware when using the mobile application. You will simply be asked to mentally note how many images of a particular type you think you saw projected around you when you were completing your data entry tasks and tell us that number after you have completed the data entries.

Secondly, while you are entering data, we will be playing construction site noise using a surround sound system in the lab. This noise will be in the range of 70 dB to 100 dB. To provide a frame of reference, a whisper is about 30 dB, conversational speech is about 60 dB, a noisy restaurant or highway traffic or a busy office is about 70 dB, an alarm clock or hairdryer is about 80 dB, and city traffic is about 90 dB. The Canadian Centre for Occupational Health and Safety (CCOHS) recommends that people wear hearing protection when exposed to noise levels at or above 85 dB. We will provide hearing protectors (model: Intruder EM7202 Head Band Earmuff) **which you must agree to wear at all times as instructed by the evaluator in order to take part in this study.** At the lowest end of the noise range, wearing these hearing protectors will reduce the level of noise you hear to 49 dB (less than conversational speech) and at the highest end to 79 dB (equivalent to a

hairdryer and less than city traffic). This means that at no point will *you* actually be exposed to noise beyond the 85 dB threshold – essentially, you will not be exposed to noise levels above those which you would typically encounter in everyday life. Additionally, over the course of the experiment you will not be exposed to noise for more than 30 minutes in total.

The whole study session will take about 1 hour to complete, and you will receive a \$10.00 honorarium in recognition of your participation. Your participation in this study is entirely voluntary and you may end your participation at any time or for any reason (including if you are uncomfortable with the noise) without penalty.

If you have any questions about this study, please contact Nathan Langton or his supervisors Dr. Irina Kondratova, Dr. Joanna Lumsden, or Prof. Natalie Webber as follows:

Nathan Langton
Honours Thesis Student
UNB Faculty of Computer Science
nathan.langton@unb.ca

Dr. Joanna Lumsden
Adjunct Professor
UNB Faculty of Computer Science
Tel: (506) 444 0382
jo.lumsden@nrc.gc.ca

Dr. Irina Kondratova
Adjunct Professor
UNB Dept. of Civil Engineering
Tel: 506 444 0489
e-Mail: irina.kondratova@nrc.gc.ca

Prof. Natalie Webber
UNB Faculty of Computer Science
Tel: (506) 452 6328
e-Mail: nwebber@gmail.com

If you would like to speak to someone about this study who is not directly associated with this research, please contact:

Virendra C. Bhavsar (Dean)
UNB Faculty of Computer Science
bhavsar@unb.ca

Risks

Given the provision and compulsory use of the industry standard hearing protectors, the risks associated with participating in this study are minimal.

Confidentiality and Data Storage

All data collected during this study will be kept confidential. Your responses to the study questions will remain anonymous. No one other than members of the research team will have access to the data collected. Individuals will not be identifiable from the data and will not be identified in any publications related to this research. The data will be stored in a secure fashion and will be destroyed 5 years from the date of this study.

Consent

I, the undersigned, acknowledge that I have been given sufficient time to consider my participation in this study. I confirm that I have received, read, and understood all the information above and give my full and informed consent to participate in the study.

I confirm I have understood the information provided regarding the use of construction site noise and agree to wear the hearing protectors provided at all times when instructed by the evaluator.

I understand that my participation in this study is entirely voluntary and that I am free to end my participation in the study at any time or for any reason without penalty. I also understand that any member of the research team can end my participation in the study for financial, scientific, or ethical reasons at any time.

I understand that by signing this form, I give my full and informed consent to the research team to use the data collected for the purpose of this research and any related research that follows.

Signed at: Fredericton, N.B.

Print Name: _____

Signature: _____

Date: _____

If you would like to receive information about the outcome of this study, please enter details of the most appropriate means by which this information can be communicated to you:

Appendix B: Experiment Instructions

Appendix B contains the form used by the coordinator supervising the experiment, all the training and experimental instructions given to participants throughout the experiment, the TLX workload tests and preference questionnaire.

Experiment Coordinators Form

Date: _____

Time: _____

Experiment Information

1. Participant Number: _____

2. Faculty (background): _____

3. Noise Level :

 70-80 dBA 80-90 dBA 90-100 dBA4. **Session 1:** Speech Stylus5. **Session 2:** Speech Stylus

Session 1	
# of Distractions by (Participant)	
# of Actual Distractions	
Start Time	
Stop Time	

Session 2	
# of Distractions by (Participant)	
# of Actual Distractions	
Start Time	
Stop Time	

Comments:

Some Background Information

To let us assess how relevant elements of people's life experience and/or physical makeup impact on their use of mobile devices and applications, we would ask you to answer the following series of questions. **Please note, this information will remain anonymous; you are not required to complete any question that you do not wish to answer.** If you have any queries, please do not hesitate to ask the researcher.

1. Please indicate your age by ticking the box next to the age-range that applies to you:

15 – 20 years
 21 – 25 years
 26 – 30 years
 31 – 35 years

36 – 40 years
 41 – 45 years
 46 – 50 years
 51 – 55 years

56 – 60 years
 61 – 65 years
 66 + years

2. Please indicate your gender by ticking the appropriate box:

Male Female

3. Please indicate the extent to which you have previously used a handheld mobile device by ticking the box next to the experience level that applies to you:

I have never used a Pocket PC
 I have used a Pocket PC

└ please indicate how often you use such systems _____

4. Please indicate the extent to which you have previously used speech-based input to interact with an application by ticking the box next to the experience level that applies to you:

I have never interacted with an application using speech
 I have used speech-based interaction before

└ please indicate how often you have used speech to interact with an application _____

Introduction

You are being asked to help evaluate a multimodal field data entry (MFDE) application. This is a software application which has been developed to run on a mobile device (in this case a Pocket PC – see Figure 1) to allow concrete engineers on a building/construction site to enter concrete test results. **Please remember that it is the technology that is being evaluated, not you!**



Figure 1 – You will use a Pocket PC (similar in look to this one).

In this study, you will use the Pocket PC to enter a series of data items into a simple form (see Figure 2). The form has been designed to allow a user to enter data using two different input techniques: (a) by speaking the data (speech-based input) and (b) by using the stylus to select letters and numbers from a keypad on the touch screen on the Pocket PC (stylus-based input).

Figure 2 – An example of the type of form you will complete.

To help us compare the usability of these two input techniques, you will be asked to complete the form once using speech and once using a stylus (the order will be determined by the researcher). In each case, before you complete the form, you will be given some training on how to use the relevant input technique and you will be able to try it out so that you are comfortable with the input technique before you fill out the main form.

In each case, after you have completed the main form, you will be asked to complete a series of ratings concerning how difficult you felt the task was to complete. After you have completed the form twice you will be asked to indicate your preference for the two input techniques you have used.

The Experimental Environment

To ensure the usability results you are about to help generate are meaningful, it is very important that we evaluate the usability of the form and the two input techniques under conditions that reflect the environment in which the form will ultimately be used – in this case, when the user is physically mobile on a construction site. To do this we will ask you to move around the lab as you complete the form – the data you will be required to enter into the form will be shown on a series of tables around the lab space and you will need to walk to each of the tables in turn to read the data and then enter it. As you are doing this, we will randomly project photographs onto the walls of the lab and ask that you take mental note of the number of times you see a particular image. This reflects the need for someone on a construction site to be aware of physical hazards around him/her.

Finally, since a building site is a noisy place, we will introduce representative construction noise to the lab space. The implications of the levels at which this noise will be delivered were explained to you in the consent form – if you wish to review this, please ask the researcher. **You will be required to wear hearing protectors (see Figure 3) at all times when the construction noise is being played – refusal to do so will mean that your participation in the study will be terminated by the researcher.**



Figure 3 – The ear protectors which you will be required to wear during the experiment.






You will be provided with detailed instructions at each stage of your participation. If now, or at any time, you have any questions please do not hesitate to ask the researcher.

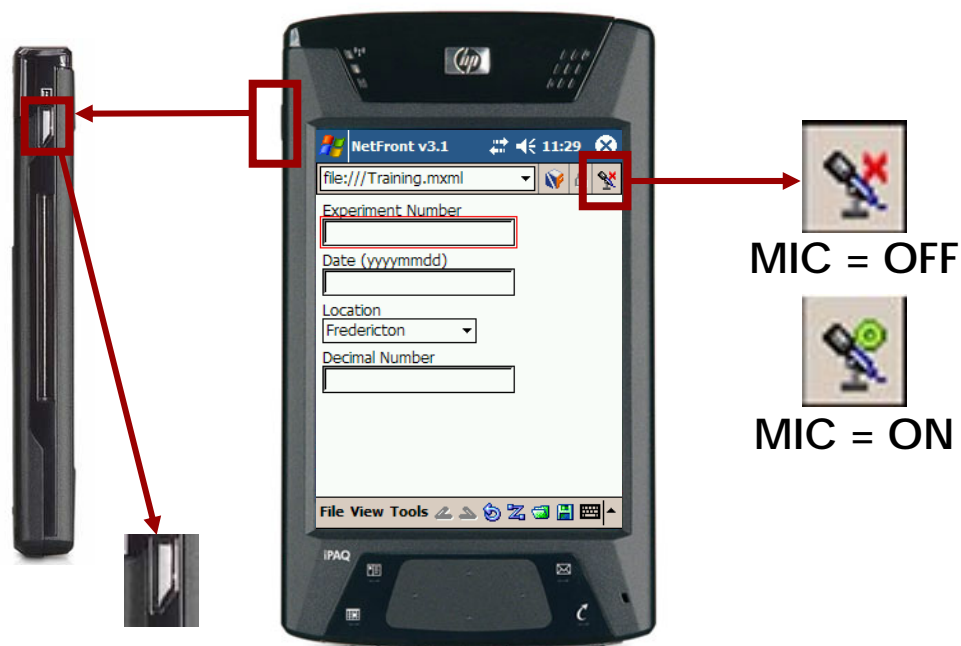
Speech Input :: Training Session

Before you begin to complete the main form, we need to explain how to use the speech input to enter data. To do this, we have developed a sample form which will allow you to test out the speech input without worrying about what your actions might do! Please feel free to ask questions at any point during the training session – it is important that you are comfortable with data entry using the speech input technique before you proceed to using the main form.

There are two components to learning how to use the speech input: first, you need to become familiar with the mechanics of using the microphone (MIC) on the Pocket PC and second, you need to learn how to enter different types of data.

Operating the MIC:

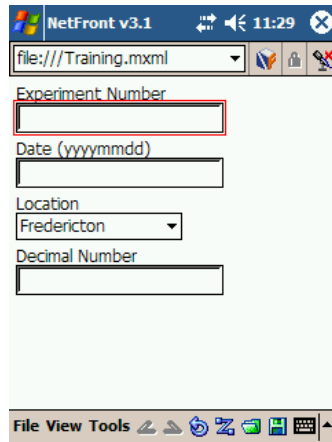
1. The top right button () indicates that the MIC is turned OFF. 
2. When the MIC is turned ON the top right button appears green (). You must PRESS and HOLD Button 1 (see below) for the MIC to be ON (it might take a couple of seconds for this to happen).
3. When the MIC is turned ON – you must remember to PRESS & HOLD Button 1 – you can enter speech input via the microphone on the Pocket PC. ONLY ENTER DATA when the MIC button appears green ().
4. When you have finished entering speech, you simply release Button 1 and the MIC will switch off – the top right button will return to OFF ().
5. Repeat this process each time you want to enter speech to the system.



Button 1:
Press (& hold)
to turn MIC ON

Entering Data:

To help you familiarise yourself with the speech input, we would like you to practice entering some data into a form (see below).




The Pocket PC showing the training form. Notice that the cursor is blinking in the first field (“Experiment Number”).


Since you will be required to walk from table to table to enter the data when you are completing the main form, we would like you to do the same for this training session. Please read how to enter each type of data and then follow the subsequent directions.

To enter data in any given field, the cursor must first be flashing in that field (or in the case of a drop-down list, the field must be highlighted). This is done automatically for you as you move through the list of fields in sequence.


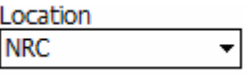
To enter a number – e.g. “685237”:

1. PRESS & HOLD Button 1 to turn MIC ON ().
2. Enter the digits by saying “six eight five two three seven”.
3. Release Button 1 to turn MIC OFF.
4. The field will display 685237 if the speech entry was accepted.



To enter a date – e.g. 10th January 2006:

1. PRESS & HOLD Button 1 to turn MIC ON ().
2. Enter the date either by saying:
 - a. “tenth of January, two thousand and six”,
 - or*
 - b. “January, tenth, two thousand and six”
3. Release Button 1 to turn MIC OFF.
4. The field will display 20060110 if the speech entry was accepted.

To select something from a drop-down list – e.g. to select “NRC” from the list of possible locations:

1. PRESS & HOLD Button 1 to turn MIC ON ().
2. Select the location by saying “NRC”.
3. Release Button 1 to turn MIC OFF.
4. The field will display  if the speech entry was accepted.

To enter a number with a decimal point – e.g. “47.31”:

1. PRESS & HOLD Button 1 to turn MIC ON ().
2. Enter the decimal number by saying “four seven point three one”.
3. Release Button 1 to turn MIC OFF.
4. The field will display  if the speech entry was accepted.

Once you are comfortable that you have understood the mechanism for entering the different data types, please put your ear protectors on. Go to the table in the lab corresponding to the first field on the sample form - i.e. the table labelled “Experiment Number”. You will see a data sheet that looks like this:

FIELD:

“EXPERIMENT NUMBER”

Please enter

685237

into the

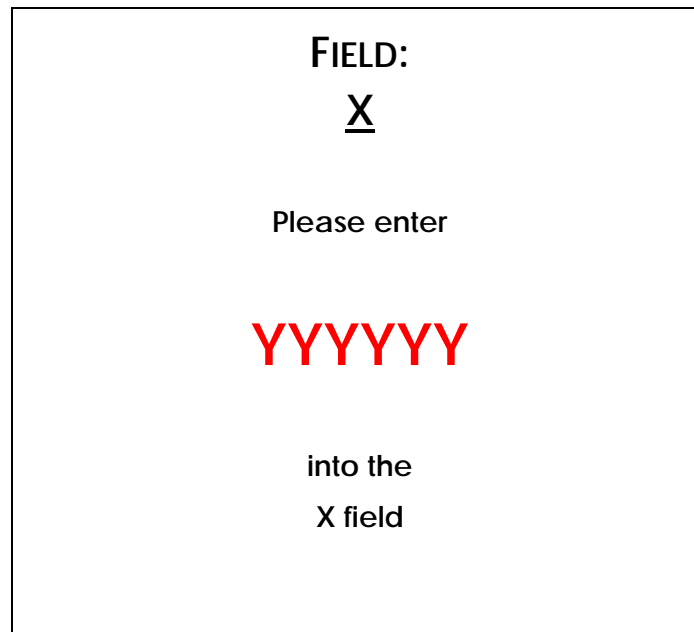
Experiment Number field

Enter the data as directed. Once you have entered the specified data, and the data is recognized, the cursor will automatically move to the next field. Wait until you see the cursor flash in the next field (or in the case of a drop-down box, the field become highlighted). Look to see the name of that field (i.e. the next field in the form). Go to the table in the room that has this label, enter the data shown, and repeat this process until you have filled in all four fields. Feel free to ask the researcher for assistance at any point. You may repeat the whole process for all four fields if you would like further practice – just ask the researcher.

Speech Input :: Experimental Session

You should now be familiar with how to enter data into a form on the Pocket PC using speech. Please read the following instructions carefully. If you have any questions before you begin, please feel free to ask the researcher. Please remember that you will not be able to ask any questions once the session begins due to the noise levels.

You will begin by standing at the Start table in the room. The researcher will give you a Pocket PC on which will be the main experimental data entry form. This is identical to the one you have just used in the training session except that it has a few more fields. You will see the cursor flashing in the topmost field on your form. Just as in the training session, you will go to the table in the room corresponding to the first field in your form. On that table you will see an instruction sheet similar to the following:



FIELD:
X

Please enter

YYYYYYY

into the
X field

Enter the data indicated. Once you have finished entering the data, the cursor will advance to the next field in the form. Wait until you see the cursor flash in the next field (or in the case of a drop-down box, the field become highlighted). You should move to the table in the room that corresponds to that field. Read the data entry instructions, enter the data, and then continue this process until you have visited all the fields/tables. When you are finished return to the researcher at the Start table.

We are interested in evaluating the effectiveness of the speech input technique when used under such noisy conditions. As a result of the noise levels, it is possible that your speech input will not always be successfully recognised. This will either result in a field being left blank or incorrect data being displayed. It is important for us to see when this happens so **PLEASE DO NOT CORRECT ANY ERRONEOUS ENTRIES** – even if you notice an error or omission, please just move on to the next field in the form.

While you are entering data, five projectors will be used to project a series of photographs onto the walls of the lab. The majority of the photographs will show “safe” construction sites (see the example below).



An example of a "safe" construction site photograph

At random intervals, a "danger" photograph will be displayed by one of the projectors. There is only ONE type of "danger" photograph (see below).



The "danger" photograph

While you are entering data into the form on the Pocket PC, please keep a mental tally of how many times you think the "danger" photograph has been displayed. When you have finished filling out the electronic form and you have returned to the Start table, please let the researcher know how many "danger" photographs you have counted.

When you are ready to begin, please let the researcher know. He will make sure you have your ear protectors on correctly before commencing the experimental session.

REMEMBER:

- **DO NOT remove your ear protectors until the researcher has indicated to you it is safe to do so.** If, at any time during the session, you wish to stop please approach the researcher directly and he will take this as an indication that you wish to stop.
- **DO NOT CORRECT ERRONEOUS ENTRIES – move onto the next field even if you spot an error or an omission.**
- **WATCH OUT FOR THE "DANGER" PHOTOGRAPHS!**

Speech Input :: Workload Assessment - Introduction

We would like you to complete some tables designed to find out about your experiences during the tasks you have just completed. We are examining the "workload" you experienced. The factors that influence your experiences when interacting with the system to complete the tasks may come from the system itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by these factors may change as you use the different input techniques (i.e. the speech and the stylus). The physical parts of workload are easy to measure but the mental ones are harder.

Since workload is something that is experienced individually by each person, we need to measure it by asking each person to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually. This set of 6 scales was developed for you to use in evaluating your experiences in different tasks.

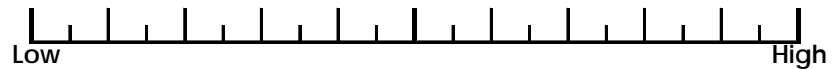
Please read the definitions of the scales carefully. If you have a question about any of the scales please ask the researcher. It is extremely important that they be clear to you.

Now that you have completed your data entry using speech input, we would like you to fill in the 6 scales. **You should evaluate the session by marking each scale at the point which matches your experience. Each line has a description at each end: please consider each scale individually. Please consider your responses carefully.** Your ratings will play an important role in the evaluation being conducted so your active participation is essential to the success of this experiment and is greatly appreciated. Please remember that you are free to terminate your contribution to this study at any time and you do not have to complete any of the scales if you do not wish to. If you have any questions, please do not hesitate to ask the researcher.

Speech Input :: Workload Assessment

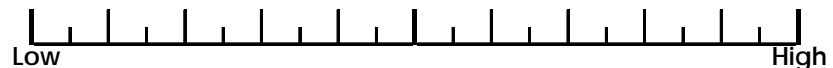
MENTAL DEMAND How much mental, visual, and auditory activity was required? (e.g. thinking, deciding, calculating, remembering, looking, searching, listening, scanning, looking etc.)

MENTAL DEMAND



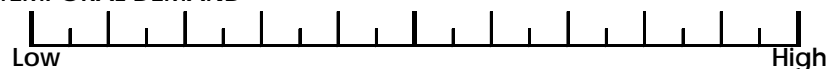
PHYSICAL DEMAND How much physical activity was required? (e.g. pushing, pulling, turning, controlling, activating etc.)

PHYSICAL DEMAND



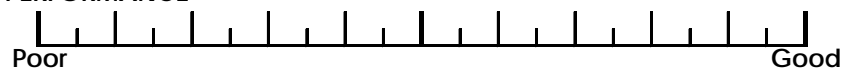
TEMPORAL DEMAND How much time pressure did you feel due to the rate or pace at (TIME PRESSURE) which the task elements occurred? (e.g. slow, leisurely, rapid, frantic)

TEMPORAL DEMAND



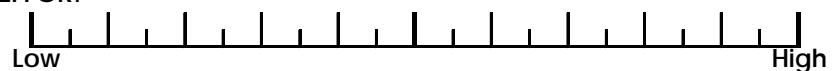
PERFORMANCE LEVEL How successful do you think you were in accomplishing the **ACHIEVED** task(s) set by the experimenter? How satisfied were you with your performance? Don't just think of your 'score', but how you felt you performed.

PERFORMANCE



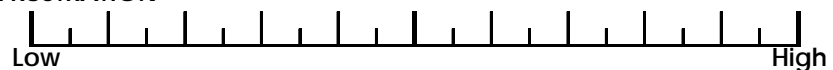
EFFORT EXPENDED How hard did you have to work (mentally and physically) to accomplish your level of performance?

EFFORT



FRUSTRATION LEVEL How much frustration did you experience? (e.g. were you relaxed, content, stressed, irritated, discouraged?)

FRUSTRATION




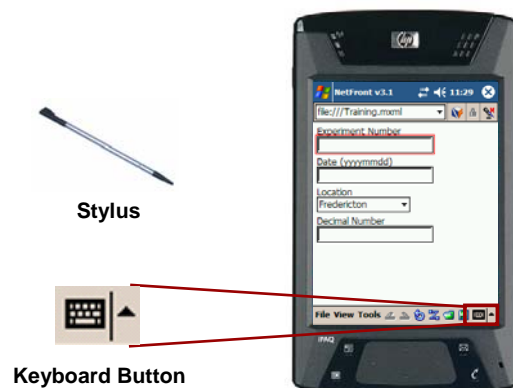
Stylus Input :: Training Session

Before you begin to complete the main form, we need to explain how to use the stylus and keypad to enter data. To do this, we have developed a sample form which will allow you to test out the stylus-based input without worrying about what your actions might do! Please feel free to ask questions at any point during the training session – it is important that you are comfortable with data entry using the stylus and keypad before you proceed to using the main form.

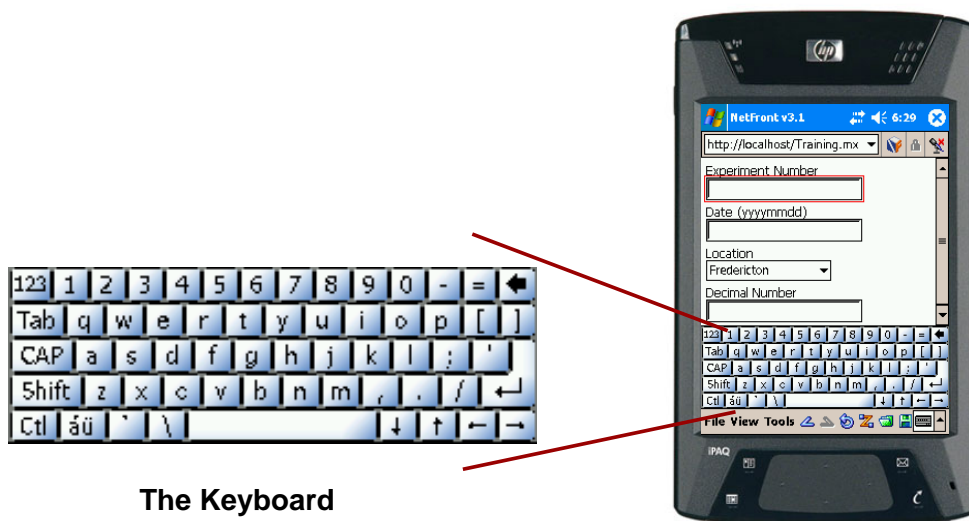
There are two components to learning how to use the stylus-based input: first, you need to become familiar with the mechanics of the stylus and the touch screen keypad on the Pocket PC and second, you need to learn how to enter different types of data.

Activating & Using the Touch screen Keyboard:

1. The Keyboard button is located on the bottom right ().
2. Push this button by tapping (single tap) the corresponding part of the touch screen with the stylus.

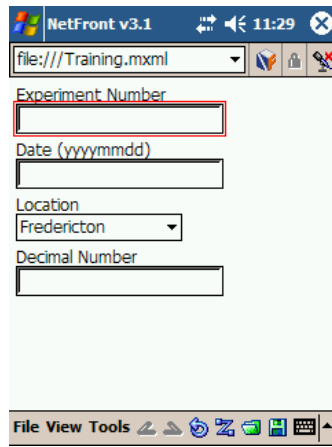


3. Once the Keyboard button has been pushed the keyboard appears on the bottom of the screen (see below).



4. To enter data in a field on the form, the field must first be selected. To select a field, touch the field with the stylus. For text fields, the cursor will flash in the selected field; for drop-down boxes, the field will become highlighted.
5. Enter data using the stylus and keyboard. To select an entry from a drop-down box, tap the small arrow to the right of the field and from the list of items that will appear, select the desired one by tapping (once) on the item. To enter text or numbers, select the letters/numbers you wish to enter into the field by tapping (single tap) on them using the stylus.

To help you familiarise yourself with the stylus-based input, we would like you to practice entering some data into the form (see below).




The Pocket PC showing the training form. Notice that the cursor is blinking in the first field ("Experiment Number").


Since you will be required to walk from table to table to enter the data when you are completing the main form, we would like you to do the same for this training session. Please read how to enter each type of data and then follow the subsequent directions.

To enter data in any given field, the cursor must first be flashing in that field (or in the case of a drop-down list, the field must be highlighted). To do this, simply tap the field with the stylus.

To enter a number – e.g. "685237":

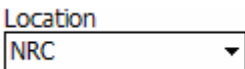
1. Press the Keyboard button () to display the keyboard (if it is not already showing)
2. Using the stylus, tap the following sequence of digits on the keyboard: 685237
3. Using the stylus, select the next field
4. The field will display 685237 if the data entry was correct.

To enter a date – e.g. 10th January 2006:



1. Press the Keyboard button () to display the keyboard (if it is not already showing)
2. Using the stylus, tap the following sequence of digits on the keyboard (yyyymmdd format): 20060110
3. Using the stylus, select the next field
5. The field will display 20060110 if the data entry was correct.

To select something from a drop-down list – e.g. to select “NRC” from the list of possible locations:

1. Tap on the little down arrow to the right of the field (▼) to display the list of options.
2. Select the option “NRC” by tapping (single tap) on the word “NRC”.

3. The field will display  if the list selection was accepted.

To enter a number with a decimal point – e.g. “47.31”:

1. Press the Keyboard button () to display the keyboard (if it is not already showing)
2. Using the stylus, tap the following sequence of digits on the keyboard: 47.31
3. Using the stylus, select the next field
4. The field will display  if the data entry was correct.

Once you are comfortable that you have understood the mechanism for entering the different data types, please put your ear protectors on. Go to the table in the lab corresponding to the first field on the sample form - i.e. the table labelled “Experiment Number”. You will see a data sheet that looks like this:

FIELD:

“EXPERIMENT NUMBER”

Please enter

685237

into the

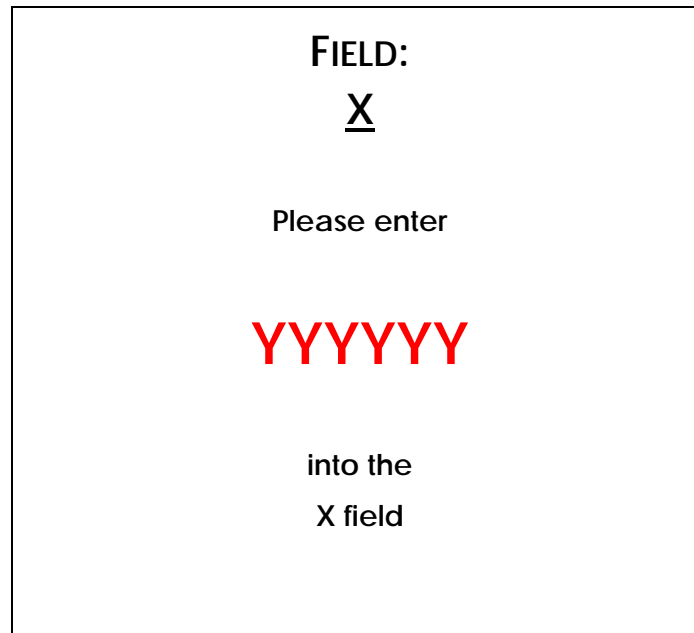
Experiment Number field

Enter the data as directed. Once you have entered the specified data, look to see the name of the next field in the form. Go to the table in the room that has this label, move the cursor to that field by tapping on the field, and enter the data shown (follow appropriate procedures if the field is a drop-down box). Repeat this process until you have filled in all four fields. Feel free to ask the researcher for assistance at any point. You may repeat the whole process for all four fields if you would like further practice – just ask the researcher.

Stylus Input :: Experimental Session

You should now be familiar with how to enter data into a form on the Pocket PC using the stylus and keyboard. Please read the following instructions carefully. If you have any questions before you begin, please feel free to ask the researcher. Please remember that you will not be able to ask any questions once the session begins due to the noise levels.

You will begin by standing at the Start table in the room. The researcher will give you a Pocket PC on which will be the main experimental data entry form. This is identical to the one you have just used in the training session except that it has a few more fields. You will see the cursor flashing in the topmost field on your form. Just as in the training session, you will go to the table in the room corresponding to the first field in your form. On that table you will see an instruction sheet similar to the following:



The image shows a rectangular box containing the following text:

FIELD:
X

Please enter

YYYYYYY

into the
X field

Enter the data indicated. Once you have finished entering the data, move the cursor to the next field then walk to the table in the room that corresponds to that field. Read the data entry instructions, enter the data, and then continue this process until you have visited all the fields/tables. When you are finished return to the researcher at the Start table.

We are interested in evaluating the effectiveness of the stylus input technique when used under mobile conditions. As a result of the fact that you are standing and moving when entering the data, it is possible that your stylus input will not always be correct. This will either result in a field being left blank or incorrect data being displayed. It is important for us to see when this happens so **PLEASE DO NOT CORRECT ANY ERRONEOUS ENTRIES** – even if you notice an error or omission, please just move on to the next field in the form.

While you are entering data, five projectors will be used to project a series of photographs onto the walls of the lab. The majority of the photographs will show “safe” construction sites (see the example below).



An example of a "safe" construction site photograph

At random intervals, a "danger" photograph will be displayed by one of the projectors. There is only ONE type of "danger" photograph (see below).



The "danger" photograph

While you are entering data into the form on the Pocket PC, please keep a mental tally of how many times you think the "danger" photograph has been displayed. When you have finished filling out the electronic form and you have returned to the Start table, please let the researcher know how many "danger" photographs you have counted.

When you are ready to begin, please let the researcher know. He will make sure you have your ear protectors on correctly before commencing the experimental session.

REMEMBER:

- **DO NOT remove your ear protectors until the researcher has indicated to you it is safe to do so.** If, at any time during the session, you wish to stop please approach the researcher directly and he will take this as an indication that you wish to stop.
- **DO NOT CORRECT ERRONEOUS ENTRIES – move onto the next field even if you spot an error or an omission.**
- **YOU MAY NEED TO SCROLL down through the form using the scrollbar on the right hand side to reach the bottom fields when the keyboard is displayed.**
- **WATCH OUT FOR THE "DANGER" PHOTOGRAPHS!**

Stylus Input :: Workload Assessment - Introduction

We would like you to complete some tables designed to find out about your experiences during the tasks you have just completed. We are examining the "workload" you experienced. The factors that influence your experiences when interacting with the system to complete the tasks may come from the system itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by these factors may change as you use the different input techniques (i.e. the speech and the stylus). The physical parts of workload are easy to measure but the mental ones are harder.

Since workload is something that is experienced individually by each person, we need to measure it by asking each person to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually. This set of 6 scales was developed for you to use in evaluating your experiences in different tasks.

Please read the definitions of the scales carefully. If you have a question about any of the scales please ask the researcher. It is extremely important that they be clear to you.

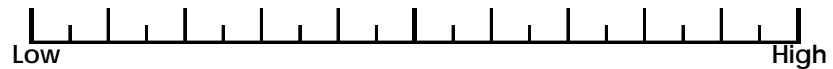
Now that you have completed your data entry using the stylus and keyboard, we would like you to fill in the 6 scales. **You should evaluate the session by marking each scale at the point which matches your experience. Each line has a description at each end: please consider each scale individually. Please consider your responses carefully.** Your ratings will play an important role in the evaluation being conducted so your active participation is essential to the success of this experiment and is greatly appreciated. Please remember that you are free to terminate your contribution to this study at any time and you do not have to complete any of the scales if you do not wish to. If you have any questions, please do not hesitate to ask the researcher.

Participant #: _____

Stylus Input :: Workload Assessment

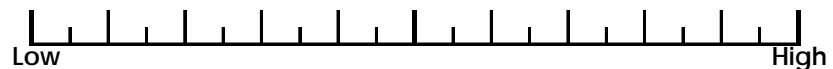
MENTAL DEMAND How much mental, visual, and auditory activity was required? (e.g. thinking, deciding, calculating, remembering, looking, searching, listening, scanning, looking etc.)

MENTAL DEMAND



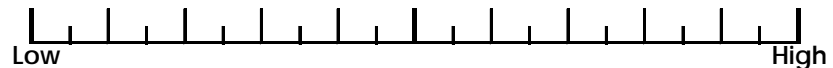
PHYSICAL DEMAND How much physical activity was required? (e.g. pushing, pulling, turning, controlling, activating etc.)

PHYSICAL DEMAND



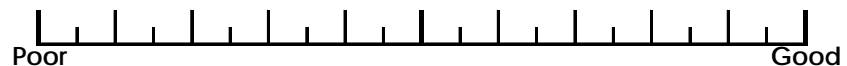
TEMPORAL DEMAND How much time pressure did you feel due to the rate or pace at (TIME PRESSURE) which the task elements occurred? (e.g. slow, leisurely, rapid, frantic)

TEMPORAL DEMAND



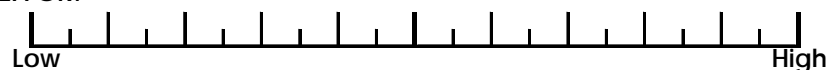
PERFORMANCE LEVEL How successful do you think you were in accomplishing the **ACHIEVED** task(s) set by the experimenter? How satisfied were you with your performance? Don't just think of your 'score', but how you felt you performed.

PERFORMANCE



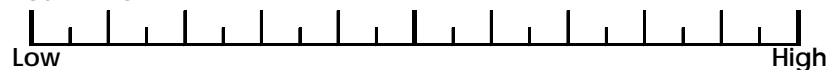
EFFORT EXPENDED How hard did you have to work (mentally and physically) to accomplish your level of performance?

EFFORT



FRUSTRATION LEVEL How much frustration did you experience? (e.g. were you relaxed, content, stressed, irritated, discouraged?)

FRUSTRATION



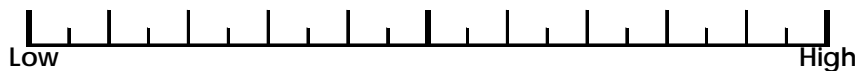
Participant #: _____

Preference Assessment

You have now completed the form using both input techniques. We would like you to tell us which method of interaction (stylus-based or speech-based) you preferred. Below is a scale identical to those that you have been using to indicate your workload. We would like you to complete it as before but this time put an 'S' on the scale to indicate your preference for the speech-based interaction and a 'SK' on the scale to indicate your preference for the stylus & keyboard based interaction. So, for example, if you preferred the stylus-based interaction, you would put the 'SK' nearer the High end of the scale than the 'S'; if you preferred the speech-based interaction, you would put the 'S' nearer the High end of the scale than the 'SK'.

Please now indicate your preferences on the following scale.

OVERALL PREFERENCE Please rate your preference for the interaction styles using the following identifiers:
 (S) for speech-based interaction
 (SK) for stylus & keyboard based interaction



You might have preferred one input style over another for certain types of data entry. For each of the following, please indicate your preference by ticking the appropriate box:

For entering numbers (without decimal points): No Preference
 Speech
 Stylus & Keyboard

For entering numbers (with decimal points): No Preference
 Speech
 Stylus & Keyboard

For entering dates: No Preference
 Speech
 Stylus & Keyboard

For selecting items from a drop-down list box: No Preference
 Speech
 Stylus & Keyboard

If you have any other comments regarding this experiment, please take a moment to note them here:

You have now completed this experiment. Thank you for your participation.

Appendix C: Experiment Results

Appendix C contains the raw data collect from participants during the experiment. Data Included with Appendix C are:

- Participants Background Information
- Session 1 – Recorded Data for Speech
- Session 1 – Recorded Data for Stylus
- Session 1 – NASA TLX Workload Test
- Session 2 – Recorded Data for Speech
- Session 2 – Recorded Data for Stylus
- Session 2 – NASA TLX Workload Test
- Overall and Individual Preferences

Participants Background Information

Participant #	Faculty	Age	Sex	Used a PPC?	How often	Used Speech based input?	How often
6	Business Admin	21-25	F	Have	A couple of Times	Have	Very few times
16	Civil Engineering (Masters in Concrete)	26-30	F	Have	Not Often	Have	Not Often
13	Civil Engineering (Masters)	21-25	F	Have	once a month	Have	once a year
5	Computer Science	21-25	F	Have	Infrequently (once a month)	Have	Not Much
1	Civil Engineering	21-25	M	Have	Seldom	Never	-
15	Computer Science	15-20	M	Have	not often, 2-3 months/yr	Never	-
2	Electrical Engineering	21-25	M	Have	Only a couple times	Have	Very few times
4	Forestry	21-25	M	Have	15 Time in past yr	Never	-
17	Forestry Technologist	21-25	M	Have	one week	Never	-
7	Other (Sales Clerk)	15-20	M	Have	on occasion	Never	-
8	Science	15-20	M	Have	Few Times a Week	Never	-
3	Business Admin	15-20	F	Never	-	Never	-
12	Civil Engineering	21-25	F	Never	-	Never	-
18	Nursing	21-25	F	Never	-	Never	-
10	Other (Home-Maker)	41-45	F	Never	-	Never	-
14	Civil Engineering	21-25	M	Never	-	Never	-
9	Civil Engineering (Masters)	21-25	M	Never	-	Never	-
11	NB Power (Projcet Manager)	41-45	M	Never	-	Never	-

Session 1 – Recorded Data for Speech

			Contract Number	Date (yyyymmdd)	Contractor	Material Used	Shipped Concrete	Placed Concrete	Test Loca
#	Noise Level	Modality	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
9	70-80	Speech	058391	20051127	Alex Boyd	Steel	30.8	082.09	Lower Fred
12	70-80	Speech	058391	20051122	Alex Boyd	Chip Rock	30.8	82.09	Lower Fred
16	70-80	Speech	258391	20051122	Alex Boyd	Steel	330.8	82.09	Lower Fred
1	80-90	Speech	2580391	20051122	Alex Boyd	Steel	30.8	82.09	Biggs St
6	80-90	Speech	22583910	20051129	Alex Boyd	Chip Rock	98.8	8882.09	Lower Fred
13	80-90	Speech	258391	2005112	Alex Boyd	Steel	30.583	8200.09	Lower Fred
2	90-100	Speech	120580391	20051122	Alex Boyd	Chip Rock	03.8	882.09	Lower Fred
5	90-100	Speech	12583095	20051122	Connie Adsett	Chip Rock	3000.8	804801	Lower Fred
8	90-100	Speech	058391	20011102	Alex Boyd	Chip Rock	8.8	882.09	Lower Fred

Session 1 – Recorded Data for Stylus

			Contract Number	Date (yyyymmdd)	Contractor	Material Used	Shipped Concrete	Placed Concrete	Test Loca
#	Noise Level	Modality	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
10	70-80	Stylus	258139	20051122	Alex Boyd	Tar	30.8	82.09	Lower Fred
14	70-80	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
15	70-80	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
3	80-90	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
7	80-90	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
11	80-90	Stylus	28	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
4	90-100	Stylus	258391	22nd NOVember 2005	Alex Boyd	Steel	30.8	82.09	Lower Fred
17	90-100	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred
18	90-100	Stylus	258391	20051122	Alex Boyd	Steel	30.8	82.09	Lower Fred

Legend

Blue Data = Correct data that should have been entered

Session 1 – NASA TLX Workload Test

Participant #	Noise Level	Modality	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
1	80-90	Speech	16	2	2	15	6	1
2	90-100	Speech	15	8	10	3	16	15
3	80-90	Stylus	20	5	14	15	17	9
4	90-100	Stylus	15	5	13	15	15	17
5	90-100	Speech	11	3	14	14	16	5
6	80-90	Speech	9	4	7	4	6	9
7	80-90	Stylus	16	6	18	17	19	18
8	90-100	Speech	11	1	8	6	11	4
9	70-80	Speech	8	4	4	15	6	3
10	70-80	Stylus	18	2	11	9	12	16
11	80-90	Stylus	15	8	12	7	13	15
12	70-80	Speech	15	2	4	17	11	11
13	80-90	Speech	3	1	3	20	3	3
14	70-80	Stylus	4	1	2	18	3	1
15	70-80	Stylus	17	18	16	17	18	16
16	70-80	Speech	12	2	7	17	5	3
17	90-100	Stylus	11	4	11	17	8	6
18	90-100	Stylus	2	2	4	18	4	2

Session 2 – Recorded Data for Speech

			Date (yyyymmdd)	Contractor	Contract Number	Material Used	Test Location	Shipped Concrete	Placed Concrete
#	Noise Level	Modality	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
10	70-80	Speech	20060207	Sally Wilson	007452	Concrete	Plaster Rock	9.058	43.37
14	70-80	Speech	20060207	Sally Wilson	8007452	Concrete	Plaster Rock	9.058	14.37
15	70-80	Speech	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
3	80-90	Speech	20060102	Sally Wilson	00705820	Concrete	Plaster Rock	09.058	14.37
7	80-90	Speech	20030602	Sally Wilson	007452	Concrete	Biggs Street	9.3058	14.35
11	80-90	Speech	20061102	Sally Wilson	0074508	Concrete	Plaster Rock	8.058	4.037
4	90-100	Speech	<Blank>	Connie Adsett	700745238	Concrete	Plaster Rock	9.058	14.37
17	90-100	Speech	20060207	Sally Wilson	900174752	Concrete	Plaster Rock	98.058	14.37
18	90-100	Speech	20060102	Sally Wilson	00074523	Chip Rock	Plaster Rock	90.058	41.37

Session 2 – Recorded Data for Stylus

			Date (yyyymmdd)	Contractor	Contract Number	Material Used	Test Location	Shipped Concrete	Placed Concrete
#	Noise Level	Modality	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
9	70-80	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
12	70-80	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
16	70-80	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.39
1	80-90	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
6	80-90	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
13	80-90	Stylus	20060702	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
2	90-100	Stylus	20060602	Sally Wilson	07452	Concrete	Plaster Rock	9.058	14.37
5	90-100	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37
8	90-100	Stylus	20060602	Sally Wilson	007452	Concrete	Plaster Rock	9.058	14.37

Legend

Blue Data = Correct data that should have been entered

Session 2 – NASA TLX Workload Test

Participant #	Noise Level	Modality	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
1	80-90	Stylus	12	6	1	18	10	1
2	90-100	Stylus	18	5	18	16	12	12
3	80-90	Speech	19	7	9	8	19	15
4	90-100	Speech	12	4	9	9	9	17
5	90-100	Stylus	16	4	16	14	15	14
6	80-90	Stylus	7	4	6	18	9	5
7	80-90	Speech	19	11	17	6	17	19
8	90-100	Stylus	5	1	6	18	4	1
9	70-80	Stylus	11	10	5	14	6	4
10	70-80	Speech	4	5	5	11	5	10
11	80-90	Speech	13	9	12	15	14	7
12	70-80	Stylus	17	13	13	14	14	12
13	80-90	Stylus	6	1	3	20	2	6
14	70-80	Speech	5	1	1	18	3	2
15	70-80	Speech	15	10	14	17	13	11
16	70-80	Stylus	18	4	15	17	14	8
17	90-100	Speech	6	11	8	17	5	7
18	90-100	Speech	8	2	4	10	8	10

Overall and Individual Preferences

Participant #	Noise Level	Overall Preference Stylus	Overall Preference Speech	Preferred Modality	Numbers	Decimal Numbers	Dates	Drop-Down Lists
1	80-90	14	18	Speech	Speech	Speech	Speech	Speech
2	90-100	14	13	Stylus	Stylus	Stylus	Speech	Speech
3	80-90	18	15	Stylus	Stylus	Stylus	Speech	Speech
4	90-100	11	15	Speech	Speech	Speech	NP	Speech
5	90-100	8	14	Speech	Speech	Stylus	Speech	Speech
6	80-90	17	13	Stylus	Speech	Stylus	Stylus	Speech
7	80-90	17	5	Stylus	Stylus	Stylus	Stylus	Speech
8	90-100	19	14	Stylus	Stylus	Stylus	Stylus	Speech
9	70-80	12	15	Speech	Speech	Speech	NP	NP
10	70-80	nd	18	Speech	Speech	Stylus	Speech	Speech
11	80-90	7	14	Speech	Speech	Speech	Speech	Stylus
12	70-80	15	18	Speech	Speech	Stylus	Speech	Speech
13	80-90	3	16	Speech	Speech	Speech	Speech	Speech
14	70-80	17	15	Stylus	Stylus	Stylus	NP	NP
15	70-80	11	18	Speech	Speech	Speech	Speech	Speech
16	70-80	15	18	Speech	NP	Stylus	Speech	Speech
17	90-100	13	10	Stylus	Speech	Stylus	Speech	Stylus
18	90-100	20	11	Stylus	NP	Stylus	Stylus	NP

Appendix D: Experiment Statistics

Appendix D contains the statistical analysis performed on Appendix C's raw data. These statistical analyses are mostly calculations that are used for further analysis using Minitab. Included with Appendix D are:

- Participant # 1-9 – Combined Experiment Statistics
- Participant # 10-18 – Combined Experiment Statistics
- Speech - Combined Experiment Statistics
- Stylus - Combined Experiment Statistics
- Avg. Task Time Statistics
- Avg. Distractions
- TLX Workload Statistics
- Precision/Accuracy Rate Statistics
- Average Data Precision Rates
- Isolated Session 2 – Distraction Awareness
- Preferences

Participant # 1-9 – Combined Experiment Statistics

Pt #	Group	Modality	Task Time (seconds)	Actual # Distraction	Identified # Distractions	% Distractions Identified	Number	Date	DDL1	DDL2	DDL3	Decimal	Decimal	All DDLs	All Decimals	Entry Precision Rate	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall Workload
1	2	Speech	77	2	0	0.0%	0	1	1	1	0	1	1	66.7%	100.0%	71.4%	16	2	2	15	6	1	7.0
1	2	Stylus	148	4	3	75.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	12	6	1	18	10	1	8.0
2	3	Speech	90	3	2	66.7%	0	1	1	0	1	0	0	66.7%	0.0%	42.9%	15	8	10	3	16	15	11.2
2	3	Stylus	159	5	5	100.0%	0	1	1	1	1	1	1	100.0%	100.0%	85.7%	18	5	18	16	12	12	13.5
3	2	Speech	100	3	3	100.0%	0	0	1	1	1	1	1	100.0%	100.0%	71.4%	19	7	9	8	19	15	12.8
3	2	Stylus	103	3	3	100.0%	0	1	1	1	1	1	1	100.0%	100.0%	85.7%	20	5	14	15	17	9	13.3
4	3	Speech	112	3	3	100.0%	0	0	0	1	1	1	1	66.7%	100.0%	57.1%	12	4	9	9	9	17	10.0
4	3	Stylus	156	5	4	80.0%	0	1	1	1	1	1	1	100.0%	100.0%	85.7%	15	5	13	15	15	17	13.3
5	3	Speech	89	3	3	100.0%	0	1	0	0	1	0	0	33.3%	0.0%	28.6%	11	3	14	14	16	5	10.5
5	3	Stylus	144	4	3	75.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	16	4	16	14	15	14	13.2
6	2	Speech	97	3	2	66.7%	0	0	1	0	1	0	0	66.7%	0.0%	28.6%	9	4	7	4	6	9	6.5
6	2	Stylus	137	4	4	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	7	4	6	18	9	5	8.2
7	2	Speech	90	2	3	150.0%	1	0	1	1	0	0	0	66.7%	0.0%	42.9%	19	11	17	6	17	19	14.8
7	2	Stylus	94	3	3	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	16	6	18	17	19	18	15.7
8	3	Speech	85	2	3	150.0%	0	0	1	0	1	0	0	66.7%	0.0%	28.6%	11	1	8	6	11	4	6.8
8	3	Stylus	118	3	3	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	5	1	6	18	4	1	5.8
9	1	Speech	102	3	3	100.0%	0	0	1	1	1	1	1	100.0%	100.0%	71.4%	8	4	4	15	6	3	6.7
9	1	Stylus	163	5	5	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	11	10	5	14	6	4	8.3

Legend

Blue Labels = NASA TLX WorkLoad Statistics

Green Labels = Precision Rates (0 = Incorrect Data Entry, 1 = Correct Data Entry)

DDL = Drop-Down Lists

Participant # 10-18 – Combined Experiment Statistics

Pt #	Group	Modality	Task Time (seconds)	Actual # Distraction	Identified # Distractions	% Distractions Identified	Number	Date	DDL1	DDL2	DDL3	Decimal	Decimal	All DDLs	All Decimals	Entry Precision Rate	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall Workload
10	1	Speech	123	4	5	125.0%	1	0	1	1	1	1	0	100.0%	50.0%	71.4%	4	5	5	11	5	10	6.7
10	1	Stylus	161	5	4	80.0%	0	1	1	0	1	1	1	66.7%	100.0%	71.4%	18	2	11	9	12	16	11.3
11	2	Speech	106	3	4	133.3%	0	0	1	1	1	0	0	100.0%	0.0%	42.9%	13	9	12	15	14	7	11.7
11	2	Stylus	120	4	1	25.0%	0	1	1	1	1	1	1	100.0%	100.0%	85.7%	15	8	12	7	13	15	11.7
12	1	Speech	110	3	3	100.0%	0	1	1	0	1	1	1	66.7%	100.0%	71.4%	15	2	4	17	11	11	10.0
12	1	Stylus	164	5	4	80.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	17	13	13	14	14	12	13.8
13	2	Speech	82	2	3	150.0%	1	0	1	1	1	0	0	100.0%	0.0%	57.1%	3	1	3	20	3	3	5.5
13	2	Stylus	116	3	3	100.0%	1	0	1	1	1	1	1	100.0%	100.0%	85.7%	6	1	3	20	2	6	6.3
14	1	Speech	136	4	5	125.0%	0	0	1	1	1	1	1	100.0%	100.0%	71.4%	5	1	1	18	3	2	5.0
14	1	Stylus	138	4	4	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	4	1	2	18	3	1	4.8
15	1	Speech	109	3	3	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	15	10	14	17	13	11	13.3
15	1	Stylus	139	4	4	100.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	17	18	16	17	18	16	17.0
16	1	Speech	114	3	2	66.7%	1	1	1	1	1	0	1	100.0%	50.0%	85.7%	12	2	7	17	5	3	7.7
16	1	Stylus	160	4	3	75.0%	1	1	1	1	1	1	0	100.0%	50.0%	85.7%	18	4	15	17	14	8	12.7
17	3	Speech	88	2	3	150.0%	0	0	1	1	1	0	1	100.0%	50.0%	57.1%	6	11	8	17	5	7	9.0
17	3	Stylus	99	3	2	66.7%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	11	4	11	17	8	6	9.5
18	3	Speech	110	3	2	66.7%	0	0	1	0	1	0	0	66.7%	0.0%	28.6%	8	2	4	10	8	10	7.0
18	3	Stylus	138	4	5	125.0%	1	1	1	1	1	1	1	100.0%	100.0%	100.0%	2	2	4	18	4	2	5.3

Legend

Blue Labels = NASA TLX WorkLoad Statistics

Green Labels = Precision Rates (0 = Incorrect Data Entry, 1 = Correct Data Entry)

DDL = Drop-Down Lists

Speech - Combined Experiment Statistics

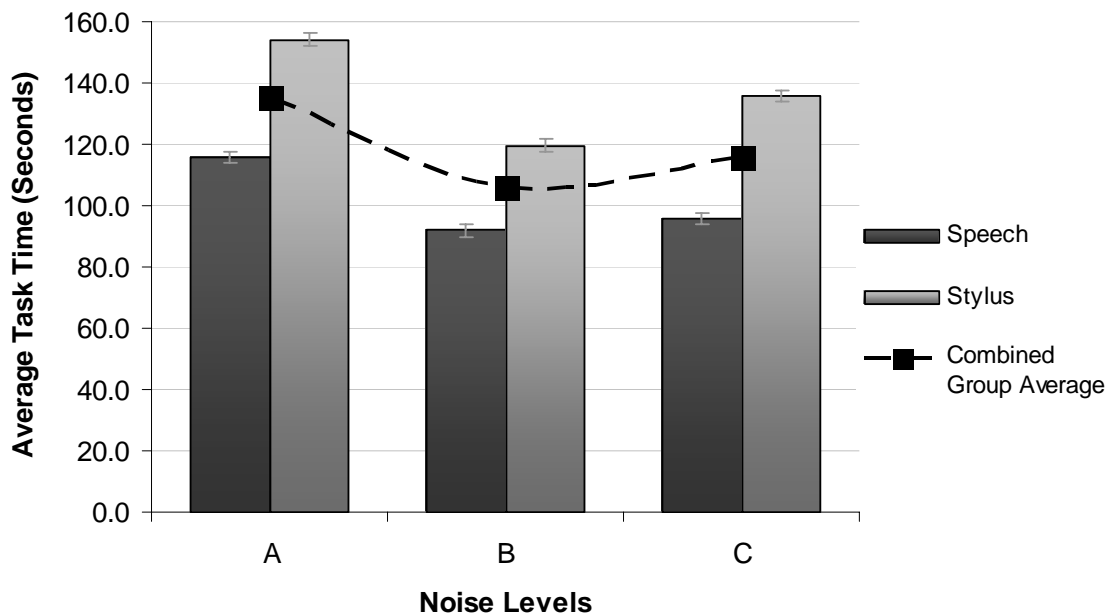
Pt #	Group	Modality	Task Time (secs)	Actual # Distraction	Identified # Distractions	% Distractions Identified	Number	Date	DDL1	DDL2	DDL3	Decimal	Decimal	All DDLs	All Decimals	Entry Precision Rate	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall Workload
9	1	Speech	102	3	3	1.0	0	0	1	1	1	1	1	1.0	1.0	0.7	8	4	4	15	6	3	6.7
10	1	Speech	123	4	5	1.3	1	0	1	1	1	1	0	1.0	0.5	0.7	4	5	5	11	5	10	6.7
12	1	Speech	110	3	3	1.0	0	1	1	0	1	1	1	0.7	1.0	0.7	15	2	4	17	11	11	10.0
14	1	Speech	136	4	5	1.3	0	0	1	1	1	1	1	1.0	1.0	0.7	5	1	1	18	3	2	5.0
15	1	Speech	109	3	3	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	15	10	14	17	13	11	13.3
16	1	Speech	114	3	2	0.7	1	1	1	1	1	0	1	1.0	0.5	0.9	12	2	7	17	5	3	7.7
1	2	Speech	77	2	0	0.0	0	1	1	1	0	1	1	0.7	1.0	0.7	16	2	2	15	6	1	7.0
3	2	Speech	100	3	3	1.0	0	0	1	1	1	1	1	1.0	1.0	0.7	19	7	9	8	19	15	12.8
6	2	Speech	97	3	2	0.7	0	0	1	0	1	0	0	0.7	0.0	0.3	9	4	7	4	6	9	6.5
7	2	Speech	90	2	3	1.5	1	0	1	1	0	0	0	0.7	0.0	0.4	19	11	17	6	17	19	14.8
11	2	Speech	106	3	4	1.3	0	0	1	1	1	0	0	1.0	0.0	0.4	13	9	12	15	14	7	11.7
13	2	Speech	82	2	3	1.5	1	0	1	1	1	0	0	1.0	0.0	0.6	3	1	3	20	3	3	5.5
2	3	Speech	90	3	2	0.7	0	1	1	0	1	0	0	0.7	0.0	0.4	15	8	10	3	16	15	11.2
4	3	Speech	112	3	3	1.0	0	0	0	1	1	1	1	0.7	1.0	0.6	12	4	9	9	9	17	10.0
5	3	Speech	89	3	3	1.0	0	1	0	0	1	0	0	0.3	0.0	0.3	11	3	14	14	16	5	10.5
8	3	Speech	85	2	3	1.5	0	0	1	0	1	0	0	0.7	0.0	0.3	11	1	8	6	11	4	6.8
17	3	Speech	88	2	3	1.5	0	0	1	1	1	0	1	1.0	0.5	0.6	6	11	8	17	5	7	9.0
18	3	Speech	110	3	2	0.7	0	0	1	0	1	0	0	0.7	0.0	0.3	8	2	4	10	8	10	7.0

Stylus - Combined Experiment Statistics

Pt #	Group	Modality	Task Time (secs)	Actual # Distraction	Identified # Distractions	% Distractions Identified	Number	Date	DDL1	DDL2	DDL3	Decimal	Decimal	All DDLs	All Decimals	Entry Precision Rate	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall Workload
9	1	Stylus	163	5	5	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	11	10	5	14	6	4	8.3
10	1	Stylus	161	5	4	0.8	0	1	1	0	1	1	1	0.7	1.0	0.7	18	2	11	9	12	16	11.3
12	1	Stylus	164	5	4	0.8	1	1	1	1	1	1	1	1.0	1.0	1.0	17	13	13	14	14	12	13.8
14	1	Stylus	138	4	4	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	4	1	2	18	3	1	4.8
15	1	Stylus	139	4	4	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	17	18	16	17	18	16	17.0
16	1	Stylus	160	4	3	0.8	1	1	1	1	1	1	0	1.0	0.5	0.9	18	4	15	17	14	8	12.7
1	2	Stylus	148	4	3	0.8	1	1	1	1	1	1	1	1.0	1.0	1.0	12	6	1	18	10	1	8.0
3	2	Stylus	103	3	3	1.0	0	1	1	1	1	1	1	1.0	1.0	0.9	20	5	14	15	17	9	13.3
6	2	Stylus	137	4	4	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	7	4	6	18	9	5	8.2
7	2	Stylus	94	3	3	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	16	6	18	17	19	18	15.7
11	2	Stylus	120	4	1	0.3	0	1	1	1	1	1	1	1.0	1.0	0.9	15	8	12	7	13	15	11.7
13	2	Stylus	116	3	3	1.0	1	0	1	1	1	1	1	1.0	1.0	0.9	6	1	3	20	2	6	6.3
2	3	Stylus	159	5	5	1.0	0	1	1	1	1	1	1	1.0	1.0	0.9	18	5	18	16	12	12	13.5
4	3	Stylus	156	5	4	0.8	0	1	1	1	1	1	1	1.0	1.0	0.9	15	5	13	15	15	17	13.3
5	3	Stylus	144	4	3	0.8	1	1	1	1	1	1	1	1.0	1.0	1.0	16	4	16	14	15	14	13.2
8	3	Stylus	118	3	3	1.0	1	1	1	1	1	1	1	1.0	1.0	1.0	5	1	6	18	4	1	5.8
17	3	Stylus	99	3	2	0.7	1	1	1	1	1	1	1	1.0	1.0	1.0	11	4	11	17	8	6	9.5
18	3	Stylus	138	4	5	1.3	1	1	1	1	1	1	1	1.0	1.0	1.0	2	2	4	18	4	2	5.3

Avg. Task Time Statistics

Noise Level	Speech	Stylus	
A (70-80 dB)	115.7	154.2	134.9
B (80-90 dB)	92.0	119.7	105.8
C (90-100 dB)	95.7	135.7	115.7
	101.1	136.5	



Avg. Distractions

Noise Level	Speech	Stylus	
A (70-80 dB)	102.8%	89.2%	96.0%
B (80-90 dB)	100.0%	83.3%	91.7%
C (90-100 dB)	105.6%	91.1%	98.3%
	102.8%	87.9%	

TLX Workload Statistics

Avg. Mental Demand

Noise Level	Speech	Stylus	
A (70-80 dB)	9.8	14.2	12.0
B (80-90 dB)	13.2	12.7	12.9
C (90-100 dB)	10.5	11.2	10.8
	11.2	12.7	

Avg. Physical Demand

Noise Level	Speech	Stylus	
A (70-80 dB)	4.0	8.0	6.0
B (80-90 dB)	5.7	5.0	5.3
C (90-100 dB)	4.8	3.5	4.2
	4.8	5.5	

Avg. Temporal Demand

Noise Level	Speech	Stylus	
A (70-80 dB)	5.8	10.3	8.1
B (80-90 dB)	8.3	9.0	8.7
C (90-100 dB)	8.8	11.3	10.1
	7.7	10.2	

Avg. Performance

Noise Level	Speech	Stylus	
A (70-80 dB)	15.8	14.8	15.3
B (80-90 dB)	11.3	15.8	13.6
C (90-100 dB)	9.8	16.3	13.1
	12.3	15.7	

Avg. Effort

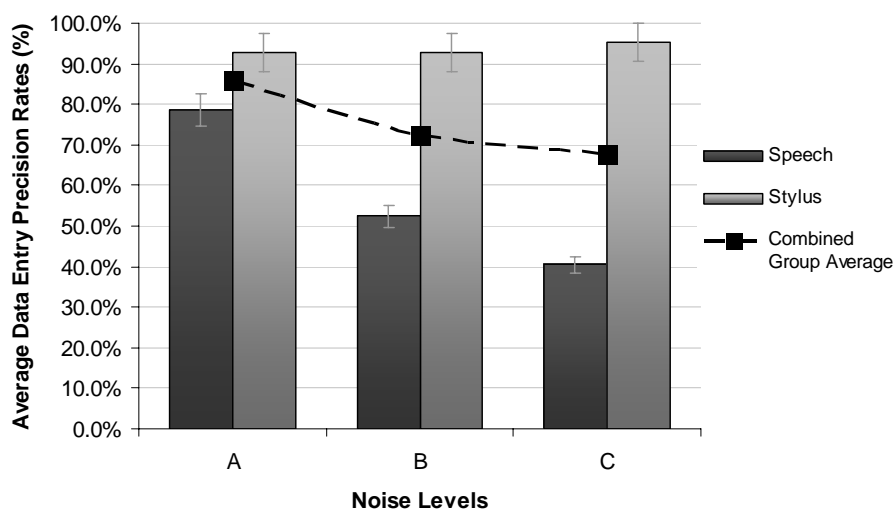
Noise Level	Speech	Stylus	
A (70-80 dB)	7.2	11.2	9.2
B (80-90 dB)	10.8	11.7	11.3
C (90-100 dB)	10.8	9.7	10.3
	9.6	10.8	

Avg. Frustration

Noise Level	Speech	Stylus	
A (70-80 dB)	6.7	9.5	8.1
B (80-90 dB)	9.0	9.0	9.0
C (90-100 dB)	9.7	8.7	9.2
	8.4	9.1	

Avg. Overall Workload

Noise Level	Speech	Stylus	
A (70-80 dB)	8.2	11.3	9.8
B (80-90 dB)	9.7	10.5	10.1
C (90-100 dB)	9.1	10.1	9.6
	9.0	10.7	

Precision/Accuracy Rate Statistics**Avg. Entry Precision Rate**

Noise Level	Speech	Stylus	
A (70-80 dB)	78.6%	92.9%	85.7%
B (80-90 dB)	52.4%	92.9%	72.6%
C (90-100 dB)	40.5%	95.2%	67.9%
	57.1%	93.7%	

Avg. Number Precision Rate

Noise Level	Speech	Stylus	
A (70-80 dB)	50.0%	83.3%	66.7%
B (80-90 dB)	33.3%	66.7%	50.0%
C (90-100 dB)	0.0%	66.7%	33.3%
	27.8%	72.2%	

Avg. Date Precision Rate

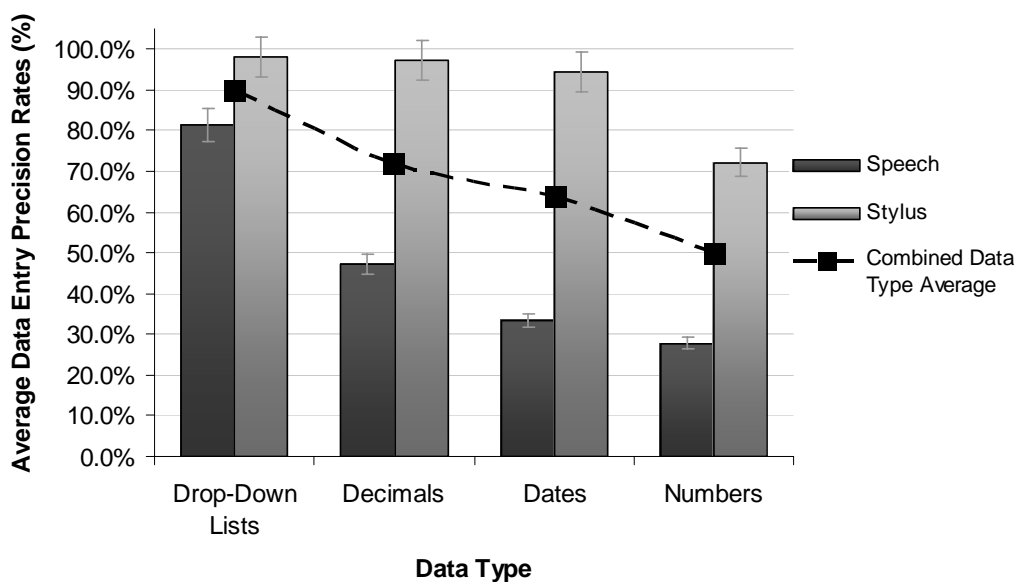
Noise Level	Speech	Stylus	
A (70-80 dB)	50.0%	100.0%	75.0%
B (80-90 dB)	16.7%	83.3%	50.0%
C (90-100 dB)	33.3%	100.0%	66.7%
	33.3%	94.4%	

Avg. Drop-Down List Precision Rate

Noise Level	Speech	Stylus	
A (70-80 dB)	94.4%	94.4%	94.4%
B (80-90 dB)	83.3%	100.0%	91.7%
C (90-100 dB)	66.7%	100.0%	83.3%
	82.6%	98.3%	

Avg. Decimal Precision Rate

Noise Level	Speech	Stylus	
A (70-80 dB)	83.3%	91.7%	87.5%
B (80-90 dB)	33.3%	100.0%	66.7%
C (90-100 dB)	25.0%	100.0%	62.5%
	47.2%	97.2%	

Average Data Precision Rates (1)

Data Type	Speech	Stylus	
Drop-Down Lists	81.5%	98.1%	89.8%
Decimals	47.2%	97.2%	72.2%
Dates	33.3%	94.4%	63.9%
Numbers	27.8%	72.2%	50.0%
	47.5%	90.5%	

Note: The next three pages (pg. 108-110) are the detailed results that combined make up the above chart and table.

Average Data Precision Rates (2)

Group	Modality	Data Type	Accuracy
1	Speech	Drop-Down Lists	1.0
1	Speech	Drop-Down Lists	1.0
1	Speech	Drop-Down Lists	0.7
1	Speech	Drop-Down Lists	1.0
1	Speech	Drop-Down Lists	1.0
1	Speech	Drop-Down Lists	1.0
2	Speech	Drop-Down Lists	0.7
2	Speech	Drop-Down Lists	1.0
2	Speech	Drop-Down Lists	0.7
2	Speech	Drop-Down Lists	0.7
2	Speech	Drop-Down Lists	1.0
2	Speech	Drop-Down Lists	1.0
3	Speech	Drop-Down Lists	0.7
3	Speech	Drop-Down Lists	0.7
3	Speech	Drop-Down Lists	0.3
3	Speech	Drop-Down Lists	0.7
3	Speech	Drop-Down Lists	1.0
3	Speech	Drop-Down Lists	0.7
1	Speech	Decimals	1.0
1	Speech	Decimals	0.5
1	Speech	Decimals	1.0
1	Speech	Decimals	1.0
1	Speech	Decimals	0.5
2	Speech	Decimals	1.0
2	Speech	Decimals	1.0
2	Speech	Decimals	0.0
2	Speech	Decimals	0.0
2	Speech	Decimals	0.0
2	Speech	Decimals	0.0
2	Speech	Decimals	0.0
3	Speech	Decimals	0.0
3	Speech	Decimals	1.0
3	Speech	Decimals	0.0
3	Speech	Decimals	0.0
3	Speech	Decimals	0.5
3	Speech	Decimals	0.0
1	Speech	Dates	0.0
1	Speech	Dates	0.0
1	Speech	Dates	1.0
1	Speech	Dates	0.0
1	Speech	Dates	1.0
1	Speech	Dates	1.0
2	Speech	Dates	1.0
2	Speech	Dates	0.0
2	Speech	Dates	0.0
2	Speech	Dates	0.0
2	Speech	Dates	0.0
2	Speech	Dates	0.0

Average Data Precision Rates (4)

Group	Modality	Data Type	Accuracy
2	Stylus	Decimals	1.0
2	Stylus	Decimals	1.0
2	Stylus	Decimals	1.0
2	Stylus	Decimals	1.0
2	Stylus	Decimals	1.0
2	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
3	Stylus	Decimals	1.0
1	Stylus	Dates	1.0
1	Stylus	Dates	1.0
1	Stylus	Dates	1.0
1	Stylus	Dates	1.0
1	Stylus	Dates	1.0
1	Stylus	Dates	1.0
2	Stylus	Dates	1.0
2	Stylus	Dates	1.0
2	Stylus	Dates	1.0
2	Stylus	Dates	1.0
2	Stylus	Dates	0.0
3	Stylus	Dates	1.0
3	Stylus	Dates	1.0
3	Stylus	Dates	1.0
3	Stylus	Dates	1.0
3	Stylus	Dates	1.0
3	Stylus	Dates	1.0
1	Stylus	Numbers	1.0
1	Stylus	Numbers	0.0
1	Stylus	Numbers	1.0
1	Stylus	Numbers	1.0
1	Stylus	Numbers	1.0
2	Stylus	Numbers	1.0
2	Stylus	Numbers	0.0
2	Stylus	Numbers	1.0
2	Stylus	Numbers	1.0
2	Stylus	Numbers	0.0
2	Stylus	Numbers	1.0
3	Stylus	Numbers	0.0
3	Stylus	Numbers	0.0
3	Stylus	Numbers	1.0
3	Stylus	Numbers	1.0
3	Stylus	Numbers	1.0
3	Stylus	Numbers	1.0
3	Stylus	Numbers	1.0

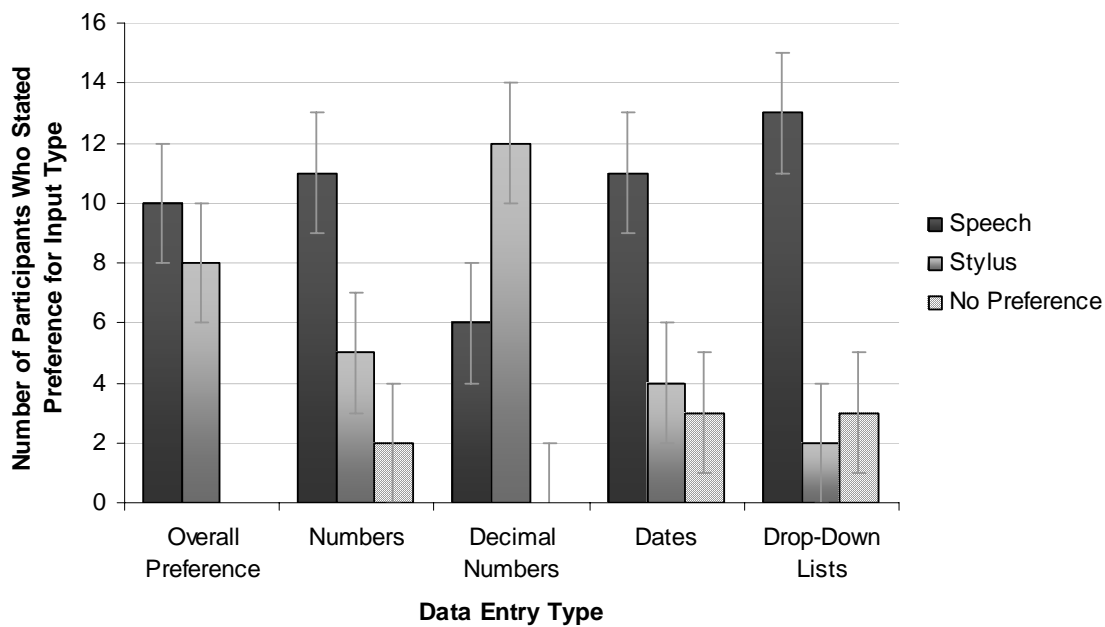
Isolated Session 2 – Distraction Awareness

Participant	Noise Level	Modality Session 2	Actual No. Distraction	Identified No. Distractions	% Distractions Identified Session 2
10	1	Speech	4	5	1.25
14	1	Speech	4	5	1.25
15	1	Speech	3	3	1.00
9	1	Stylus	5	5	1.00
12	1	Stylus	5	4	0.80
16	1	Stylus	4	3	0.75
3	2	Speech	3	3	1.00
7	2	Speech	2	3	1.50
11	2	Speech	3	4	1.33
1	2	Stylus	4	3	0.75
6	2	Stylus	4	4	1.00
13	2	Stylus	3	3	1.00
4	3	Speech	3	3	1.00
17	3	Speech	2	3	1.50
18	3	Speech	3	2	0.67
2	3	Stylus	5	5	1.00
5	3	Stylus	4	3	0.75
8	3	Stylus	3	3	1.00

Avg. Awareness

<u>Group</u>	<u>Speech</u>	<u>Stylus</u>	
1	1.2	0.9	1.0
2	1.3	0.9	1.1
3	1.1	0.9	1.0
	1.2	0.9	

Preferences (1)



<u>Preference</u>	<u>Overall Preference</u>	<u>Numbers</u>	<u>Decimal Numbers</u>	<u>Dates</u>	<u>Drop-Down Lists</u>
<i>Speech</i>	10	11	6	11	13
<i>Stylus</i>	8	5	12	4	2
<i>No Preference</i>		2	0	3	3

Note: The next page contains detailed results that produce the above chart and table.

Preferences (2)

P #	Group	Overall Preference Stylus	Overall Preference Speech	Overall Preference	Numbers	Decimal Numbers	Dates	Drop-Down Lists
1	2	14	18	1	1	1	1	1
2	3	14	13	2	2	2	1	1
3	2	18	15	2	2	2	1	1
4	3	11	15	1	1	1	0	1
5	3	8	14	1	1	2	1	1
6	2	17	13	2	1	2	2	1
7	2	17	5	2	2	2	2	1
8	3	19	14	2	2	2	2	1
9	1	12	15	1	1	1	0	0
10	1		18	1	1	2	1	1
11	2	7	14	1	1	1	1	2
12	1	15	18	1	1	2	1	1
13	2	3	16	1	1	1	1	1
14	1	17	15	2	2	2	0	0
15	1	11	18	1	1	1	1	1
16	1	15	18	1	0	2	1	1
17	3	13	10	2	1	2	1	2
18	3	20	11	2	0	2	2	0
		14	14					

Appendix E: Research Paper Accepted

As a result of this experimental study, we have submitted a paper to a conference in Montreal which has been accepted. Information regarding this paper can be found by the reference:

Kondratova, I., Lumsden J., and Langton, N., (2006), *Multimodal Field Data Entry: Performance and Usability Issues*, to appear in Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering, Montreal, Canada, June 14-16

Appendix F: Honours Thesis Schedule/Timetable

UNIVERSITY OF NEW BRUNSWICK
FACULTY OF COMPUTER SCIENCE
CS4997 SUMMARY SHEET
Fall/Winter Session 2005-2006

STUDENT NAME: Nathan Seth Langton

STUDENT SIGNATURE: _____

ID #: 3103085

E-MAIL: nathan.langton@unb.ca

PHONE: (506) 444 – 0486 (NRC)

(506) 474 – 0150 (Home)

THESIS TITLE: Multimodal Field Data Entry System: Usability Testing

SUPERVISOR: Dr. Irina Kondratova, Dr. Joanna Lumsden and Professor Natalie Webber

DATE SUBMITTED: Thursday, November 03, 2005

PHASE TITLE	ESTIMATE		ACTUAL	
	PERSON-HOURS	COMPLETION DATE	PERSON-HOURS	COMPLETION DATE
1. Design & Implement Usability Study	45	16 January 2006	150	08 February 2006
2. Prepare & Apply for Ethics Approval	10	20 January 2006	20	30 January 2006
3. Conduct Usability Study	40	22 February 2006	30	15 February 2006
4. Analyse Data	15	01 March 2006	5	05 March 2006
5. Write Report	40	29 March 2006	73	29 March 2006
6. Prepare Presentation	10	24 March 2006	12	24 March 2006
	Total: 160		Total: 290*	

* High even though it's a conservative estimation.