AN INVESTIGATION OF EYE GAZE VERSUS MANUAL INPUT FOR TARGET SELECTION

by

JEFFREY STEVEN SHELL

A thesis submitted to the School of Computing
in conformity with the requirements for
the degree of Master of Science

Queen's University
Kingston, Ontario, Canada
November, 2006

Your file Votre référence
ISBN: 978-0-494-26531-4
Our file Notre référence
ISBN: 978-0-494-26531-4

Abstract

A growing number of appliances and digital devices in the environment compete for attention. As a result, people are often interrupted by new tasks before they have completed the current ones, creating an interaction overload. Initiatives such as Attentive User Interfaces (AUIs) address this problem by filtering interruptions based on their present usefulness to the user. This is accomplished by inferring activity and interest using sensors that monitor cues such as proximity and looking behavior. We present a controlled experiment that compares the eye tracking and keyboard input modalities using a selection task. The task is an abstraction of the real world interaction overload scenario, where a person may be simultaneously confronted by many requests for attention of variable importance. Importance was conveyed by a colour coding scheme, which dictated the order in which targets could be acquired. As the complexity of the selection decision increased by way of a greater number of targets and priority options, eye tracking continued to perform better than manual, keyboard input. In addition, results indicated that subjects may have employed a more efficient planning process when using eye gaze as opposed to manual input. This suggests that outside the confines of the desktop scenario, where the complexity of day-to-day activities and virtual connections intersect, eye gaze may be an appropriate modality to select devices. Subjects in the experiment also preferred, and deemed eye tracking input to be faster, easier, and more accurate than the keyboard only configuration.

Acknowledgements

My Queen's computer science career began in a class taught by Dr. Robin Dawes. Like many students, I was inspired by Dr. Dawes to think differently and to think better. Over the last 10 years you have been my friend, teacher and mentor, always doing your best to point me in a direction that would lead to happiness, and I will forever appreciate it. Beyond measure, your teaching legacy has been showing your students how to love ideas, which is the greatest gift a student can receive.

I would like to thank Debby Robertson and Debbie Rashotte, for going the extra mile to keep me registered, which often required a Herculean effort. I would like to thank all of the students that I have had the pleasure of working with at the Human Media Lab. Thinking back to our humble beginnings, it is amazing what we have been able to accomplish together - when we did great things, it was always together. In particular, Alexander Skaburskis who for the last 5 years has been the sounding board for my life, and Connor Dickie, who I will one day proudly tell people I meet that I once went to school with Connor Dickie.

A very special thanks to Bill Surphlis, who has been exceedingly supportive of every version of me that I have exposed him to. You are truly a great leader and a caring person. I'm sure if I had the talent could write a best seller based on the lessons that I've learned through working with you.

Dr. Roel Vertegaal, we've come around the mountain. Nobody in the world can bring greatness out of me like you can. I look forward to our last journal article, our last soapbox rant and maybe one last hack at New Orleans.

Coming back and completing this 5 year process has been a very humbling experience. No matter how successful I was outside of school, not a day went by where deep down part of me didn't feel like a failure. In many ways it has been a real barrier to me moving forward in life, and in even more ways I have used it as my own barrier. I am now finally ready.

In closing, I would like to dedicate this work to my family. Without you, there would be nothing. To Baba for taking my frantic calls telling her how much I had to do with no time to go, and always rewarding me with an emphatic "You can do it!" You were right Baba, although it may take me a little while before I remember that I don't need to hang my head anymore. Special thanks to Sarah for all of your editing help. Uncle Benny, I will always remember how excited you were to hear that I got accepted into the Queen's MSc. Program and had a chance to follow my dreams – this provided constant motivation for me to get back to the laptop and try and figure out where I was going.

We should always strive to be who we wanted to be, and not let other people's visions of our lives get in the way. The path that we chose is our own to live and if we don't like where it is heading, there's always time for a detour.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Introduction

Computers have changed. From massive mainframes, to everyday appliances, computers are growing in capability, complexity and numbers. Futurists such as Donald Norman [35], William Buxton [4] and Mark Weiser [57] envisioned computers becoming ubiquitous in the environment, woven into the fabric of everyday life. In many senses they have, and in most senses they are heading in that direction. One would be hard pressed to find someone actively thinking about the computer that is hiding in their wristwatch - that computer is invisible, unless it stops working. In many applications computers have transparently replaced their mechanical forbearers.

Norman and Buxton believed that with the way technology was evolving, there was a danger that new devices would be designed using a generic one size fits all approach. Like a Swiss army knife, devices would be adequate for many task scenarios, but great at none of them. Instead, they believed that regardless of the features included, a device such as a digital telephone, should still look, feel and operate like a great telephone. Simply put, features should not trump functionality. Yet in many cases, technology allows us to see beyond the restrictions of the tool, to the overall objective of the artifact. A telephone no longer needs to be viewed as simply a talking tool; it is now a communication device that can also relay text, email and even be used to surf the web.

Like modern cellular phones, computer technology allows us to do new things and to think of existing things differently. Using new communication devices, we can now be perpetually connected to friends, workers, and social networks loosely threaded by common interest. The dedicated 9am to 5pm white collar work day is a thing of the past as new technology and usage behaviours are challenging the old limitations of distance and time [7]. However, by replacing the land line, physical office and secretary with portable gadgets, we are now connected with no filter. Consequently, we are always vulnerable to interruption, which is the topic that we will focus on in this thesis.

1.2 Attention Fragmentation – Bark Versus Byte

To make sure that communication gadgets have our attention, they might use beeps, rings, vibrations or visual distracters. Receiving a notification whenever information is transmitted is analogous to a dog barking every time someone has come to the door, except that the interruption caused by the dog is likely to be appropriate as someone coming to the door generally takes priority over other events. Unlike devices, the dog may understand whether its notification has been received based on the verbal and non verbal cues provided by its master, such as the voice intonation and physical movement. Depending on its demeanor and intelligence, it may stop barking or continue until an overt action like opening a door has taken place. Most people prefer dogs that immediately stop barking.

Unlike dogs, most devices have no reliable, hands free way of determining whether their message has been received. This is the reason why an explicit action, like pressing a physical button to confirm receipt may be required to get the device to stop barking.

Expanding this analogy to a household, it is also useful, for example, to know when someone has stopped the dishwasher prematurely, when the laundry is done or when food is cooked. The reason why we do not train dogs to bark whenever these events occur is because it could disrupt other useful activities and be very, very annoying. The only one of these situations that might warrant immediate human attention is when food is overcooked, for taste and safety reasons. However, if the timing mechanism has determined that the food is cooked, it can reduce the heat, managing the potential for catastrophe on its own. This sensible activity would eliminate the need for immediate human intervention, allowing for fewer interruptions.

Often by considering the scenario in which it is used, a device can be designed to be silently as effective as a device that produces an audio notification, while having the considerable advantage of not disturbing current activity. Achieving this generally requires a coordination effort with other devices, and an awareness of the environment. This is why it is important that interaction management systems, or virtual secretaries, have a detailed understanding of the activity that is taking place in the environment, and the ability to put events in a suitable context. Although these systems aspire to fill the vacant role of the reasonable secretary who takes a message or interrupts activity as it is appropriate, given the current state of popular technology, great gains can be achieved if we can reach the lofty heights of a well trained pug.

1.3 Paying Attention to Attention

The goal of interaction management systems such as Attentive User Interfaces (AUIs) is to have devices, and remote individuals identify and respond to a person's distribution of attention, and use it to as a basis to gage the priority of notifications. AUIs, more formally introduced in [40] primarily use attention for two purposes: as a protocol for turn taking between humans and groups of computers, and as disambiguation criterion for systems that receive multimodal input [41].

Towards achieving these goals, devices obtain an awareness of the environment using sensors such as those that detect eye gaze and monitor speech. Collecting, interpreting and disseminating this information can allow computers and people communicating through computers to come up with an informed assumption about what the recipient is doing. This can be used to venture whether the recipient is available for interruption and if so, how the message can be communicated in a way that causes minimal disruption.

Under this view, human attention is considered to be a competitive resource that is exchanged as an empowering commodity, allowing ‘conversants’ to speak and be heard, rather than intrude at their own behest. This process draws heavily on a wider analogy of how humans cope with information in the real world. Humans use perceptual channels to filter out detail, efficiently allocating finite resources in order to navigate highly complex environments. Consider a person walking through a park. The amount of information that a person is presented with in the form of sights, sounds, smells, and physical sensation is significantly greater than the quantity of information a person would encounter by using virtually any combination of popular technology artifacts. Clearly, it is not the amount of information that people are presented with that creates the problem, it is the form of the presentation [57].

The key difference in the Attentive User Interface approach is that rather than forwarding the arbitrary sequencing of interruptions caused by day to day events, a filtration process is applied which is designed to be complementary to how humans exchange the floor when communicating in group situations. Just as humans exchange up to eight non verbal cues to regulate turn taking [45], AUIs monitor non-verbal attentional channels, such as eye gaze, to determine when, whether and how to communicate with a user in a non-intrusive manner. Ideally devices, like people, will first signal their intention, and then wait for interest signified by eye gaze before taking a turn, allowing human device communication to become efficient, reliable and sociable.

Figure 1-1 *When nobody is watching it The Attentive Television pauses its feed. When the audience looks back at the television, the program resumes.*

Figure 1-2. *AuraLamp only responds to spoken commands when it receives eye contact.*

1.4 Prototype Systems - It's in the Eyes

Prototype systems have been created to demonstrate specific benefits of recognizing attentional input and appropriately applying it. Eye gaze is often used because of the relative simplicity of detecting it, and because it indicates what or whom a person is interested in [22, 33, 45]. The Noncommand User Interface [33] was the first theoretical perspective to discuss the benefits of the eye gaze input channel. Noncommand User Interfaces center in on what the audience is interested in without explicit instruction. For example, The Little Prince system [48] tells a story about the planet where The Little Prince lives. If the audience visually studies a particular element of the planet, then the virtual story teller adapts the topic of the story to speak about the feature in detail. If the audience demonstrates unfocused looking behaviour, then the tale is about the planet in general. Without explicit commands, the story adapts to accommodate the audience's interest.

Similar to virtual environments like the Little Prince, 'Wizard of Oz' Experiments have demonstrated that in physical environments augmented with speech recognition capability, people naturally look at devices that they are interested in before speaking commands [27, 37]. Simply put, we tend to look at what we are interested in. A number of devices have been created to exploit this property, including the Attentive Television [41, 43], EyeBlog [10, 44] AuraLamp [41, 43], EyeProxy [19], SUITOR [26], Attentive Messenging System (AMS) [44], eye-aRe [39] and the Attentive Cell Phone [44, 53]. The Attentive Television (Figure 1-1) plays when people are watching it, and otherwise pauses its feed. AuraLamp, seen in Figure 1-2, responds to speech commands only when it receives eye contact, reducing the number of false positives generated by unrelated chatter; and the Attentive Cell Phone only disrupts face to face conversations when there is an important incoming call, waiting for a break in the conversation before ringing [44].

1.5 Background Summary

As can be seen from these examples, prototype devices that effectively utilize attention as context to aid in the interpretation of events have the potential to improve interactions with technology. Attentive User Interfaces strive to recognize and respond to the nonverbal cues that humans constantly exchange. These cues may allow us to efficiently communicate over many channels simultaneously, indicate turn taking in group conversation [45] and fill in the blanks in statements such as "put that there" [1]. From a notification processing perspective, AUIs are meant to act as a polite, effective secretary, keeping unimportant detail in the periphery, while promoting important notifications to the focus of our attention space [40].

1.6 Contribution

Although there are multiple theories espousing the value of attention and context based approaches, and examples of systems that demonstrate specific benefits, there are few examples of controlled experiments that investigate eye gaze as an input mechanism in a manner that could scale to real world applications. Research has tended to focus on simple tasks, and compared eye gaze to alternate input mechanisms that are not appropriate outside the confines of the desktop scenario.

This thesis presents an empirical study that aims to improve our understanding of the applicability of eye gaze as a modality to interact with users in complex scenarios. Eye gaze was compared to a manual keyboard option that functionally resembles a remote control. We chose a task that simulates a subject responding to competing requests for attention of differing priority. The experimental goal was to address the tasks in the correct order of importance. Secondarily, we investigated whether the reaction time for the first decision in this composite task was proportional to the entire task time. This was done with the intention of determining whether subjects employed a different planning process when using eye gaze, as opposed to manual keyboard input when presented with different stimuli. The central contribution is to improve our knowledge of how eye gaze compares to manual input, particularly when the task becomes more difficult. The results indicate that when using eye gaze, subjects performed better than using the manual option, and engaged in a more effective planning process. We believe that this will encourage further research towards multimodal interface design featuring eye gaze input.

2 Related Work

2.1 Introduction

There is a variety of research on eye gaze in the sciences, particularly in the field of psychology. From uncovering relationships between eye movements and learning [21], to developing a platform to validate a rehabilitative technique for autistic children [38], to examining where expert radiologists look when reading mammograms [29], there is a lot to be gained from studying looking behaviour. A more detailed summary of the broad applications of eye gaze research can be found in Duchowski [12], and in the proceedings of the biannual ACM symposium on Eye Tracking Research and Applications (ETRA) [13].

In the HCI community, eye gaze has been studied as an input mechanism for over 25 years [1]. The focus has been on understanding the conditions where eye gaze can be effectively integrated into the desktop scenario [13, 22, 34, 46, 55, 60], analyzing where subjects look given certain stimuli [21, 47, 59], compressing images based on where subjects don't look [16, 33, 48, 52] and developing interfaces that improve the ability of physically disabled people to communicate [7, 17,20, 28]. In this chapter, we will begin by describing a well known approach to predicting response time for target acquisition tasks. This will be followed by a review of empirical research which evaluates eye gaze for HCI applications. To further motivate the experiment reported in this paper, this discussion will be restricted to evaluations of eye gaze as a selection modality, and concluded with the justification and description of some of the design decisions we made in our experiment.

2.2 Predicting Selection Time

In [6], in regards to predicting the amount of time to fulfill a task, the authors segregated task completion time into two components. The time to acquire a task, and the time to execute the task. Acquisition is the process of building a mental representation of the task, and execution involves invoking the selection.

$$T_{task} = T_{acquire} + T_{execute}$$

Using the language of this experiment, acquisition time is the time to recognize the stimulus and map the solution to the correct key press, and execution time is the amount of time that it takes to press the key. Acquisition can in part be modeled by Hick's law [19] as discussed in [5]. Hick's law predicts the decision time based on the number of equally probable selection options:

$$Decision\ Time = b * \log_2(n + 1)$$

For n equally probable alternatives, b is a constant that can be determined empirically, which is often approximated by 150 ms [6].

The time to execute the decision is simply the amount of time that it takes to press a key. For an expert typist who works at 135 words per minute this translates to 80 ms for a key press; for an average skilled typist who works at 55 words per minute, this corresponds to

200 ms [6]. All subjects were touch typists, so we use the range of 80-200 ms to approximate execution time.

For each target requiring a key press, the selection time can therefore be modeled by:

$$T_{task} = b * \log_2(n + 1) + T_{execute}$$

Where $b = 150$ ms and $T_{execute} = 80 \sim 200$ ms

2.3 Hick's Law Limitations

Looking specifically at modeling selection using eye gaze, Kveraga et al [23] demonstrated that response time is not governed by Hick's law when the task is selecting a target using eye gaze. Selection, in the experiment was made by a visual saccade. Despite varying the number of response targets from 1 to 8, the mean response time remained in the order of 200 ms. The number of alternatives did not affect response time because, as discussed in the feature-integration theory of attention, target recognition using eye gaze occurs quickly, automatically and in parallel across the visual field [48]. Selection using eye gaze is a preattentive task, where subjects do not need to be cognitively aware of the search, or of expressing the activity through a secondary channel like the hands – they just look at the target. The classic example of preattentive selection is the pop-out effect, where an item that is distinct by a low-level feature such as color immediately triggers eye movement [49]. Broadbent [3] and Cherry [8] showed that a similar principle applies to the auditory system.

To contrast this effect, Kveraga et al [23] included two other conditions in this experiment. In one condition, a subject had to deliberately look in the opposite direction of the target, and in the other, selection was made using manual key presses. In both of these conditions, subjects had to decide how to map the selection to a response, and Hick's law appropriately applied as this is the situation that it was designed to model.

In the experiment reported in this thesis, selecting the target using eye gaze is a preattentive task which can not be modeled using Hick's law. However, Hick's law does apply when using the manual condition and when specifying priority, as responses had to be mapped to a key. We believe that Hick's law is not a suitable methodology to use to model this experiment because it does not apply equally to all of the experimental conditions.

Although we did not use this approach, formulating and validating a predictive model of response time has the potential to show insight into the hidden mechanics of selection. In the next section we will look at experimental research geared towards testing eye gaze in simple selection scenarios. The task in these experiments is to choose a target using one of a number of input devices and mechanics. Described in chronological order, as the technology improved, researchers were able to better test the concept as opposed to the technology, but a well defined performance model has yet to be shown.

2.4 HCI Experimental Research

Ware and Mikaelian [56] compared three different modes of selecting targets using eye gaze:

1) Prolonged fixation, otherwise known as gaze dwell time
2) Gaze together with a hardware button press
3) Prolonged fixation followed by a fixating on a virtual on-screen button

All three of these input techniques demonstrated faster response time than a traditional mouse as reported in [5]. Fixation using the virtual on-screen button was the slowest among the eye gaze conditions. Dwell time was slower than the hardware press condition, but more accurate. The authors conclude that if the price of hardware is unimportant, the user's hands are occupied, and targets are large and widely spaced apart, eye gaze may be best selection option.

Sibert and Jacob [46] were able to demonstrate a statistically significant benefit to using eye gaze over an alternative input mechanism. In two experiments, they compared eye gaze dwell time to mouse selection. Similar to [23], the first task reported was to select a coloured circle from a group of non-coloured circles. Eye gaze was 85% faster than mouse selection. In the second experiment reported, the task was to select the letter shown on the monitor that corresponded to the letter that was spoken to the subject. Eye gaze was 31% faster than the mouse. The gap between the mean scores narrowed in both absolute and relative terms as the cognitive load of the selection increased.

Fono and Vertegaal [13] were also able to isolate a compelling performance advantage to using eye gaze for a selection task. In a Window selection task, eye gaze selection (fixation + key press) outperformed manual options (mouse and keyboard) by an average of 72%. Of particular note to this paper, the authors stated that the differences in performance between the eye fixation modality and the gaze and key press option may be accounted for by the amount of time that it took to press a key. This implies that recognizing a target and pressing the key (spacebar) is performed in serial sequence, even when the subject's hands could be positioned on the single activation key prior to stimulus presentation.

As can be seen by these experiments, eye gaze has been demonstrated to be faster than mouse and manual input for selection tasks where the acquisition component is simple compared to execution [13, 23, 46, 56]. However, a different input modality is more than a channel to communicate the same information to a computer. Eye gaze affords an opportunity to change the way that people interact with technology. The next section reviews two experiments that do not show significant statistical improvements, but forward ideas that have the potential to enrich the human computer dialogue.

2.5 Eye Gaze to Test Concepts

Wang, Zhai and Su [55] evaluated the Eye Assisted Selection and Entry (EASE) system for Pinyin character entry. Pinyin is a phonetic representation of Chinese characters using Roman characters. Subjects had to select the Chinese character that corresponded to the meaning of the term displayed in Pinyin. This conversion is more than mechanical as Chinese is not an alphabetic language and many Chinese characters are homophonic with many others. Integrating eye gaze provided an opportunity to improve the process of translation. To choose the most appropriate subjects had to read all of the options before making a selection. Selections were made one of two ways: by looking at the target and pressing a key, or by typing a character that corresponded to the position of the targets. This experiment did not show a statistically significant difference between using EASE over a keystroke only method [55]. The authors believe that this is mainly because in the complex task, the decision time dominated the selection time, obscuring the effect of the input device.

Zhai, Morimoto and Ihde developed Manual and Gaze Input Cascaded (MAGIC) pointing, a non traditional approach to using eye tracking in a standard desktop scenario [60]. MAGIC pointing combines both eye gaze and mouse movements as an alternate to traditional mouse interaction. Eye gaze is used to identify the cursor's starting position (context), and mouse movements are used for precise motor control (focus). The evaluation of the MAGIC system showed mixed results. The 'liberal' implementation of the system, where the cursor followed the subject's eye gaze, slightly outperformed the manual option. Because a mouse pointer following the subject's eye gaze would be distracting, it can be deemed to be impractical in real scenarios. The more practical 'conservative' implementation, where the cursor was not visible until there was mouse movement was slower than the manual alternative.

Although the experimental results were not compelling, the MAGIC system showed how eye tracking and mouse movements can be combined in a multimodal interface to potentially improve usability. The MAGIC model can be understood as eye tracking specifying the region of interest (attention) while the mouse movement conveys the subject's intention within that region.

The attention / intention or focus / context dualities are fundamental to how AUIs function. Just like people, AUIs recognize that the focus of nonverbal perceptual channels implies attention, which is used as an empowering and interpretive vehicle for communication. As demonstrated by experiments and prototypes, eye gaze is a faster selection modality than the traditional mouse or keyboard alternative in many circumstances. Eye Gaze recognition technologies are also improving and hold the considerable advantage of requiring less explicit input from the user.

The rest of this chapter is devoted to motivating the design decisions and differentiating the experiment reported in this thesis from the other experiments that were reviewed in this chapter.

2.6 Dwell Time Limitations

Unlike [13, 23, 46], we chose not to include gaze dwell time in our experiment because there is a strict limitation to the amount of information that it can convey. Either someone fixates at a target, or does not, there are only two options. Without additional information, it is difficult to attribute meaning outside of very specific circumstances. Also, because eyes are principally an input device for people, a system that only uses dwell time would interpret natural eye movements as a catalyst for action. This has been dubbed the Midas touch effect, where everything a user looks at gets (often inappropriately) activated [22]. Eye gaze is most effectively used as context to support the interpretation of intentional acts (i.e. key presses, speech commands) as found in many AUIs [10, 40, 41, 42, 43, 44, 48].

2.7 Eye-Trigger Selection Limitations

In [13, 55, 56], a subject had to hit an action key to indicate that she was looking at the target. With only one action key to specify intention, the selection is equivalent to a fixation with a trigger, or a universal remote with only an on / off button. To be able to generalize to interacting with groups of computers in an unconstrained environment, more information would need to be conveyed. At the least, the *'trigger'* would need to offer more than one option, analogous to a limited function remote control. Thus making a decision requires consideration of not only the stimulus, but how to express the stimulus using a keypad. We believe this additional step is fundamental to be able to generalize as to how eye gaze can scale to more complex scenarios. This is the primary reason why we included more than one selection option as opposed to a simple trigger in our experiment. Colour was used to attribute priorities to targets which specified the order in which they could be acquired.

2.8 Summary and Statement of Uniqueness

There are four characteristics that made this experiment unique.

1. **Prioritization Scheme –** The colour coded prioritization is meant to be a 'Wizard of Oz' rendition of how an AUI may present requests for attention of different priority. It can be argued that the experimental task, further described in section 3.4, of choosing green targets over yellow ones is a decision far more trivial than

the application we are generalizing to. However, we hope that the presentation is viewed as malleable.

2. **Selection involves disambiguation** - We used both eye gaze (attention) and the keyboard (intention) together to specify action, which is an abstraction of the process utilized by some AUIs. As described in section 2.7, we believe an action key to specify intention is too simplistic to generalize to the experience of using computers in unconstrained environments.

3. **Complex task with Many Activities per Trial** – Real world selection is not simply mechanical. By looking at trials that may involve planning, prioritizing and multiple selections, versus just looking at a single selection task, we get a richer insight into how eye gaze compares with manual input. This approach allows us to see how performance is affected when there are many paths to complete a task. The relationship between complexity and performance may be most relevant to understanding applications in multimodal, unconstrained scenarios.

4. **Understanding the Decision Process** – As further discussed in chapter 7, we isolated the first of many selections and compared it to the total task completion time. If the time to make the first selection is not proportional to the entire trial time, then it is evidence that subjects are engaging in a planning process before issuing their selections. If this consistently varies by experimental conditions, particularly by input mechanism, then we will conclude that the conditions affect how subjects plan. Different planning processes may be more or less attuned to

difficulty, providing insight into how these conditions may scale to unencumbered scenarios.

3 Task Description & Hypotheses

3.1 Introduction

In this chapter, we will describe the main experimental task which is an evaluation of eye gaze versus manual keyboard input in a target selection task of varying complexity. We will begin by discussing a scenario that illustrates the purpose and the primary application of the experiment. This will be followed by a discussion of the experimental task, a formal introduction of the experimental variables, and a statement of our hypotheses.

3.2 Illustrative Scenario

Jonathan has about twenty appliances and digital devices in his apartment, which include his traditional appliances (microwave, dishwasher, oven, refrigerator etc), his general electronic devices (VCR, television, stereo, DVD Player, alarm clock), his house controls (thermostat, security alarm, light switches) and his communication devices (telephones, fax machine, desktop, laptop, mobile telephone, PDA). Jonathan can purchase one of three wireless remote controls to centralize communication with all of the devices. This will reduce the number of times that he has to get up and make physical contact with devices.

3.2.1 Option #1: Manual Parallel Input – Overload

The first remote system, which we will call called Overload, has a button for each function provided by every device in the environment. Let us assume that there is an average of five important functions per device. To accommodate the twenty devices in his house, this remote would require about 100 keys, still less than the number of keys found in the keyboard of the average desktop computer, but significantly more than are found in a typical remote control. The number of keys required in the Overload unit is modeled by the following equation:

*Overload: Required Keys = Number of Devices * Number of Functions*

If Jonathan adds a new device or replaces existing devices, the number of buttons can very quickly get out of hand. This is somewhat restricted because Overload only has 120 buttons, which limits his ability to add new devices and functions. In the case that he does add a device or change a function, to maintain a logical organization scheme, he may have to move or relabel buttons. Jonathan was hoping to use the same remote to interact with technology at his cottage, but given these restrictions he will likely have to buy two Overload packages.

Graphic	+	-	▶	■	❚❚
Device Independent Meaning	More	Less	Start/Play	End	Delay
	Up	Down	On	Off	Later
	Harder	Softer	Resume	Quit	Freeze
	Louder	Quieter	Confirm	Deny	Mute
	Faster	Slower	Go	Stop	Wait
	Bigger	Smaller	Answer	Hang Up	Hold
	Closer	Further	Select	No	Maybe
	Zoom In	Zoom Out	Open	Close	Neutral
	Next	Previous	Execute	Lock	Pause

Figure 3-1. *Potential meanings of generic symbols*

3.2.2 Option #2: Manual Serial Input – TwoSteps

The second option, which we'll call TwoSteps, allows Jonathan to press one of twenty buttons to select the device of interest, followed by one of five options to tell the device what Jonathan wants done. Using only 25 keys, TwoSteps provides functionality comparable to the Overload System. This is because devices have similar basic behaviors that can be modeled using generic symbols. As illustrated in Figure 3-1, these symbols have an obvious meaning in most situations. The number of keys required in the TwoSteps unit is modeled by the following equation:

TwoSteps: Required Keys = Number of Devices + Number of Functions

Jonathan is excited to learn that he could greatly increase the number of keys to accommodate a very aggressive number of action options without even approaching the number of keys in the Overload model. Additionally, the buttons can be simply organized in a well thought-out manner, perhaps mimicking the physical location of the devices in the environment, similar to the light switch panel in [36]. However, this may constrain adding new devices, so like Overload, he may occasionally have to relabel buttons to keep the remote organized.

3.2.3 Option #3: Eye Gaze Serial Input - Oz

The third alternative is for Jonathan to buy a new remote called Oz that only sends the commands from the remote control to the product that he is looking at. Like the TwoSteps system, Oz uses generic symbols to convey basic behaviours. To effectively use Oz, he will have to memorize the location of the buttons and what they mean so he can keep his gaze on the target instead of the remote. Because there are only between five and ten buttons on the remote, Jonathan does not consider this daunting. The number of keys required in the Oz unit is modeled by the following equation:

Oz: Required Keys = Number of Functions

3.2.4 Scenario Summary

This scenario illustrated what this experiment is intended to model – a person communicating with technology using one of various remote controls. The TwoSteps approach is represented by manual input in the experiment, where a subject had to press a button to signal the target followed by a disambiguating second key press to signify what should be done with the target. In the TwoSteps illustration above, it could be lowering the volume or turning off a television. In the experiment, it simply indicates a property of the target. The Oz approach is represented by eye gaze, where the subject looks at the appropriate target and simultaneously presses a button indicating a property of the target. Due to its impracticality, we did not include the Overload option in this experiment.

3.3 Application to Attentive User Interfaces

In the experiment the key press disambiguated colour, conveying the priority of the target. The priority classification was chosen to mimic the ordering of interruptions that attention management systems like AUIs produce. Attentive User Interfaces try to filter diverse events that require attention into a few simple and suitable notifications. Unlike the experiment, if a typical laundry cycle has completed, an AUI may not actually flash a coloured light, it would find an appropriate, not necessarily interruptive way to convey its message. If something of greater importance is occupying the user's attention, then the system would wait until the task is completed before telling the user that the laundry is done. If nothing of higher priority is happening, then the laundry notification would be communicated immediately. The goal is to keep the user informed of the most pressing event occurring, or a small set of events of roughly equivalent importance.

There are two forms of communication between the AUI and the User. The AUI system samples the user's activities to determine both his or her availability and the least interruptive mode of notification. In this sense, the AUI is passively watching and formulating a model of the user's state, and filtering incoming requests for attention. At the same time, the infrastructure is also actively watching the user, awaiting commands that he or she may implicitly or explicitly issue. In a perfect AUI scenario, a user would have the benefit of looking at a part of the room representing a light switch, say off, and have the lights turn off [41], or looking at a fridge to ask what meals can be created given its inventory [42]. As the active infrastructure facilitates this hands free computing, the passive AUI adds the events to the pool of knowledge that it draws on in performing the role of the virtual secretary, keeping the user focused on the most pressing tasks at hand. Without the ability to communicate acceptance or rejection of a notification in a manner similar, but simpler to acknowledging a barking dog, AUIs will still demand disruptive physical interaction with the various devices and appliances in the environment, just less frequently.

The experiment in this thesis is more attuned to understanding how effective eye gaze may be as a part of a multimodal hands free selection and notification acknowledgment system. However, results remain implicitly linked to the more significant goal of understanding and promoting the virtual secretary, who interacts with the user via hands free computing.

If we consider the more general problem of interacting with technology in the environment, the experimental task takes on a different meaning, but the results are still interesting. From this perspective, it is an investigation of the scenario described in section 3.2 of *TwoSteps* versus *Oz*. The task breaks down the two elements of the selection process; choosing the target (attention) and the colour (intention) using a serial selection process.

Experimental Task Algorithm
Stimulus Presentation (1-4 Targets of variable Priority)
For all Active Targets (Green or Yellow)
    Select Target of Highest Priority using Input Mechanism
        If in Eye Gaze Condition
            Fixate on Target
        If in Manual Condition
            Press Key indicating Target (4, 5, 6 or + on the keypad)
        If selection is not highest priority repeat
    Target Selected – Feedback - orange box around Target
    Select Priority
        Press Key indicating Priority (1 or 2 on keypad)
        If incorrect priority is chosen repeat selection
    Target turns red
Repeat until all targets in the Trial have been selected (red)

Figure 3-2. *Experimental Task Algorithm - Italicized Steps indicate User Action, Green Steps indicate success, Red Steps indicate error.*

3.4 Experimental Task

As shown in Figure 3-2, the subject is presented with 1 to 4 active targets per trial. When many targets are simultaneously activated, it serves as a metaphor for situations in which users are confronted by multiple requests for attention. Green targets represent high priority notifications (i.e. an important telephone call) which must be processed immediately, while yellow targets denote lower priority notifications (i.e. laundry cycle has completed) which can only be processed after all of the high priority notifications. Red lights act as a zero priority placeholder that required no action and thus are not considered targets. Figure 4-1 shows a picture of the experiment interface. Targets have to be individually selected in the order dictated by the colour coding. Selection involved choosing the target and then specifying the priority (colour of the light). The order of selection is not important among multiple equal priority notifications. The trial is concluded after all of the targets are selected in the appropriate order.

3.5 Independent Variables

We investigated the effect of three independent variables in our task:

1. Input Mechanism
2. Number of Priorities
3. Number of Targets

3.5.1 Input Mechanisms

There were two selection mechanisms in this experiment.

Figure 3-3. *The keyboard Interface for the Experiment.*

Eye Gaze: the subject selected the target of highest priority by looking at it, and subsequently specified the colour of the target (green or yellow) by pressing one of two keys. As seen in Figure 3-3, subjects pressed 1 on the keypad for a green target, and 2 on the keypad for a yellow target. Unlike the illustration, the green and yellow keys were not coloured for the experiment.

Keyboard Manual: the subject chose the target by pressing one of four keys, corresponding to the target of highest priority, and then specified the colour of the target (green or yellow) by pressing one of two keys. As seen in Figure 3-3, subjects pressed the following keys on the keypad to select targets:

- 4 on the keypad to select target 1
- 5 on the keypad to select target 2
- 6 on the keypad to select target 3
- + on the keypad to select target 4

Subjects pressed 1 on the keypad for a green target, and 2 on the keypad for a yellow target. The subjects positioned their dominant hand on the keyboard at all times.

Although using two hands together may produce faster results, we wanted to mimic the typical one hand operation of a remote control.

3.5.2 Number of Priorities

For each trial, either one or two priority options were activated. Subjects were to select the highest priority targets visible. Targets were assigned priorities randomly.

3.5.3 Number of Targets

Up to four targets were active in a trial. Each target needed to be selected using eye gaze or manual input in order of priority to complete a trial. There were 10 trials for each number of targets. These were presented in random order for each input mechanism.

3.6 Hypotheses

We now present our experimental hypotheses.

1. *In a single placeholder single priority selection task, eye gaze input will be faster than manual input*

This represents the base case for the input mechanism analysis. In a single placeholder selection task, only 1 target gets activated, and the priority of the light is always high (green colour). There is little complexity to the decision process. When the target is activated (light turns green), using eye gaze the subject will look at the target, and press 1 on the keypad to indicate high priority. Using the manual option, the subject will still have to look at the screen to notice the activation of the target, and press a button to select the target (4 on the keypad) followed by a second button to indicate the priority of the target (1 on the keypad). The only difference should be the extra key press required in the manual condition, and the extra key press should take between 80-200 ms [6].

The rest of the hypotheses correspond to the main experiment, where up to 4 targets were activated in random priority per trial.

2. *Using Eye Gaze, subjects will have a lower Response Time (RT) than when using Manual input*
    - *RT(EG) < RT(Manual)*

For each target, manual input requires two key presses: the first to indicate the target of the selection, and the second to specify the colour of the target. Eye gaze only requires one key press, indicating the colour of the target. As they both require eye contact to recognize the activation of the target, the extra key press required by manual input should add between 80-200 ms [6] to the response time for each additional active target per trial.

3. *Increasing the number of targets that require action per trial increases Response Time:*
    - *RT(n targets) < RT(n+1 targets)* for n = 1 to 3

For each additional target, in the eye gaze condition, 1 key needs to be pressed, and in the manual condition, 2 keys need to be pressed. In either case, adding a target increases the number of key presses per trial, so we expect the response time to grow accordingly.

4. *Increasing the number of priorities (colours) that are active per trial increases Response Time*
    - *RT(1 colour) < RT(2 colours)*

When there are two colours as opposed to one, the order of selection matters. Applying Hick's law [19], this will affect the decision time of each target in a trial by changing the priority selection from b*log(2) to b*log(3), which increases predicted response time by 88 ms.

Interaction effects

5. *Number of Targets x Number of Priorities will not demonstrate a significant interaction effect*

Each additional target adds 1 to 2 key presses (1 for eye gaze, 2 for manual input), averaging 1.5 additional key presses. Applying [6] this should add 80-200 ms x 1.5 = 120 – 300 ms, to the response time per target. On the contrary, each additional priority option should only add 88 ms to the response time per target [19]. The effect of the number of targets will dominate the effect of the number of priorities, obscuring a possible interaction effect.

6. *Number of Priorities x Input Mechanism will not demonstrate a significant interaction effect*

Adding a priority will affect the input mechanisms equally. Varying the number of priorities in the task does not add or remove a key press, for both input mechanisms it just prescribes the order of selection.

7. *Number of Targets x Input Mechanism will demonstrate a significant interaction effect*

Adding a target per trial increases the number of key presses by 2 for the manual option, and only by 1 for the eye gaze option, so the number of targets should affect input mechanisms differently.

8. *Number of Lights x Number of Colours x Input Mechanism will demonstrate a significant interaction effect*

Increasing the complexity of the decision process will favor the input mechanism that requires less work. In the manual keyboard condition, increasing the number of targets adds to the number of key presses at double the rate that it does for the eye gaze condition. Manual keyboard input will also be more sensitive to an increase in the number of priorities because more buttons need to be pressed in a trial than in the eye gaze condition, complicating the coordination problem.

3.7 Subjective Evaluation

After completing the experiment, subjects filled out a post experiment questionnaire. The questionnaire consisted of eight Likert-style questions, measuring the subject's preference and perceptions of ease of use, response time and accuracy of the eye gaze and manual input mechanisms. The questionnaire used a 5 point scale in which 1 indicated a strong disagreement and 5 indicated a strong agreement with the associated statements.

3.8 Subjects & Experimental Design

Fourteen subjects participated in this experiment, three of whom were students at the lab. The remaining nine participants were Queen's students who responded to public posters advertising the experiment. These participants represented a roughly even split between science and humanities students. Six of the subjects were female, and eight were male. The average age was 23. They all described themselves as touch typists. Six subjects wore glasses, which were removed for the duration of the experiment to improve eye tracking performance. In none of the cases did this affect their ability to detect or respond to the targets presented. One subject described himself as somewhat colour blind, but upon presentation of the colours in the interface, he had no trouble differentiating between them. Six subjects had some experience using an eye tracker prior to the experiment. A within subjects design was used, and each subject was presented with both input device conditions. The order of the conditions was randomized by the experiment software. For each input mechanism, subjects were presented with 50 trials, 10 for each number of active targets (1 to 4) in random order, and 10 for the base case, where subjects were informed in advance that only 1 target would be active in the trial and that it would be of high priority. Active targets were randomly assigned a priority. In total, each subject was presented with 100 trials. Subjects were informed that the objective was to complete the task quickly, but that trials which had errors would be excluded.

3.9 Training

Subjects were presented with training sets of 10 trials for each input mechanism. The task was to select two targets of variable priority. To successfully complete training, subjects had to have completed 8 of the 10 trials without making an error. Errors were defined as trials where a minimum of one key over the required number was pressed. Subjects also had to not have shown a difference of 20% between the average of the first 5 trials and the last 5 trials per set. Upon achieving these objectives, subjects had the option of whether to repeat the training procedure if they felt as though they were not adequately proficient.

4 Implementation

4.1 Apparatus

The main experiment software was run on a Pentium 4 2.6 GHz processor using a 17" flat screen LCD monitor. The software was written in MAX/MSP 4.3 for Windows [29]. A standard USB Macintosh keyboard was used. The only keys that were active in the experiment were on the keypad as described in section 3.5.1. The LC Technologies Eyegaze [24] system handled eye tracking input. It uses a near-infrared LED located at the centre of the camera lens which generates a corneal reflection and a "bright pupil effect". These are used to calculate the gaze point on screen. The camera samples at 30 Hz. The Eyegaze system was set in dual computer mode, where one dedicated computer handles all of the computationally intensive image processing and streams the eye gaze coordinates to the main experiment computer using a 100 Mbps direct Ethernet connection. For every key press in the experiment, data were time stamped and recorded in a text file for subsequent analysis.

4.2 Dwell Time

In the eye gaze condition, subjects needed to fixate on the target in order for it to register as a selection. A fixation was defined as 3 successive samples in the same horizontal region. At the 30 hz frame rate, this introduced a 100 ms delay compared to the manual input mechanism. Key presses denoting priority were not registered until after the target was selected.

4.3 Eye Gaze Calibration

In preparation for the eye gaze condition, subjects were presented with the standard 15 point LC Technologies calibration module, mapping eye gaze coordinates to screen coordinates [24]. All subjects were able to successfully complete the calibration routine.

4.4 Eye Tracking Margin of Error

Within the foveola, it is not possible to detect what in particular the subject is looking at as the region of highest acuity has a 1.3° visual angle [12]. At the 24" distance to the monitor used in the experiment, this maps to ¼ inch radius of error. In addition to this, the eye tracking hardware introduces up to a ½ inch error, bringing the total uncertainty to ¾ of an inch.

4.5 Eye Input Mapping

To address these inherent inaccuracies and simplify the detection of eye movements, the screen was divided into 4 separate horizontal regions. The gaze point was associated with the nearest target along the x axis in screen coordinates, reducing the selection of a target to one dimensional movement. The additional redundancy afforded by only recognizing four regions across one dimension greatly improves the eye tracking accuracy. The horizontal total width of the monitor is 13.3" [10], so each of the four regions was 3.325" wide. In the centre of each region was a 1" placeholder. This afforded a 1.4125" boundary in each direction between adjacent regions. The boundary far exceeded the ¾" radius of error caused by the eye tracking apparatus and the properties of the eye as discussed in section 4.4.

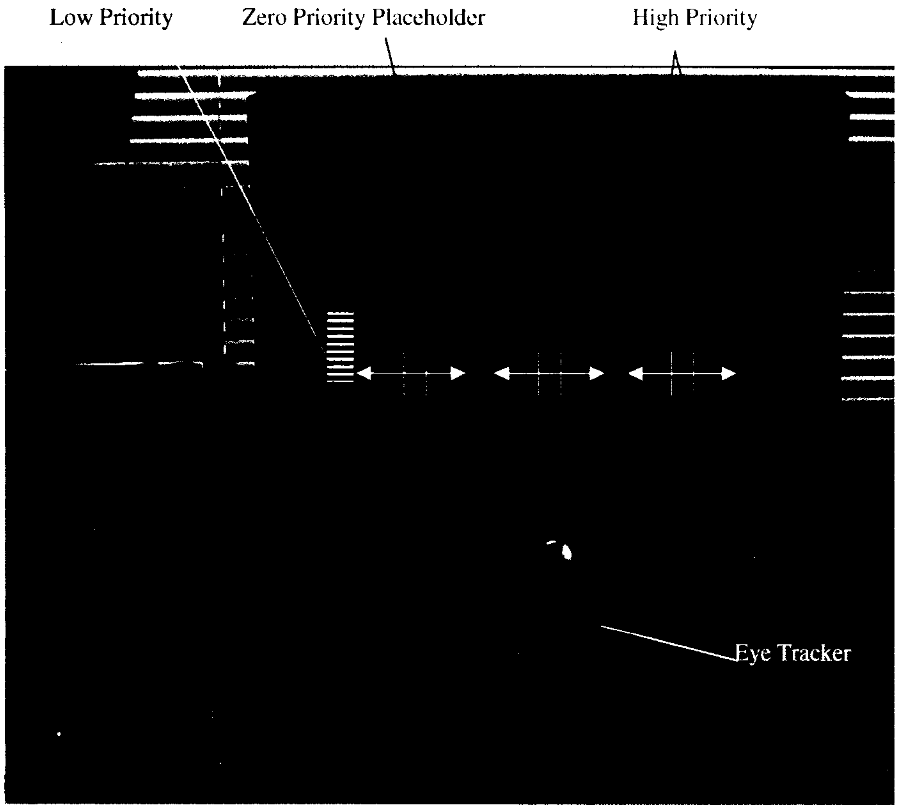

Figure 4-1. *Experiment Interface.*

4.6 Experimental Interface

As described in section 3.4, each placeholder could be red, yellow or green. These colours were chosen to make use of the familiar traffic light analogy which we believe to be a simple way of presenting prioritized options. The green and yellow colours represented targets, while red was used as a placeholder. Because people associate the colour red in the traffic light situation as requiring immediate stopping action, it may have been better to use the colours differently. However, we believe that the extended training process ensured that subjects were comfortable with the meanings of the different colours before the main experimental task began.

We chose two priorities because if we would have used red as a third priority, then we would have required either the use of a fourth colour as a placeholder for inactive target slots which would have complicated the traffic light analogy. We did not use the background colour as a placeholder because it would have confused the spatial relationship between the targets and the keys. For both conditions, to provide feedback conveying that the target was selected, a two pixel wide orange box surrounded the target until the correct colour was chosen.

5 Results

The experiment involved three independent variables; Input Mechanism (2 Levels), Number of Targets (4 Levels) and Number of Priorities (2 Levels). Between and within Subjects effects were analysed using GLM Univariate one way analysis of variance (ANOVA). Significance was evaluated at the α =.05 level. Trials containing errors as defined in section 3.9 were excluded from analysis. Within factors differences between pairs of means were detected using GLM Univariate Pairwise comparisons at the α =.05 level with Bonferroni adjustments for multiple comparisons.

5.1 Base Case - Single Placeholder

As explained in section 3.6, the single placeholder condition is where subjects were presented with only 1 target in 1 possible location. They were instructed that the target would be high priority (green). This tested the mechanics of selection for the input mechanisms.

Base Case - Input Mechanism (ms)			
Input Mechanism	**Mean**	**Std. Error**	**Results**
Eye Gaze	709.71	21.09	= Manual
Manual	705.67	28.82	= Eye Gaze

Table 5-1 *Means and Standard Error for selecting 1 target from 1 placeholder*

Effects of Input Mechanism on Response Time Given 1 Placeholder

Table 5-1 shows the data summary of the response time for the simplest case – 1 high priority target was active per trial from 1 placeholder. There was no significant difference across conditions ($F_{1,220}$ =0.12, p=0.91). As a consequence, we support the null hypothesis that eye gaze is not a faster mechanism than manual input for single placeholder, single target selection.

5.2 Multiple Placeholder – Effects of Independent Variables

For the rest of this chapter, all trials have a variable number of targets and priorities which can occur in 4 locations or placeholders.

Input Mechanism (ms)			
Input Mechanism	**Mean**	**Std. Error**	**Results**
Eye Gaze	1865.14	25.36	≠ Manual
Manual	2156.01	22.96	≠ Eye Gaze

Table 5-2 *Means and Standard Errors for the input mechanism*

5.2.1 Effects of Input Mechanism

Table 5-2 shows the data summary for the input mechanism used in the trial. Analysis of variance showed that there were significant differences across levels ($F_{1,593}$ = 72.32, $p < 0.001$). We reject the null hypothesis that the input mechanism does not have an impact on response time. Mean response time using eye gaze was 16% faster than the Manual condition.

Number of Targets (ms)			
Targets	**Mean**	**Std. Error**	**Results**
1	943.96	25.03	≠ 2 ≠ 3 ≠ 4
2	1698.90	30.97	≠ 1 ≠ 3 ≠ 4
3	2262.31	34.19	≠ 1 ≠ 2 ≠ 4
4	2756.08	37.74	≠ 1 ≠ 2 ≠ 3

Table 5-3 *Means and Standard Errors for the number of active targets per trial*

5.2.2 Effects of Number of Targets

Table 5-3 shows the data summary for the number of targets that required action per trial. Analysis of variance showed that the response times for the number of targets differed significantly across levels ($F_{3,178}$ = 655.36, $p < 0.001$). Post-Hoc comparisons with Bonferroni correction showed all levels to differ significantly from one another. One target was faster than two targets ($p<0.01$), two targets were faster than three targets ($p<0.01$), three targets were faster than four targets ($p<0.01$). We reject the null hypothesis that the number of targets does not have an impact on response time.

Number of Priorities (ms)			
Priority (ies)	**Mean**	**Std. Error**	**Results**
1	1726.14	25.33	≠ 2
2	2360.09	22.02	≠ 1

Table 5-4 ***Means and Standard Errors for the number of priorities represented per trial***

5.2.3 Effects of Number of Priorities

Table 5-4 shows the data summary for the number of priorities that were represented per trial. Analysis of variance showed that there were significant differences across levels ($F_{1,593}$ = 356.76, $p < 0.001$). We reject the null hypothesis that the number of priorities does not have an impact on response time.

5.3 Two-Way Interaction Effects

5.3.1 Effects of Targets * Priorities

Analysis of variance showed that there was not a significant interaction effect between the number of targets and the number of priorities represented per trial ($F_{2,178}$=1.537, p=0.216). We accept the null hypothesis that there was no interaction effect between the number of targets and the number of priorities.

Number of Priorities * Input Mechanism (ms)				
Priorities	Eye Gaze	(S.E.)	Manual	(S.E.)
1	1608.27	37.15	1839.30	34.53
2	2173.38	33.53	2542.25	28.68

Table 5-5 *Means, Standard Errors for Number of Priorities * Input Mechanism*

5.3.2 Effects of Priorities * Input Mechanism

Analysis of variance showed that there was not a significant interaction effect between the number of priorities represented per trial and the input mechanism ($F_{2,178}$=1.537, p=0.216). We accept the null hypothesis that there was no interaction effect between the number of priorities and the input mechanism.

Number of Targets * Input Mechanism						
Targets	Eye Gaze	(S.E.)	Manual	(S.E.)	Delta	Result
1	864.56	35.50	1023.36	35.29	158.80	EG ≠ M
2	1474.95	45.21	1922.86	42.35	447.91	EG ≠ M
3	2104.92	51.46	2402.22	45.59	297.30	EG ≠ M
4	2682.03	56.97	2830.13	49.52	148.10	EH = M

Table 5-6 *Means, Standard Errors and Delta for Number of Targets * Input Mechanism*

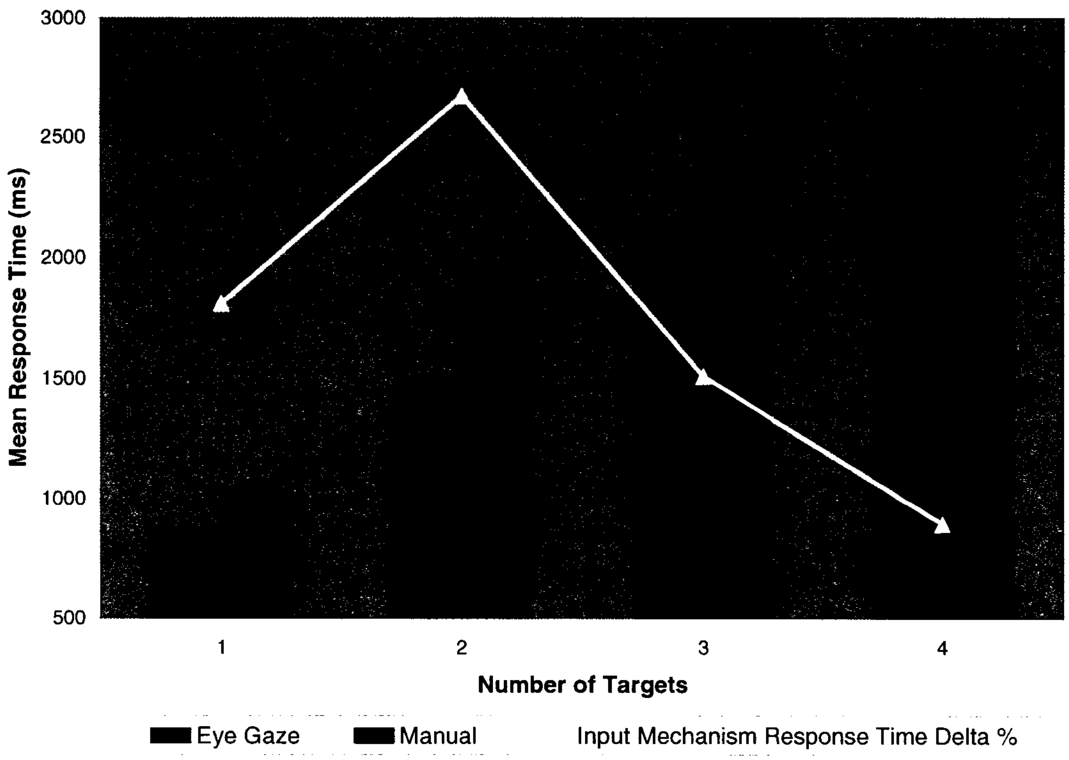

Figure 5-1. *Chart of Means and Input Mechanism Delta % for Number of Targets * Input Mechanism.*

5.3.3 Effects of Targets * Input Mechanism

Table 5-6 and Figure 5-1 show the data summary for the effect between number of targets and the input mechanism used in the trial. Response time did vary significantly across number of targets by input mechanism ($F_{3,178}$ = 3.6, p=0.013). This allows us to reject the null hypothesis that there is no interaction effect.

5.4 Three-Way Interaction Effects

Number of Targets * Number of Priorities * Input Mechanism (ms)						
Targets	Priorities	Eye Gaze	(S.E.)	Manual	(S.E.)	Result
1	1	864.56	35.50	1023.36	35.29	EG ≠ M
2	1	1421.54	66.74	1758.39	57.73	EG ≠ M
	2	1532.46	60.42	2099.97	62.22	EG ≠ M
3	1	2037.20	84.95	2379.58	81.74	EG ≠ M
	2	2162.22	62.13	2423.24	44.36	EG ≠ M
4	1	2531.39	121.08	2453.98	108.56	EG = M
	2	2778.88	51.99	3071.95	41.81	EG ≠ M

Table 5-7 ***Means and Standard Errors for Number of Targets * Number of Priorities * Input Mechanism***

5.4.1 Effects of Targets * Priorities * Input Mechanism

Table 5-7 shows the data summary for the number of targets * number of priorities * input mechanism interaction effect. Response time differences across conditions were significant ($F_{2,178}$ = 4.461, p=0.012). This allows us to reject the null hypothesis that there is no interaction effect.

5.5 Summary of Findings – Significant Differences

Test	Placeholders	Our Hypothesis	Results
Input Mechanism	1	Y	N
Input Mechanism	4	Y	Y
# Targets	4	Y	Y
# Priorities	4	Y	Y
Targets * Priorities	4	N	N
Priorities * Input Mechanism	4	N	N
Targets * Input Mechanism	4	Y	Y
Targets * Priorities * Input Mechanism	4	Y	Y

Table 5-8 ***Summary of Findings***

5.6 Error

In the eye gaze condition, the error rate was 26%. In the manual condition, the error rate was 22%. Although this error level may be too high for practical applications, it is similar to levels reported in [56], an experiment where there was only one selection per trial.

5.7 Subjective Evaluations

Attribute	Eye Gaze Mean	Manual Mean
Liked	4.07	3.36
Easy	3.64	3.29
Fast	3.71	3.14
Accurate	3.71	3.21

Table 5-9 ***Results of Likert-Style Questionnaire. Scaled from 1 (Strongly Disagree), 2 (Disagree), 3 (Neutral), 4 (Agree) to 5 (Strongly Agree).***

Table 5-9 shows the mean scores for the responses to the questionnaire. Liking was the only attribute that showed significance in the Chi-squared analysis (Liked Eye Gaze, $\chi^2(4)=15.29$, $p<0.01$; Liked Manual $\chi^2(4)=12.43$, $p<0.02$). In general, participants preferred using Eye Gaze and found it easier, faster, and more accurate than the Manual alternative.

6 Discussion

6.1 Introduction

In this section, possible explanations for our findings will be discussed, relating results to our experimental predictions. First, we will go through the independent variables. This will be followed by a detailed look at the interactions between the independent variables and a presentation of our explanations for the observed phenomena. In particular, we will focus on the unexpected pattern of difference in response time between input mechanisms as the number of active targets in a trial changed. This discussion leads into the introduction of a theory which motivates the first key press response time analysis in Chapter 7. Chapter 7 seeks to illuminate differences in the decision process that subjects employed under various experimental conditions.

6.2 Base Case - Single Placeholder

The single placeholder is the simplest task presented in this paper. Subjects had to select only one target fixed at high priority. Eye gaze was not shown to be faster than the manual input mechanism. We believe that this is because the differences in the input mechanism allowed subjects to take a short cut when using the manual option. Recall that in order to select a target using the eye gaze condition, subjects had to fixate on the target. This requires 3 successive samples from the eyetracker at 30Hz (100 ms), to be followed by a key press indicating the priority of the target. In this part of the experiment, subjects only needed to perceive a change in colour, as the colour was guaranteed to be green. In the manual condition subjects could saccade instead of fixate.

In fact, due to the physiological properties of the eyes they did not even have to look directly at the target. Although perception drops off outside of the foveola, in the macula region, visual acuity is still good enough to detect colour change. The expanded visual angle of 16.7° in the macula region projects to 7.46" on the screen [12]. We believe that in the manual condition subjects saved the 100 ms that is required to fixate on the target in the eye tracker condition, which offset the 80-200 ms to press the additional key [6].

Given the simplest stimulus, performance using eye gaze and manual keyboard input mechanisms varied by less than 1%. Although this is not consistent with expectation, it provides the ideal base case for understanding how changes in complexity affect response time for the different input mechanisms. For example, as we can see from Table 5-6, the 1 target 1 priority pairing in the main experiment, where stimuli were presented in random order, showed an increase in response time of 22% for the eye gaze condition and 45% for the manual condition over the base case. The only mechanical difference between the base case and in the main experiment is that in the latter targets could be either green or yellow. This shows us that by putting a simple task in the context of more complex trials, response time significantly increased. In the rest of the experiment, because the target could be green, yellow or red subjects had to look at the target before making a selection.

6.3 Multiple Placeholder – Effects of Independent Variables

6.3.1 Input Mechanism

As an aggregate, comparing means, eye gaze outperformed the manual keyboard option by 290.87ms, or 16%. Although these results were not as dramatic as the ones found in [13], they are encouraging but not uniquely interesting. To provide new insight, we need to improve our understanding of how eye gaze performs relative to the manual option when the complexity of the task changes.

6.3.2 Number of Targets

As expected, increasing the number of targets increased the response time. Of note, the transition from 1 target to 2 targets showed a much larger delta (754.94 ms) then the transition from 2 to 3 (563.41 ms) or 3 to 4 targets (493.77 ms). We believe that the most likely explanation is that when there is only 1 target there is no possibility of having two priorities activated at once. This was the one condition where subjects did not have to prioritize their selection; they simply had to make it. As a consequence, the transition from 1 to 2 targets can be more generally thought of as the transition from 1 to multiple targets, encompassing both the difficulty of choosing multiple targets, and the difficulty of prioritizing the selections.

6.3.3 Number of Priorities

Increasing the number of priorities increased the response time. Although the result was consistent with expectation, the magnitude of the spread was surprising (633.95 ms). We believed that increasing the number of priorities would have much less of an effect in an absolute sense than increasing the number of targets. However, this includes cases when there is only one active target. When there is only one target, there can only be one priority. As shown in Table 5-3, response time is fastest when there is only one target. If we remove the one target case and only analyze situations where either one or two priorities are possible, the difference narrows to 320.91ms. Pairwise comparisons still yield significant differences at the $p < 0.001$ level. This result is more in line with expectations.

6.4 Two Way Interaction Effects

6.4.1 Number of Targets * Number of Priorities

There was no significant interaction effect between the number of targets and the number of priorities. We believe that this is due to the mechanics of selection. For each increase in the number of targets, either one or two key presses were added depending on the input mechanism. Increasing the number of priorities does not increase the number of actions required to complete a trial, it only constrains the order of selection. Thus the effect of the number of targets dominated the effect of the number of priorities producing no significant interaction effect.

6.4.2 Number of Priorities * Input Mechanism

Similarly, we did not see a significant interaction effect between the number of priorities and the input mechanism. Increasing the number of priorities did not affect the input mechanisms differently. Eye gaze was 14% faster than the manual option when there was 1 priority and 17% faster when there were two.

6.4.3 Number of Targets * Input Mechanism

There was a significant interaction effect between the number of targets and the input mechanism. However, if the difference between the two input mechanisms was only mechanical, we would have seen the same relationship, but different results. For each additional target, we would expect the performance difference to grow by either a fixed or a proportional amount.

Targets	Eye Gaze (ms)	Manual (ms)	Difference (ms)	% Difference
1	864.56	1023.36	158.8	18%
2	1474.95	1922.86	447.91	30%
3	2104.92	2402.22	297.3	14%
4	2682.03	2830.13	148.1	6%

Table 6-1 ***Difference in Response Time Means for Number of Targets * Input Mechanism***

As can be seen by **Table 6-1**, there does not seem to be a clear pattern to the magnitude of difference in response time between input mechanisms for a variable number of targets. Results may be trending to Manual input becoming relatively faster.

6.5 Accounting for the Pattern of Difference

We have three possible explanations for the lack of a clear relationship. In addressing these explanations, we will discuss three-way interaction effects and motivate the first key press response time analysis in Chapter 7. The explanations are:

1) The number of priorities significantly influenced response time
2) The input mechanism affected the execution process
3) The input mechanism affected the planning and decision processes

6.5.1 The Effect of Priorities

Because we treated priority as a random as opposed to a fixed factor, as the number of targets increased, so did the probability that two priorities be present in a trial. Although this could have had a significant impact on the results if the distribution was skewed to provide significantly more data points for a particular target, priority and input mechanism combination, this was not the case.

6.5.1.1 Three Way Interaction Effects

Number of Targets * Number of Priorities * Input Mechanism (ms)					
Targets	Priorities	Eye Gaze	Manual	Difference	% Difference
1	1	864.56	1023.36	158.80	18%
2	1	1421.54	1758.39	336.85	24%
	2	1532.46	2099.97	567.51	37%
3	1	2037.20	2379.58	342.38	17%
	2	2162.22	2423.24	261.02	12%
4	1	2531.39	2453.98	-77.41	-3%
	2	2778.88	3071.95	293.07	11%

Table 6-2 ***Difference in Response Time Means for Number of Targets * Number of Priorities * Input Mechanism***

When we introduce priority into the equation, and look at the interaction between number of targets, number of priorities and input mechanism, the magnitude of difference in means does not appear to follow a more predictable pattern. Aside from the 4 target 1 priority pairing, which we believe is an outlier caused by treating priority as a random variable, eye gaze is faster than manual input under all conditions. For each input mechanism, as the number of targets and priorities increases, so does the response time, but in a seemingly non linear fashion.

In summary, there is an interaction effect between the three independent variables, and eye gaze is significantly faster than manual input. The magnitude of difference between the input mechanisms may be trending to favor manual input as the task gets harder, but this relationship is unclear. The additional granularity provided by introducing number of priorities has not impacted the appearance that task difficulty does not exacerbate performance differences between input mechanisms. Thus we turn to an alternate explanation.

6.5.2 The Execution Process Differs by Input Mechanism

A second explanation is that the subjects may have used a different execution process when using eye gaze as opposed to manual input. It has been hypothesized elsewhere that a possible benefit to using eye gaze as a selection modality is that instead of serially choosing a target and selecting it, these two tasks may be performed in parallel – *while* visually choosing a target, subjects can press a key [14]. If so, the eye gaze condition should be one key press faster (80-200ms) than the manual condition for each target per trial. If not, response time should vary by at most the 100 ms difference in time between fixating on the target (100ms) and pressing a key for the slowest typist (200ms). Neither explanation fit the observed data.

The magnitude of difference did not linearly change 0-100 ms for each target in a trial. We believe that observed results were not consistent with either prediction because both explanations assume that the selection process for both the eye gaze and manual input mechanisms is sequential and involves the same subtasks:

Total Trial Time = $\sum$ ($Time_{Acquisition}$ + $Time_{Execution}$) *For all targets*

In the eye gaze condition, a sequential acquisition and execution process is implicitly enforced regardless of the number of targets per trial, as the subject had to look at each target to select it before specifying priority. However, in the manual condition, subjects did not individually have to look at targets before making each selection. They did have to initially look at the stimulus, but not necessarily focus on the next target before each selection. Some subjects may have only looked at the stimulus once, and then concentrated on execution, producing the following formulation:

Total Trial Time = $Time_{Acquisition}$ + $\sum$ ($Time_{Execution}$) *For all targets*

In the manual condition, we are unsure of the subject's sequence of looking behaviour while in the eye gaze condition it is constrained. This is exaggerated as completing a trial may have involved a long sequence of up to eight selections, permitting many different paths to completion. The selection process may vary by input device, by number of targets, priorities, individual subjects, or even by trial. For example, when there are four targets representing two priorities, subjects may only fixate on the first target of high priority and the first target of low priority. However, when there are four active targets of the same priority, they may only need to scan the monitor to acquire the targets without the need for fixation. Consequently, we can not assume selection mechanics and apply a simple model to predict selection time in this experiment and other experiments of a complex nature as the model is unlikely to explain the underlying process.

6.5.3 Decision Process Differs by Input Mechanism

A third explanation is that the input mechanism affects more than the execution process to complete a trial. Not only may the mechanical processes of selection differ, but the very method of what we will call route planning may be affected by the experimental conditions. In the previous section, we ventured that there may be a difference in when subjects plan when using different input mechanisms. When there are more paths to fulfill a composite objective, the subject may assess the stimulus and determine a sequence of actions that efficiently achieve the objective prior to executing the solution. With this explanation, we hypothesize that there may be a difference in the efficiency of the plan that they develop. In the next section of this paper, we will investigate what we believe may be an indicator of this sort of planning.

7 First Response Analysis

7.1 Introduction

There are two components to the selection process: deciding what target to choose and executing the decision. As an observer to the experiment, it seemed as though the subjects were approaching selection differently when using eye gaze as opposed to the manual option. As trials in this experiment often required multiple selections, there may have been an element of route planning. To further understand this, we looked at the first measured response – the first key press.

7.2 First Key Press Response Time Interpretation

If the first key press response time was in proportion to the total task time, then we would believe that there is no element of planning. If this is not the case, then by looking at how it varies to expectation we have an opportunity to gain insight into the planning process. We seek to understand whether planning is consistently different for various experimental conditions, and if so, how. For this analysis, we use the same data reported in Chapter 5 and discussed in Chapter 6 without excluding errors.

First Key Press	
Input Mechanism	***Meaning***
Eye Gaze	Priority (target is selected by eyes)
Manual	Target

Table 7-1 ***Summary of the Meaning of the First Key Press for different Input Mechanisms***

7.2.1 First Selection Meaning for Input Mechanisms

As shown in Table 7-1, in order to select the first key in the manual condition, a subject must have determined that the target is of the highest active priority (1 of 4 keys). In the eye gaze condition, the subject must also have chosen the highest priority target, as the key press specifies priority (1 of 2 keys). So in either case they have made the same decision. Hick's Law predicts that it will take 110.5 ms longer when given 4 selection candidates in the manual condition as opposed to the 2 provided in the eye gaze condition. In order to conclude that subjects are using a different decision process for different input mechanisms, we need to show:

1. First response time variances between input mechanisms exceed a conservative estimate of the mechanical difference in execution
2. First response times are not proportional to the total trial time

7.3 Hypotheses

In the eye gaze and manual input conditions, the fingers of one hand are positioned over the keyboard targets. Thus execution involves 1 degree movement. In the eye gaze condition, 1 key is chosen from 2 possibilities; in the manual condition 1 key is chosen from 4 possibilities. Hick's law predicts that this will take subjects an extra 110.5 ms when using the manual condition. This will be offset by the 100 ms latency time required for eye gaze selection.

1. *The input mechanism will not effect response time for the first key press*

Number of Targets

Adding selection options will reduce response time because there are fewer restrictions in which keys can be pressed. However, more targets do not necessarily imply a greater number of appropriate candidates for the first selection. In the case that there are 2 priorities active, more targets could mean more low priority options which will not influence the number of first selection candidates.

2. *Number of targets will not affect response time*

Number of Priorities

Adding priorities will increase response time. Rather than choosing targets that are not red, subjects needed to prioritize their selection which will increase the time to make the first selection.

3. *Number of priorities will increase response time*

Number of Targets * Number of Priorities

Increasing the number of targets and priorities will demonstrate an interaction effect. If there are more targets to choose from of a single priority then we are guaranteed to have more candidates for the first selection which will reduce response time. For example, if there 2 two targets of the same priority, there will be 2 candidates for selection (2 out of 4 options), whereas if they are varied, there will only be 1 candidate for selection (1 out of 4 options). More selection options will reduce selection time.

4. *The interaction of Targets and Priorities will affect first key response time.*

Input Mechanism * Number of Targets

Differences in response time will be explained by the input mechanism, not the number of targets. As discussed above, number of targets does not imply number of selection options.

5. *Number of Targets will not have an interaction effect with input mechanism for first key response time.*

Input Mechanism * Number of Priorities

More priorities will not affect the first key response times differently across input mechanisms. In every condition, the subject must decide on the target of highest priority in advance of selection. Differences in response time will be explained by the input mechanism factor.

6. *Number of priorities will not have an interaction effect with input mechanism for first key response time.*

Input Mechanism * Number of Targets * Number of Priorities

An increasingly complex decision process will affect response times for the first key press. However, the main driver of response time differences will be the Number of Targets * Number of Priorities combination. Before the first response can be executed, the same decision process must have concluded regardless of the input mechanism, so we expect to not see a significant 3 way interaction effect.

7. *Number of targets and priorities will not have an interaction effect with input mechanism for first key response time.*

8 First Response Results

Just as in Chapter 5, first key responses involved three independent variables; Input Mechanism (2 Levels), Number of Targets (4 Levels) and Number of Priorities (2 Levels). Between Subjects effects were analyzed using GLM Univariate one way analysis of variance (ANOVA). Significance was evaluated at the α =.05 level. Trials containing errors were included in the analysis as discussed in section 7.2. Within factors differences between pairs of means were detected using GLM Univariate Pairwise comparisons at the α =.05 level with Bonferroni adjustments for multiple comparisons.

8.1 Effects of Independent Variables on First Key Press Time

Input Mechanism (ms)			
Input Mechanism	**Mean**	**(S.E.)**	**Results**
Eye Gaze	851	9	≠ M
Manual	231	10	≠ EG

Table 8-1 *Means and Standard Errors for input mechanism condition*

8.1.1 Effects of Input Mechanism on First Key Response Time

Table 8-1 shows the data summary for the input mechanism used in the trial. Analysis of variance showed that there were significant differences across levels ($F_{1,188}$ = 2032.394, $p < 0.001$). We reject the null hypothesis that input mechanism did not affect response time. Post-hoc comparisons with Bonferroni correction showed that conditions differed significantly from each other. Manual input was 268% faster than Eye Gaze.

Number of Targets (ms)			
Targets	Mean	(S.E.)	Results
1	615	11	≠ 2 ≠ 3 ≠ 4
2	523	11	≠ 1 = 3 = 4
3	528	13	≠ 1 = 2 = 4
4	527	16	≠ 1 = 2 = 3

Table 8-2 ***Means and Standard Errors for the number of active targets per trial***

8.1.2 Effects of Targets on First Key Response Time

Table 8-2 shows the data summary for the number of targets that required action per trial. Analysis of variance showed that the response times for the number of targets did not differ significantly across levels ($F_{3,188}$ = 1.793, p = 0.147). We accept the null hypothesis that number of targets did not affect response time. Post-hoc comparisons with Bonferroni correction showed that only 1 target differed from the other levels.

Number of Priorities (ms)			
Priority (ies)	Mean	(S.E.)	Results
1	634	10	≠ 2
2	421	8	≠ 1

Table 8-3 ***Means and Standard Errors for the number of priorities represented per trial***

8.1.3 Effects of Priorities on First Key Response Time

Table 8-3 shows the data summary for the number of priorities that were represented per trial. Analysis of variance showed that there were significant differences across levels ($F_{1,188}$ = 183.236, $p < 0.001$). We reject the null hypothesis that the number of priorities does not affect response time.

8.2 Two Way Interaction Effects

Number of Targets * Number of Priorities (ms)				
Targets	1 Priority	(S.E.)	2 Priorities	(S.E.)
1	642	10		
2	684	14	393	17
3	706	21	408	11
4	706	29	422	10

Table 8-4 ***Means and Standard Errors for the interaction between number of targets and number of priorities represented per trial***

8.2.1 Effects of Targets * Priorities on First Key Response Time

Table 8-4 shows the data summary for the effect between the number of targets and the number of priorities. Analysis of variance showed that there is not a significant interaction effect between the number of targets and the number of priorities represented per trial ($F_{2,281}$=0.447, p=0.640). We accept the null hypothesis that there was no interaction effect between number of targets and number of priorities on response time.

Number of Priorities * Input Mechanism (ms)				
Input Mechanism	1 Priority	(S.E.)	2 Priorities	(S.E.)
Eye Gaze	1041	15	615	11
Manual	234	15	228	11

Table 8-5 ***Means and Standard Errors for the interaction between number of priorities and the input mechanism***

8.2.2 Effects of Priorities * Input Mechanism on First Key Response Time

Table 8-5 shows the data summary for the effect between the number of priorities and the input mechanism used in the trial. Analysis of variance showed that there was a significant interaction effect between the number of targets and the number of priorities represented per trial ($F_{2,188}$=21.082, $p < 0.001$). We reject the null hypothesis that there was no interaction effect between number of priorities and input mechanism on response time.

Number of Targets * Input Mechanism (ms)				
Targets	Eye Gaze	(S.E.)	Manual	(S.E.)
1	902	15	330	15
2	792	16	254	16
3	843	18	202	20
4	898	22	185	22

Table 8-6 ***Means and Standard Errors for the interaction between Number of Targets and Input Mechanism***

8.2.3 Effects of Targets * Input Mechanism on First Key Response Time

Table 8-6 shows the data summary for the effect between number of targets and the input mechanism used in the trial. Analysis of variance showed that there was a significant interaction effect between the number of targets and input mechanism ($F_{3,188}$ = 26.119, p < 0.001). We reject the null hypothesis that there was no interaction effect between number of targets and input mechanism on response time.

8.3 Three Way Interaction Effects

8.3.1 Effects of Targets * Priorities * Input Mechanism on Time on First Key Response Time

Number of Targets * Number of Priorities * Input Mechanism (ms)					
Targets	Input Mechanism	1 Priority	(S.E.)	2 Priorities	(S.E.)
1	Eye Gaze	902	15		
	Manual	330	15		
2	Eye Gaze	1002	24	582	22
	Manual	229	22	281	23
3	Eye Gaze	1072	32	614	18
	Manual	193	36	212	17
4	Eye Gaze	1249	46	648	16
	Manual	174	44	194	17

Table 8-7 ***Means and Standard Errors for Number of Targets * Number of Priorities * Input Mechanism***

Table 8-7 shows the data summary for the number of targets * number of priorities * input mechanism effect. Response time differences across conditions were not significant ($F_{2,188}$ = 0.654, p=0.52). We accept the null hypothesis that there was no interaction effect between number of priorities, number of targets and input mechanism on response time.

8.4 Summary of Findings – Significant Differences

Test	Our Hypothesis	Results
Input Mechanism	N	Y
# Targets	N	N
# Priorities	Y	Y
Targets * Priorities	Y	N
Priorities * Input Mechanism	N	Y
Targets * Input Mechanism	N	Y
Targets * Priorities * Input Mechanism	N	N

Table 8-8 ***Summary of findings for First Response***

9 First Response Discussion

9.1 Introduction

In this section, we will discuss our findings and relate them to our predictions, providing explanations for differences when they occur. We will begin by discussing the effect that the number of targets, the number of priorities and their interaction had on first response time. Then we will proceed to the principle question of this analysis; do subjects employ a different planning process when using the eye gaze as opposed to manual input? In this discussion, we will compare the results of the first response analysis to the main experimental results that were communicated in Chapter 5.

9.2 Effects of Targets and Priorities

9.2.1 Number of Targets

As expected, the Number of Targets condition did not have a significant effect on the first response time. We believe that this is because targets do not imply first strike candidates. If both low and high priority targets are active in a trial, then low ones would qualify as targets, but would not be appropriate for a first key press.

However, if we look at the interaction between the number of targets and the input mechanism, as shown in Table 8-6, we see that each additional target reduced the first response time in the manual condition by 21%, while seemingly not affecting the response time for the eye gaze condition. Because of the unclear relationship between targets and candidates, we are more able to account for this relationship when we include the effect of priorities.

9.2.2 Number of Priorities

As predicted, the number of priorities present in a trial significantly affected first response time. However, we expected priorities to positively affect first response time whereas it produced a negative effect of 51%. This may be because two active priorities guarantees that at least one target was of high priority. The contrast provided by the additional colour may allow the subject to quickly rule out the low priority option and isolate the high priority target(s).

Like the effect of the number of targets, the impact of the number of priorities is different across input mechanisms. Table 8-5 shows that when using eye gaze, going from 1 priority to 2 reduced the first response time by 426 ms, or 41%. Using manual input this change had no effect on first response time (6 ms improvement). The pronounced effect in the eye gaze condition may be partially due to the immersive quality of eye gaze as an input device. As reported in [22], by using a human input mechanism as an output device used to select targets, we forced subjects to focus their attention, drawing them into the task. This effect is exaggerated by the selection mechanics in the eye gaze condition, where subjects had to individually fixate on targets to select them. We believe that this may have put subjects in the habit of focusing on individual targets throughout the entire task, changing the search pattern from parallel preattentive to a sequential visual search [49, 50]. Thus while the studying the stimulus prior to the first selection, as opposed to just searching for the first high priority candidate, subjects may have fixated on each of the highest priority targets visible.

9.2.3 Number of Targets * Number of Priorities

Although no interaction effect is shown, Table 8-4 indicates that the number of priorities has a greater effect on first strike response times than the number of active targets. This is consistent with expectations.

9.2.4 Experimental Condition Effect Trends

Independent Variables Effect Trends		
Input Mechanism	**Priorities**	**Targets**
Eye Gaze	Negative	Positive
Manual	Positive	Negative

Table 9-1 ***Summary of Relationship Trends with respect to First Response Time by Input Mechanism***

Examining Table 9-1, we see the number of priorities and the number of targets seemed to affect first key response time for the input mechanisms in opposite ways. When there are fewer targets subjects take less time to make the first choice when using manual input, but more time when using eye gaze. Accordingly, more priorities markedly reduces the first selection response time when using eye gaze, but slightly increases it when using manual input.

9.3 Do Subjects engage in a Different Selection Process when using Different Input Mechanisms?

Prior to executing the first response, in all conditions, subjects have identified a target which is a member of the highest priority that is active in the trial. As stated in section 7.2, to establish that they use a different process to arrive at their selection when using different input mechanisms we believe that two criteria must be satisfied:

1. Response time variances between input mechanisms must exceed the amount of time that can be attributed to execution differences (110.5 ms)
2. Response time variances must not be proportional to total trial time

To this we add a third criterion:

3. There is no alternative explanation that better accounts for the observed data

In this section, we will attempt to establish these three criteria.

9.3.1 First Response Time Variances between Input Mechanisms

As found in Table 8-1, as an aggregate, first key response time was 268% faster using manual input than it was when using eye gaze, with an absolute difference of 620 ms. This greatly exceeds the 110.5 ms threshold predicted by Fits law. Even if we counter-intuitively sum the predicted effects of what we assumed would be the greatest sources of variability; the eye tracker latency, and the range in speed between proficient typists, the result is 210.5 ms. This has only accounted for 34% of the observed variance.

However, this may be an unfair comparison. The first response for the eye gaze condition specified both the target and priority, while in the manual keyboard condition the first response only selects the target. If we account for time that it takes to subsequently select the priority by very conservatively[1] doubling the first response time for the manual condition, it is still 84% faster than eye gaze.

The data suggests that the difference in the first response times for the input mechanisms exceeds what we would expect to be caused by the mechanical process of execution. However this alone is not sufficient to conclude that a different decision process was employed when using the different input mechanisms. If the results remain proportional to the total task time, then we may simply be observing the comparative effect of the experimental conditions on total task time as described in Chapter 6.

9.3.2 Proportional Variance to Expectation: Analysis Method

Priorities	Targets	Eye Gaze (ms)	Manual (ms)
1	1	37.00	(181.68)
	2	291.48	(211.00)
	3	393.27	(203.20)
	4	615.77	(132.51)
2	2	(184.12)	(244.02)
	3	(106.50)	(191.71)
	4	(46.68)	(189.56)

Table 9-2 – ***Differences from Proportional Expectation – Negative values, displayed in red and in parentheses indicate faster response times than proportionately expected***

Table 9-2 was constructed to show how results of the first key press proportionately relate to the entire task time. The total time recorded for a trial was divided by the number of key presses required to complete the trial to give an expected time for the first key press. This figure was compared to the first key response time.

[1] This is conservative as subjects have already decided on the target, and determined that it is of highest priority in preparation for the first selection. For the second selection subjects simply have to map the solution to 1 of 2 buttons representing priority and execute the selection.

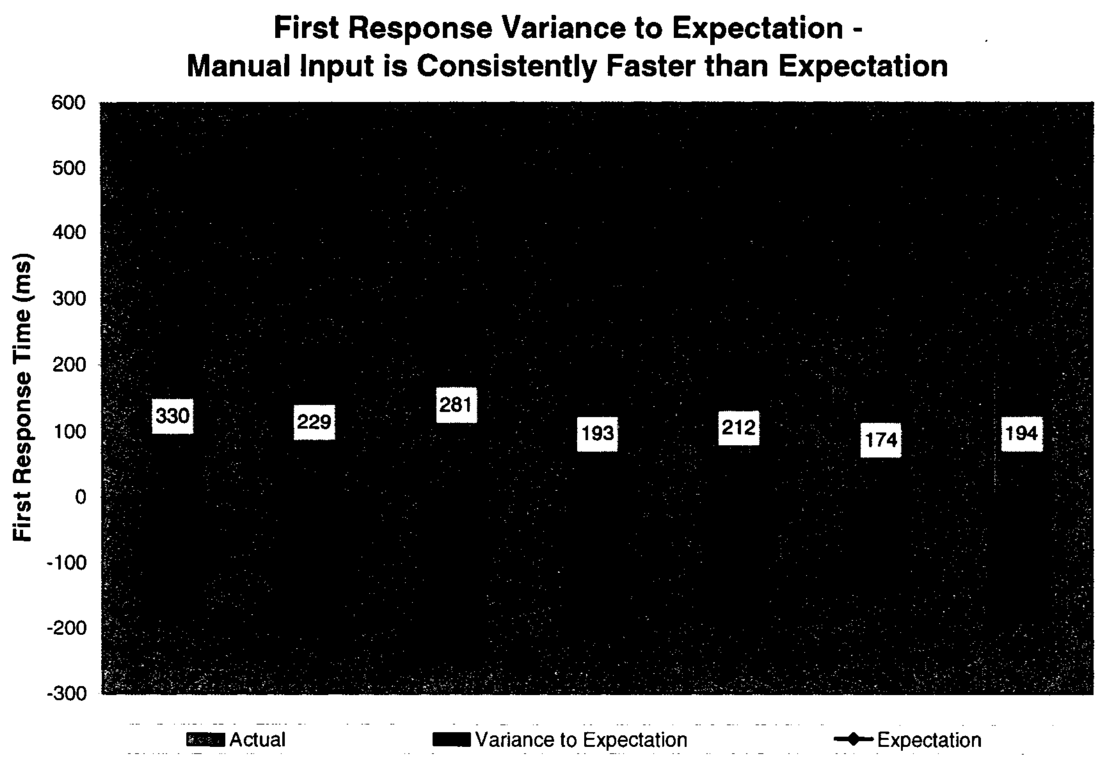

Figure 9-1. *Chart of First Response Variance to Expectation for Manual Input.*

9.3.3 Proportional Variance to Expectation: Manual Input

As can be seen in Table 9-2 and in Figure 9-1, the manual condition is consistently faster than expectation in all scenarios. There does not appear to be a relationship between the number of active targets or priorities and the proportional difference to expectation; that is to say, the magnitude of the proportional difference is consistent. This, in conjunction with the lack of variance as compared to eye gaze in Table 8-7 indicates that subjects did not employ a route planning process.

Returning to Table 9-2, the first response sets a faster pace than the subsequent responses follow. We believe that the time which was lost throughout the trial was mainly due to the iterative switching from target selection to priority specification. Although there were only 6 total buttons to choose from, grouping buttons into different meaning brackets presented a coordination challenge to subjects. Our strongest validation for this theory is the 1 button, 1 target pairing. As seen in Table 8-7, the first response took 330 ms. Table 5-7 shows that using manual input, subjects took 1023 ms on average to complete the trial. The difference between the two cases is the second press which consequently took approximately 673 ms. Recall that prior to the first press, subjects had already determined the target and its priority. After the first press, they simply had to map the solution to 1 of 2 priority keys and execute the second key press. This is a less complex decision and execution than the first response, but took 102% more time. Thus we believe that the switching from the target specification group to the priority group of buttons must have disproportionately affected response time. For the manual condition, the mechanical process of selection appears to account for the proportional difference in the first response time as compared to the total task time.

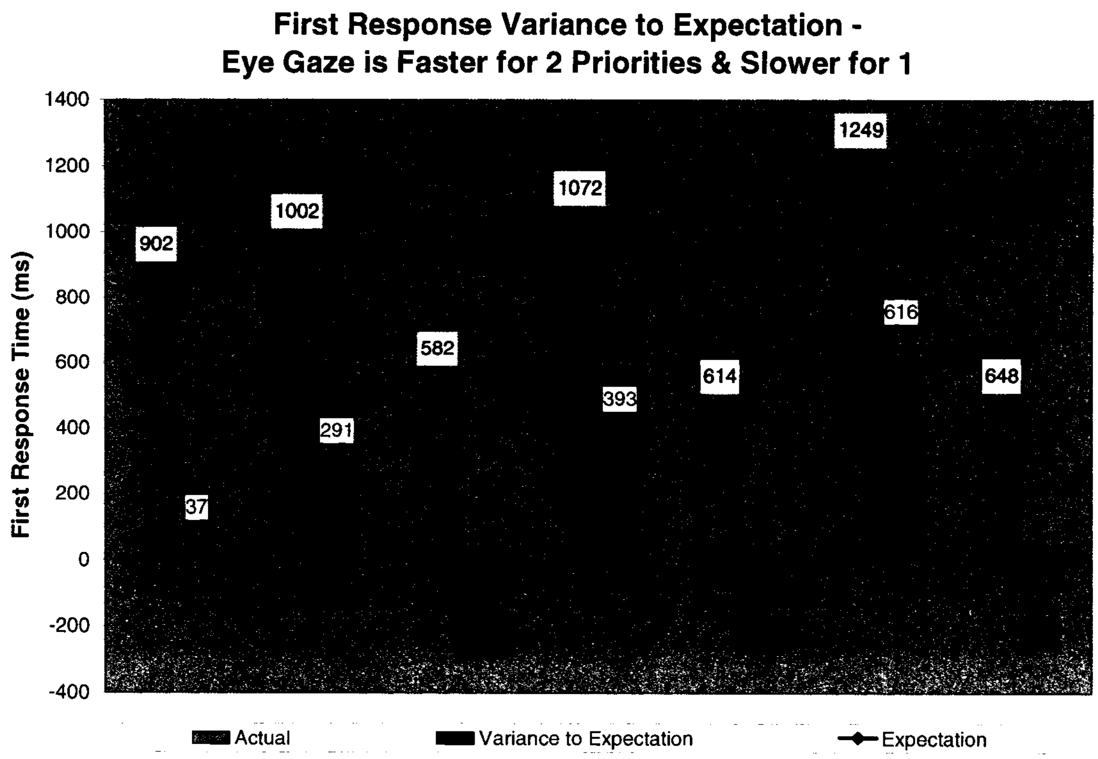

Figure 9-2. *Chart of First Response Variance to Expectation for Eye Gaze.*

9.3.4 Proportional Variance to Expectation: Eye Gaze

As can be seen in **Table 9-2** and Figure 9-2, there are two prominent first response variances to proportional expectation when using eye gaze:

1. Response time decreases as number of priorities represented in a trial increases.
2. Response time increases as the number of targets is increased.

Thus first response time is proportionately slowest when there is only 1 priority present, and 4 targets to choose from. This is surprising as in this case, all of the targets are first strike candidates and thus subjects have the most freedom to arbitrarily choose any of the

4 targets shown. In section 9.3.1, we hypothesized that when there are 2 priorities, the colour contrast allows subjects to quickly subdivide the problem and identify suitable targets. However, if contrast improved first response time, why would it make the total task time worse and only proportionately affect response time when using eye gaze?

By forcing subjects to fixate on targets before making a keyboard selection, we may have changed the subject's search pattern from preattentive pattern matching to a visual search, individually considering all of the options of equal priority before making a selection [50]. While considering the options, we believe that subjects formulated a route plan which initially caused a delay visible in the first response time. As the number of first response candidates grew, so did first response time in both absolute terms (Table 8-7), and in proportion to the total task time (Table 9-2). The more selection freedom subjects had, the more time they spent planning their route. This planning led to fast intermediate selections, producing a total task time significantly less than both the manual option, which did not exhibit the characteristics of similar planning, and trials which contained 2 active priorities. When there were 2 priorities in a trial, subjects had less flexibility to determine their selection order with fewer options to consider in the route plan. Consequently, first reaction time was faster, and intermediate selections were slower, producing a slower overall result (Table 5-7).

We believe that in trials with a less constrained order of selection, when using eye gaze subjects do not choose an arbitrary sequence; they plan their route. It should be noted that it is possible that the experimental selection constraints, not the eye gaze modality was the cause of the route planning that lead to faster intermediate selections and total trial times. By requiring fixations solidified by a disambiguating key press that forced subjects to map selections to an alternate modality, we may have inadvertently disrupted the preattentive selection pattern, and created the conditions to support planning. If this is the case, then this approach may be the most appropriate model to use to interact with computers in unconstrained environments. As described in section 9.3.6 results in this analysis may also been affected by a context or *transition pause* when switching from high to low priority selection.

9.3.5 Proportional Variance to Expectation: Summary

Overall, neither the eye gaze nor the manual input condition demonstrated results proportional to the total task time. The observed data appears to support the explanation that subjects employ a different selection process for the different input conditions. In the manual condition, subjects did not appear to employ a planning process, and the reduction in speed during trials is attributable to the difficulty of iteratively switching from target selection to priority specification. In the eye gaze condition, during the scenario where there was the greatest number of first response options, subjects took the longest to execute their first decision, but completed the trial quickest. This suggests that subjects took time to plan a selection route that yielded significantly faster response times for the rest of the trial. More flexibility for the eye gaze condition appeared to allow for better planning. However, before drawing a conclusion, we will consider an alternative explanation.

9.3.6 Alternative Explanation: Execution Process

An alternative explanation for the observed data is that execution mechanics, not the planning process was the cause of the variance between first response and total task performance. To begin, let's choose the 1 priority eye gaze case which demonstrated the largest variance to proportional expectation. As discussed, the immersive quality of the eye gaze selection modality makes it harder for the subject to quickly evaluate first response targets unless there is obvious colour contrast. Thus it takes subjects longer to make the first selection, but because the mechanics of selection are simpler, subjects pick up time throughout the trial. Like the manual option, where the process of alternating between selecting targets and priorities slows down selection, when subjects have 2 priorities in a trial they have to change from pressing the button to signify high priority to the button that signifies low priority. At this point, they are affected by a *transition pause* which accounts for total task performance being faster when there is 1 priority as opposed to 2.

This is a simpler explanation and requires far fewer assumptions. However, it does not account for why eye gaze does not show improvements or declines at a consistent rate relative to the manual option. Rather, the magnitude of difference is generally stable but occasionally inexplicably variable across conditions, but may be trending to favor manual input. If the differences in response times were driven by the mechanics of selection, we would expect to see a linear magnitude of difference in response times between factors. As discussed in detail in section 6.5, the failed attempt to find simple relationships between trial complexity and response times over input mechanisms is what initially inspired the first response investigation.

Thus we believe that subjects invoked a process of planning when using eye gaze, particularly when there was flexibility in the path to complete the trial. Although this account is more complicated, and relies on more assumptions, we believe it to be more reasonable, and not conceptually irreconcilable with the alternative explanation. As stated in the alternative explanation, we believe that the *transition pause* delayed the response time for selections subsequent to the first response.

9.4 First Response Time Summary

We believe that results support the theory that the mechanical use of the input mechanism may not be the principle driver of differences in response times. When using specific input mechanisms, the actual process of planning and decision making changes. Increasing the number of targets and the number of priorities appeared to affect the first response times across input mechanisms in opposite ways. Both eye gaze and manual input exhibited first response times that were disproportionate to the total task time. As a whole, the result of this investigation was that the first response was significantly faster when using manual input as opposed to eye gaze. Relating this to the results shown in section 6.3.1, despite the first response taking 268% longer, using eye gaze the total task was completed 16% faster. Thus we believe that the difference in initial planning time produces a more efficient and effective execution result, particularly when there is greater flexibility in the response path.

Conducting an experiment that involves many conditions over composite trials introduces significant variability, which is why many factors may conspire to produce a nonlinear model. Thus the attempt to forecast response time when selection complexity is a parameter may be unrealistic. However, we hope to be able to test the hypothesized planning effect in future experiments that are more targeted towards evaluating the merit of our theory.

9.5 Limitations

Because the subjects were chosen from a pool of technically competent university students, our ability to generalize from the results is restricted to the group that they represent. Additionally, the subjects who did have experience using an eye tracker may have been more proficient using the technology, which could have skewed the data. In the future, we should try to only use subjects who have an equivalent amount of experience and technical competence, or alternatively use the experience that they have as an experimental parameter. We also may have benefited from experimenting with different key placement layouts to provide the fastest and most accurate platform to use in comparison to eye gaze. We also may have overestimated the significance of our results as we analyzed repeated measures using a factorial analysis [30]. Had we not taken this approach, the targets * input device, and targets * priorities * input device interactions may have not shown significance.

10 Conclusions

In this thesis, we presented an investigation of eye gaze versus manual input for target selection in conditions of variable complexity. In the base case where subjects had to select a target of known priority appearing in a known location, eye gaze and manual input exhibited similar results (variance of 1%). When the order of targets and priorities were varied and randomized, on average, eye gaze was 16% faster than manual input and 37% faster in the best case (Table 5-2). As shown Table 5-7, with one exception, for all combinations of number of targets and priorities, eye gaze was consistently and significantly faster than the manual option. Subjects also preferred using eye gaze and found it easier, faster and more accurate than the manual alternative even though there were more errors. Consequently, it stands to reason that eye gaze may be an appropriate choice to select devices in unconstrained environments.

Although eye gaze remained faster than manual input with increased task complexity, the magnitude of difference did not change in an obvious fashion. We expected that as the task complexity increased, differences between the input devices would become more pronounced. If so, one would assume that the mechanics of selection between the input mechanisms would be the cause of the spread. This, in turn, would allow us to develop a preliminary performance model that used selection complexity as a predictive parameter. However, our findings do not support this conclusion. Thus we presented an alternative explanation for the observed data; if selection mechanics were not explaining the response time variance between input mechanisms, then perhaps the planning and decision process would.

To investigate this, we segregated the first measured response from the total trial time, and compared results across experimental conditions. As seen in Table 8-1, we found that as an aggregate, subjects took 268% more time to make their first selection using eye gaze then they did using manual input. This is contrary to what we would expect as visually selecting a coloured target is a preattentive task that should favor eye gaze [49, 50]. It can be argued that these selections are not equivalent - using eye gaze, the first selection involves specifying both target and priority, which specifies the same information as the first two selections using manual input. However, even if we double the amount of time that it takes to make the first selection using manual input, it is still faster than the first response time using eye gaze under all conditions, by an average margin of 84%.

Both eye gaze, and manual input demonstrated first response times that were disproportional to the total measured task time. As seen in **Table 9-2**, when using the manual option, subjects' first response time was faster than expectation for all conditions. We believe that in the manual condition, subjects did not engage in route planning prior to execution, and were slowed throughout the trial by the difficulty of iteratively switching between target specification and priority specification.

Using eye gaze, subjects were slightly faster when there were two priorities, but significantly slower when there was only one, implying that the intermediate selections were faster than the first response. Unexpectedly, in the case where there was only 1 active priority, the first response time was the slowest, while the total task performance was the fastest. We believe that this was a consequence of route planning. When subjects had the most freedom to make a selection, they did not make an arbitrary first selection. Instead they took extra time to plan all of their selections and then rapidly executed them. These results are not consistent with purely preattentive selection. We believe that fixating on the target when using the eye gaze condition may have caused the subjects to use a sequential visual search inadvertently enabling the more effective planning process [50].

10.1 Applications

One of the main benefits to Attentive User Interfaces is that they recognize and respond to the non verbal cues that people routinely exchange [40]. These cues are interpreted to indicate what or whom people are interested in, as people do in group conversation [45]. AUIs use this attentive context to regulate the sequencing of interactions in an ongoing discourse between people and communicative digital devices.

This experiment showed that in virtually all cases, eye gaze was a faster modality to choose targets in a selection task designed to be analogous to a user opening up communication channels with digital appliances. In situations where there are the least number of constraints on subject behaviour – they could choose any target they wished as they were all suitable, subjects engaged in a more efficient and effective planning process when using eye gaze. This positions the eye gaze input technology to succeed in environments where there are a lot of options, and few constraints – for instance the real world.

In addition to opening and closing communication channels, eye gaze as a selection modality enables new interactive possibilities. Using eye gaze, users can focus on the behaviors they are interested in, without having to worry about either device naming conventions or the process of specifying targets, as the target is implied by gaze. We saw the negative effect of naming conventions even in the simple case where subjects had to find a key to represent a target rather than just looking directly at it. We believe that switching from specifying the target to specifying intention was the main cause of why manual input proportionately slowing down after the first response. Using eye gaze, user focus changes from learning the language of technology to concentrating on the intended outcome of the action, facilitating an impact driven model of interaction.

As computer technology continues to populate the environment, it will be increasingly important that it helps focus, rather than distract people. Thus we see eye gaze providing the context of attention, combined with generic references to disambiguate intention as the ideal remote control for the real world. As evidenced by Google's meteoric rise, as virtual and real continue to intersect, attention may quickly become a dominant commodity.

11 Future Work

11.1 Increasing the Complexity of Trials: Towards a Performance Model

Because eye saccades are implicitly jittery [22], and the one degree size of the foveola guarantees an implicit error in the eye tracking process, we restricted the number of targets in this desktop scenario to four. It would be worthwhile to repeat this experiment, increasing the number of targets and priorities to a level where the complexity of the mapping problem will be seen across both the number of selection options and the number of priorities. Realistically, even the most effective Attentive User Interfaces will present the user with more than two notification priorities. For the manual input mechanism, each additional target or priority option would increase the number of keys (keys=priority modes + target options). For the eye gaze condition, an increase in the number of priority modes will only add one additional key (keys=priority modes). Adding more options will either show whether and at what point performance in the manual option breaks down in absolute terms relative to eye tracking because of the increased selection complexity, or whether the two methods continue to converge as suggested by the data. It will also give us an opportunity to examine not only absolute differences, but also the rate of change in how people perform when faced with increases in decision complexity. It may be possible to abstract relationships between input mechanisms and certain forms of decision complexity that can help inform next generation interface design.

As two decision axes may be too trivial to establish a principle from which we can generalize, it may sensible to add a third dimension to the decision process. This could be achieved by adding a varying size dimension. Experiments which vary the complexity of selection may provide the basis to develop a performance model that applies to interactions with groups of computers in unconstrained environments. With improvements in current eye contact sensing technology, we will soon see whether our results apply to more ecologically valid scenarios, outside of the confines of the desktop computer. This will also allow us to see whether real world spatial cues affect the response time of the various input mechanisms.

11.2 Validating and Extending the First Response Results

To understand whether the planning process is caused by the conjunction of the visual search pattern and switching modalities to execute the disambiguating key press, we may conduct a simpler follow up experiment comparing eye gaze using dwell time with eye gaze disambiguated by a key press. Analyzing the first selection response time by evaluating its proportional difference to the aggregate task time would indicate whether subjects also engage in planning when there is no mapping requirement. As discussed in section 2.6, although eye gaze dwell time is not a suitable input mechanism for unconstrained environments, this investigation would improve our understanding of the planning effect.

Many Attentive User Interfaces use speech, as opposed to manual key presses to disambiguate their intention with reference to the target of their visual attention [1, 2, 27, 40, 42, 44]. Speech may equivalently show both the same planning delay, and planning benefits as we believe that the manual input exhibited in this experiment. We would like to conduct an experiment that evaluated gaze and speech against gaze and manual key presses to explore the generalizability of the planning effect.

For both of these future experiments, we can validate many of our assumptions by studying differences in the visual scan path depending on the input mechanism and the experimental conditions.

11.3 Final Words

Although experiments that involve many conditions that are subject to a variety of effects may be contentious, we believe that they are necessary to ensure that HCI evaluation techniques are flexible and responsive enough to scale to real world applications. In a world where change is the status quo, attention is increasingly monetized as the incumbent commodity and good ideas are evaluated by the living, breathing and never sleeping marketplace, the age old techniques of evaluation in highly controlled, artificial settings may lose relevance. We believe that by using complexity as a moderately controlled parameter, we have learned more about the selection lifecycle of eye gaze then we would have through multiple, more specific investigations of simpler phenomena. Taking the opportunity to introduce a wiki citation, as the renowned artificial intelligence researcher and Turing Award recipient Herbert Alexander Simon [17] may have asked, at what point do we replace sufficiency in the context of a specific, well defined objective with the goal of satisficing a bigger idea?

12 References

1. Bolt, R.A. 'Put-That-There': Voice and gesture at the graphics interface. SIGGRAPH 1980. Seattle: ACM Press, 1980. pp. 262-270.

2. Bolt, R. A. Gaze-Orchestrated Dynamic Windows. In Proceedings of the 8th Annual Conference on Computer Graphics and Interactive Techniques. Dallas: ACM Press, 1981, pp. 109-119.

3. Broadbent, D. Perception and Communication. New York: Pergamon Press, 1958

4. Buxton, W. Less is More (More or Less), in P. Denning (Ed.), The Invisible Future: The seamless integration of technology in everyday life. New York: McGraw Hill, 2001. pp. 145 – 179.

5. Card, S.K., Moran, T.P. and A. Newell. The Psychology of Human-Computer Interaction. Mahwah: Erlbaum, 1983.

6. Card, S.K., Moran, T.P. and A. Newell. The keystroke-level model for user performance time with interactive systems, Communications of the ACM, Vol. 23 No. 7 (July 1980), New York: ACM Press. pp.396-410

7. Castells, M. The Rise of the Network Society. Oxford: Blackwell Publishers, 1996.

8. Cherry, C. Some experiments on the reception of speech with one and with two ears. Journal of the Acoustic Society of America 25, 1953, pp. 975-979.

9. Cleveland, N. R. Eyegaze Human-Computer Interface for People with Disabilities. In Proceedings of Automation Technology and Human Performance, 1994. Washington, DC. Available at http://www.lctinc.com/doc/cathuniv.htm (28.9.2001).

10.Dell Computers. Dell 1702FP. http://support.ap.dell.com/support/edocs/monitors/1702fp/en/specs.html, 2006.

11.Dickie, C., Vertegaal, R., Shell, J. S., Sohn, C., Cheng, D. and O. Aoudeh. Eye Contact Sensing Glasses for Attention-Sensitive Wearable Video Blogging. In Extended Abstracts of CHI 2004. Vienna: ACM Press, 2004. pp.769-770.

12.Duchowski, A. T. Eye Tracking Methodology: Theory & Practice. London: Springer-Verlag, 2003.

13.Eye Tracking Research and Applications (ETRA). www.e-t-r-a.org. (2006)

14.Fono, D., and Vertegaal, R. EyeWindows: Evaluation of Eye-Controlled Zooming Windows for Focus Selection. In Proceedings of CHI 2005. Portland: ACM Press, 2005. pp.151-160.

15.Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J. and R. J. Dolan. The Trouble with Cognitive Subraction. Neuroimage Vol. 4 No 2. (October 1996). Academic Press pp. 97-104.

16.Geisler W. S. and Perry J. S. 1999. Variable resolution displays for visual communication and simulation. Society for Information Display, 30, 420-423.

17.Herbert Simon. Wikipedia. http://en.wikipedia.org/wiki/Herbert_Simon. (2006)

18.Hansen, J.P. Tørning, K., Johansen, A. S., Itoh, K., and H. Aoki. Gaze typing compared with input by head and hand. In Proceedings of ETRA 2004. San Antonio: ACM Press, 2004. pp.131-138.

19.Hick, W. E. On the rate of gain of information. Quarterly Journal of Experimental Psychology. Vol 4. 1952. pp. 11-26.

20.Josephson, S. and M. E. Holmes. Visual attention to repeated internet images: testing the scanpath theory on the world wide web. In Proceedings ETRA 2002. New Orleans: ACM Press, 2002. pp. 43-49.

21.Just, M.A. and Carpenter, P.A. A theory of reading: from eye fixations to comprehension. Psychological Review Vol. 87 No. 4 (1980). pp. 329-354.

22.Jacob, R.J.K. Eye Tracking in Advanced Interface Design. In Virtual Environments and Advanced Interface Design, ed. by W. Barfield and T.A. Furness. New York: Oxford University Press, 1995. pp. 258-288.

23.Kveraga K, Boucher L, Hughes HC. Saccades operate in violation of Hick's law. Experimental Brain Research. Vol. 146 No. 3 (October 2002). pp. 307-14.

24.LC Technologies. The Eyegaze System. www.eyegaze.com. 2006.

25.Lockerd, A. and F. Mueller. LAFCam: Leveraging affective feedback camcorder. In Extended Abstracts of CHI 2002, Minneapolis: ACM Press, 2002. pp. 574-575.

26.Maglio, P. Barrett, R., Campbell, C. and T. Selker. Suitor: An Attentive Information System. In Proceedings of International Conference on Intelligent User Interfaces 2000, New Orleans: ACM Press, 2000. pp.169-176.

27.Maglio, P., Matlock, T., Campbell, C., Zhai, Z. and B. Smith. Gaze and speech in attentive user interfaces. In Proceedings of the International Conference on Multimodal Interfaces. Berlin: Springer-Verlag, 2000. pp. 1-7.

28.Majaranta, P., and K-J. Räihä. Twenty years of eye typing: systems and design issues. In Proceedings ETRA 2002, New Orleans: ACM Press, 2002. pp.15-22.

29.Mello-Thoms, C., Nodine, C.F., and H. L. Kundel. What attracts the eye to the location of missed and reported breast cancers? In Proceedings of ETRA 2002, New Orleans: ACM Press, 2002. pp. 111-117.

30.Max, L and P. Onghena, Some Issues in the Statistical Analysis of Completely Randomized and Repeated Measures Designs for Speech, Language, and Hearing Research. JSLHR (1999), Vol. 42. pp. 261-270.

31.MAX/MSP 4.3 for Windows: Cycling '74. www.cycling74.com. 2006.

32.McCarthy, J. F. Active Environments: Sensing and Responding to Groups of People. Personal and Ubiquitous Computing Vol. 5 No. 1 (February 2001), London: Springer-Verlag. pp. 75-77.

33.Murphy, H. and A. Duchowski, Gaze-Contingent Level Of Detail Rendering. In Proceedings of EuroGraphics 2001. Manchester: EuroGraphics, 2001.

34.Nielsen, J. Noncommand User Interfaces. Communications of the ACM, Vol 36 No 4 (April 1993), New York: ACM Press. pp. 83-99.

35.Norman, D. A. The Invisible Computer: Why Good Products Can Fail, The Personal Computer is So Complex, and Information Appliances are the Solution. Cambridge: MIT Press, 1998.

36.Norman, D. A. The Psychology of Everyday Things. New York: Basic Books, 1988.

37.Oh, A., Fox, H., Van Kleek, M., Adler, A,. Gajos, K., Morency, L-P, and T. Darrell. Evaluating Look-to-Talk: A gaze-aware interface in a collaborative environment. In Extended Abstracts of CHI 2002. Seattle: ACM, 2002, pp. 650-651.

38. Ramloll, R., Trepagnier, T., Sebrechts. M., and A. Finkelmeyer. A gaze contingent environment for fostering social attention in autistic children. In Proceedings ETRA 2004, San Antonio: ACM Press, 2004. pp. 19-26.

39. Selker, T, Lockerd, A. and J. Martinez. Eye-R, a Glasses-Mounted Eye Motion Detection Interface. In Extended Abstracts of CHI 2001. Seattle: ACM Press, 2001. pp. 179-180.

40. Shell, J. S., Selker, T., and R. Vertegaal. Interacting with Groups of Computers. Communications of the ACM Vol. 46 No. 3 (March 2003), New York: ACM Press. pp 40-46.

41. Shell, J. S., Vertegaal, R. and A. Skaburskis. EyePliances: Attention-Seeking Devices that Respond to Visual Attention In Extended Abstracts CHI 2003 Ft.Lauderdale: ACM Press, 2003

42. Shell, J. S., Bradbury, J., Knowles, C., Dickie, C., and R. Vertegaal. eyeCOOK: A Gaze and Speech Enabled Attentive Cookbook. In Extended Abstracts of UbiComp 2003. Seattle: Springer-Verlag, 2003.

43. Shell, J. S., Vertegaal, R., Mamuji, A., Pham, T., Sohn, C. and A. Skaburskis. EyePliances and EyeReason: Using Attention to Drive Interactions with Ubiquitous Appliances. In Extended Abstracts of UIST 2003. Vancouver: ACM Press, 2003.

44. Shell, J. S., Vertegaal, R., Cheng, D., Skaburskis, A., Sohn, C., Stewart, A. J., Aoudeh, O. and C. Dickie. ECSGlasses and EyePliances: Using Attention to Open Sociable Windows of Interaction. To Proceeds of ETRA 2004, San Antonio: ACM Press, 2004. pp. 93-100.

45.Short, J., Williams, E., and B. Christie. The social psychology of telecommunications. London: Wiley, 1976.

46.Sibert, L.E. and R.J.K. Jacob. Evaluation of Eye Gaze Interaction. In Proceedings of CHI 2000. The Hague: ACM Press. pp. 281-288.

47.Skaburskis, A.W., Shell, J.S. and Vertegaal, R. and C. Dickie. AuraMirror: artistically visualizing attention. In Extended Abstracts of CHI '03. Ft. Lauderdale: ACM Press. pp. 946-947.

48.Starker, I., and Bolt, R.A. A gaze-responsive self-disclosing display. In Proceedings of CHI'90 Seattle: ACM Press. pp. 3-9.

49.Theeuwes, J. Visual selective attention: A theoretical analysis. Acta Psychologica Vol. 83. No. 2. (November 1993). pp. 93-154.

50.Treisman, A., and G. Gelade. A feature integration theory of attention. Cognitive Psychology Vol. 12, No.1 (January 1980). pp. 97-136.

51.Vertegaal, R. The GAZE Groupware System: Mediating joint attention in multiparty communication and collaboration. In Proceedings of CHI'99. Pittsburgh: ACM Press, 1999, pp. 294-301.

52.Vertegaal, R., Weevers, I., and Sohn, C. GAZE-2: An attentive video conferencing system. In Extended Abstracts of CHI 2002. Minneapolis: ACM Press, 2002, pp. 736-737.

53.Vertegaal, R., Dickie, C., Sohn, C. and M. Flickner. Designing attentive cell phones using wearable eyecontact sensors. In Extended Abstracts of CHI 2002. Minneapolis: ACM Press, 2002, pp. 646-647.

54.Vertegaal, R. and Y. Ding, Explaining effects of eye gaze on mediated group conversations: amount or synchronization? In Proceedings of the CSCW 2002. New Orleans: ACM Press, 2002 pp. 41-49.

55.Wang, J., Zhai, S. and H. Su. Chinese Input with Keyboard and Eye-Tracking – An Anatomical Study. In Proceedings of CHI 2001. Seattle: ACM Press, 2001 pp. 349-356.

56.Ware, C. and H.H. Mikaelian. An Evaluation of an Eye Tracker as a Device for Computer Input. In Proceedings of CHI+GI 1987. Toronto: ACM Press, 1987 pp. 183-188.

57.Welford, A.T. Fundamentals of Skill. London: Methuen, 1968.

58.Weiser, M., The Computer for the 21st Century. Scientific American. Vol. 265 No.3 (September 1991). pp. 94-104.

59.Wooding, D.S. Fixation maps: quantifying eye-movement traces. In Proceedings of ETRA 2002. New Orleans: ACM Press, 2002. pp. 31-36.

60.Zhai, S., Moromoto, C., and S. Ihde. Manual and Gaze Input Cascaded (MAGIC) Pointing. In Proceedings of CHI 1999. Pittsburgh: ACM Press, 2002. pp. 246-253.