

Chapter 8

Why Are Eye Mice Unpopular? – A Detailed Comparison of Head and Eye Controlled Assistive Technology Pointing Devices

R. Bates and H.O. Istance

8.1 Head Mice and Eye Mice

8.1.1 Introduction

Head-based and eye-based assistive technology pointing devices, or head and eye mice, are used to enable computer control in a similar manner to standard hand-held desktop mice, by simply moving the computer cursor to where a user is pointing either with their head or their eyes. These devices have been in existence for many years within the motor-disabled community and over this time there has been widespread acceptance and adoption of head mice, with many disabled people using these devices to access computers and communication devices. However, anecdotal evidence suggests that eye-based pointing is widely and quickly dismissed as a usable means of pointing in assistive technology due to difficulty of use and inaccuracy, making the potential advantages of eye-based pointing difficult to realise in the face of such quick decisions not to use this approach. It is necessary to understand in detail what the limitations of eye-based pointing are and how these can be overcome. Making relative performance comparisons with head-based pointing can usefully do this, as this is an established technique of assistive technology pointing.

8.1.2 For Eye Pointing

Firstly, eye-gaze has the potential to be a very natural form of pointing, as people tend to look at the object they wish to interact with (Jacob 1991, Salvucci 2000). Secondly the speed of eye-gaze to locate a target can be very fast when compared to other pointing devices (Ware 1987, Sibert 2000). Thirdly, due to the specialised nature of the muscles controlling the eye, natural eye movements exhibit little

detectable fatigue and offer near fatigue-free pointing (Saito 1992). Finally, eye-trackers can be non-encumbering, as they do not require the user to wear anything. In contrast, head pointing requires conscious movement and steering of the head to point at an object. This form of pointing can be slow due to the high mass of the head restricting rapid movement, it is fatiguing due to the non-specialised nature of the neck muscles, which may tire after repeated head movements and it often requires the user to wear a target or device. Initially it appears that eye-based pointing has considerable advantages over head-based pointing.

8.1.3 Against Eye Pointing

Firstly the eye is not a highly accurate pointing device as it exhibits a positional tolerance (Carpenter 1991, Jacob 1995). The fovea of the eye, which gives clear vision, covers a visual angle of $\approx 1^\circ$ arc of the retina, hence when fixating a target the eye only needs to be within $\approx 1^\circ$ of the target position to clearly see the target. This gives an inaccuracy in measured gaze position. Secondly, since eye gaze position cannot easily be consciously controlled or steered, as it is driven by subconscious interest (Yarbus 1967), the eye tends to fixate briefly on targets of interest before jumping to other points of interest. Thus it requires effort to fixate steadily on a target for any extended period of time. This contrasts sharply with the deliberate, if slow, controlled movement and positional accuracy of head-based pointing. Thirdly the eye is being employed as both an input modality to the user, so the person can see the computer interface, and an output modality from the user to the interface, indicating the pointing intention of the user on the interface. This convergence of interaction point and gaze-point means that the pointing cursor, in contrast to head pointing, cannot be parked or left at a position on the screen whilst the eye momentarily looks away. Such convergence results in unwanted pointing movements at the feedback point on the computer screen as the cursor follows the eye wherever it gazes (Jacob 1995, Velichkovsky *et al.*, 1997). Finally, unlike head-trackers, eye-trackers are not widely available, and can be expensive.

Clearly, both types of pointing devices have advantages and disadvantages. The question that must be answered is to what extent each of these advantages and disadvantages influence the performance of the devices and lead to the acceptance or rejection of each device by their intended users.

8.2 Assessing the Devices

8.2.1 A 'Real World' Test

A 'real world' experimental test sequence, rather than an abstract target acquisition test, was used to test the performance of the devices. The test consisted of a total of 150 simple tasks across two domains, word-processing with Microsoft Word and web browsing with Internet Explorer, which formed a natural flow of interaction. The proportions of object usage, target sizes and interaction techniques

in the two sets of tasks mimicked as closely as possible ‘real world’ interaction based on previous observation of users. The 150 test task objects that comprised the test, such as a button or menu item, were assigned one of four size categories (0.3°, 0.6°, 0.9°, 1.2°) based on the smallest visual angle subtended by the screen object central to the task at a distance of 60cm from the screen. Typical objects for each category were spin control button or text in the 0.3° category, filenames and scrollbar buttons 0.6°, menu items and toolbar buttons 0.9°, icons and keyboard keys 1.2°. The distribution of object sizes in the test was determined by the nature of the real-world task, with 0.3° objects taking 4%, 0.6° 17%, 0.9° 23% and 1.2° 56% of the total test tasks.

8.2.2 Device Performance

The usability of the two pointing devices was assessed in terms of device *performance* based on the European ESPRIT MUSiC performance metrics method (Macleod *et al.*, 1997). Performance was defined as ‘*The quality of interaction with the device and the time taken to perform that interaction*’. Device performance was calculated by measuring the quality of interaction during the tasks by counting the number of incorrect commands (such as hitting the wrong target), the number of target misses (with no command generated), the number of cursor position corrections (unnecessary position correction cursor movements), and the time taken to complete the tasks (measured from the start of a task until the task was finished or abandoned). Tasks were initially given a quality rating of 5 (perfect) (Szczur 1994), with subsequent errors reducing the quality until the task was completed or failed, and the next task started. Task performance with each device was calculated by a simple formula (8.1) such that a task that has the highest level of quality and takes no time would give a performance of 100%, with any reduction in quality or increase in time degrading the measured performance:

$$\text{Task performance} = \frac{\text{Quality of interaction (1-5)}}{5 + \text{Time taken for interaction (secs)}} \quad (\text{equation 8.1})$$

8.2.3 Test Participants

Six able-bodied test participants were chosen for the experiment. The participants were selected to give a wide range of experience using the head and eye mice from very experienced users through to novice users with little previous experience of the devices. All participants were experienced hand mouse users. Two participants were categorised as ‘low’ experience users of both of the devices, two as ‘medium’ experience users and two as ‘high’ experience users. (The number of available participants with experience of both test devices was limited and the number of hours experience of using each device to define each category is shown later in Table 8.2).

8.2.4 Experimental Design

The experiment was a within-subjects design, with all six participants completing one session of the 150 test tasks with each of the three devices, hand mouse, head mouse and eye mouse. These sessions were conducted in random order. Prior to these taking place, each participant was given training and practice to become familiar with all of the test tasks. A hand held mouse was used for the practice sessions. Participants were also given the opportunity to practice some of the test tasks with the eye and head mouse until they felt comfortable with the use of these devices.

8.2.5 Test Apparatus

A standard PC was used for the tests. For the eye mouse a Senso-Motoric Instruments infrared video-oculography eye-tracker was used to measure eye-gaze position with a software driver used to move the cursor in response to the eye-gaze of the test participants. A Polhemus ISOTRACK electromagnetic motion tracking system was used to measure the head position of the test participants for the head mouse and a second software driver used to move the cursor in response to head position. Target selection was by a hand held micro-switch and text entry was via a WiViK on-screen keyboard. A standard desktop hand mouse was included in the tests as a baseline for comparison. Participants were seated with a head and eye to monitor screen distance of 60cm on a seat with a backrest and head support to help participants to maintain their head position and to increase seating comfort.

8.2.6 Test Analysis

All data was obtained by capturing the complete contents of the test computer screen during each session, including the cursor position, at a rate of 5 frames per second. The data was analysed by stepping through the video files and recording the quality and time taken to perform each task within a session. In addition, the time taken by any non-productive actions during the task was measured and the nature of the non-productive action was recorded.

8.3 Test Results

8.3.1 Overall Performance

The test showed that there were no overall differences between the word-processing and web-browsing performance for each device, but there were differences in performance between the devices (figure 8.1).

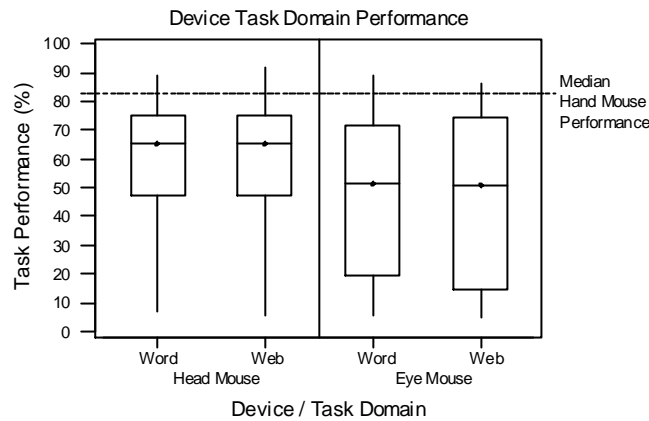


Figure 8.1. Overall device performance

Grouping the word and web results together for each device gave a performance of 65.2% for the head mouse and 51.1% for the eye mouse. This difference was statistically significant (Mann-Whitney Test, $U=301228$, $n_1=900$, $n_2=900$, $p<0.0001$), with the head mouse outperforming the eye mouse by 1.28 times. For comparative purposes participants completed the same tasks with a hand-held mouse. The performance for the hand mouse was the same over the two domains, and grouping the data gave a performance of 83.3%, considerably higher than the assistive technology devices.

8.3.2 Task Domain Performance

The similarity of device performance between the word and web task domains showed that the context of the tasks had little effect on the performance of the devices. The frequencies of target sizes present in each of the tasks in the domains were similar, which may have accounted for the similarities in measured performance much more than the context of the tasks.

8.3.3 Target Size Performance

Examining the device performance by target size shows a distinct relationship between increasing task performance and increasing target size (figure 8.2). Here again, the eye mouse was inferior to the head mouse for each of the target sizes. Non-parametric regression was used to determine the relationship of performance to target size (equations 8.2 and 8.3).

$$\text{Task performance for the head mouse} = 38.9 * \text{target size} + 21.6 \quad (\text{equation 8.2})$$

$$\text{Task performance for the eye mouse} = 32.2 * \text{target size} + 16.3 \quad (\text{equation 8.3})$$

Projecting these models to determine at what target sizes the devices have the same performance, or achieve 100% performance, should be treated with caution, as it is probable that device performance will become non-linear as efficiency approaches 100%. Of more use was translating these models into equivalent target sizes for equal device performance over the range of target sizes in the test.

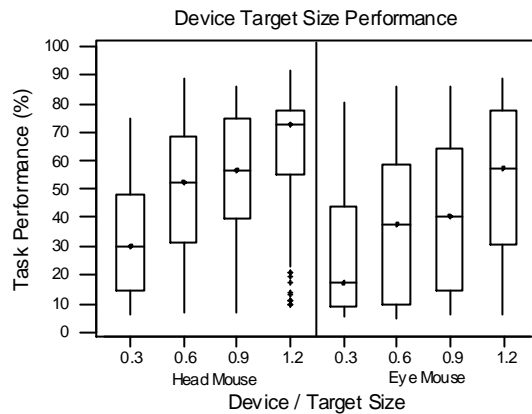


Figure 8.2. Device performance for different target sizes

For example, for the eye mouse to achieve the same efficiency as the head mouse at a target size of 0.3° , the target size for the eye mouse would need to be $((38.9 \times 0.3) + 21.6 - 16.3) / 32.2 = 0.53^\circ$ (table 8.1).

Table 8.1 Equivalent target sizes for equal performance

Head mouse target size	Required eye mouse target size	Percentage increase in eye mouse target size
0.3°	0.53°	176%
0.6°	0.89°	148%
0.9°	1.26°	140%
1.2°	1.62°	135%

It is notable that the difference in the required level of target magnification for the eye mouse decreases rapidly with increasing target size, indicating that the eye mouse is capable of performing well on larger target sizes.

Comparing the head mouse target sizes to the required eye mouse target sizes revealed a progression, where for equivalent performance the eye mouse targets must be approximately one size category larger than the head mouse targets. For example, 0.3° head mouse targets gave approximately the same performance as

0.6° eye mouse targets. This strongly suggested that if the target sizes could be increased then the eye mouse could achieve parity with the head mouse. One mechanism to achieve this is to provide the user with an ability to ‘zoom in’ on an area of the screen (Bates 1999) which effectively increases target sizes temporarily. The means of controlling this zoom mechanism and the resultant effect on performance are currently being investigated.

8.3.3.1 Elements of Performance and Target Size

The task quality and task time elements that constituted the performance metric showed steady increases in quality and steady reductions in task time with increasing target size. The head mouse was again superior to the eye mouse with the highest quality and lowest times for the target sizes. Of the two elements, the relationship between task time and target size was most interesting (figure 8.3).

The task time was broken down into six elements: productive time, the time lost generating incorrect commands, misses and cursor control corrections, the time lost whilst the eye mouse cursor was displaced looking at the feedback point on the interface, and finally the time lost during calibrations of the devices. Looking at the individual elements of task time, the most time was lost in cursor control corrections. This was the most detrimental element for both devices, indicating that considerable time was wasted correcting the cursor position onto targets.

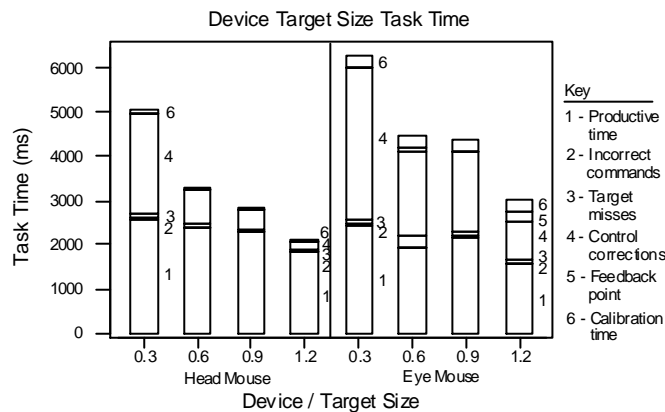


Figure 8.3. Device task time for different target sizes

Most interesting was the similarity in productive time for each of the devices. In particular, the eye mouse had shorter productive times (was more efficient) than the head mouse, indicating that it had the potential to be equal or superior to the head mouse if the non-productive elements could be reduced. The time lost in incorrect commands and misses was not significantly different between the devices.

Looking at elements of the eye mouse task time that could be reduced suggested that if target size continued to increase then the control correction time could be minimised, and the eye mouse would perform as well as the head mouse.

It is notable how these results show the performance dependency of the assistive technology devices on the size of the targets present on the computer interface.

8.3.4 Participant Experience and Performance

Table 8.2 shows the numbers of hours of device usage accumulated by participants in each of the three experience categories.

Table 8.2. Participant experience categories and exposure hours

Experience category	Head mouse hours	Eye mouse hours	Numbers of participants
Low	0.25	1 – 2	2
Medium	0.5 – 1	6 – 8	2
High	2 – 3	15 - 30	2

There was a definite relationship between the experience rating of test participants with the devices and their performance, with increasing experience resulting in increasing task performance (figure 8.4).

Here again, the eye mouse was inferior to the head mouse for the low and medium experienced participants. However, the two devices achieved parity of performance, at 73.0% for the head mouse and 73.5% for the eye mouse in the high experience participants. While the number of hours device experience for these participants is very different between devices, the data suggests that the eye mouse could approach the performance levels of the head mouse if participants are sufficiently practiced. In steady-state, performance with neither device is likely to exceed the 83% performance baseline obtained from the hand mouse.

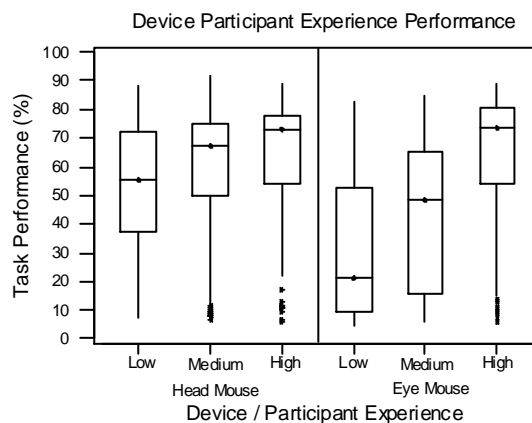


Figure 8.4. Device performance for different participant experience

The performance of the medium group of head mouse users after 0.5 to 1 hour experience is far higher than the low experience eye mouse group with 1 to 2 hours experience. The learning times coupled with the poor performance results for low and medium experienced participants for the eye mouse strongly suggested that considerable effort and dedication would be required to obtain even a moderate level of performance with the eye mouse. This was a strong indicator as to why eye mice appeared to be rarely used, and suggested that more work is needed to investigate the most efficient means of training participants to use the eye mouse in order to reduce the time needed to achieve these levels of performance.

8.3.4.1 Elements of Performance and Experience

Examining the task quality and task time elements that constituted the performance metric showed steady increases in quality and steady reductions in task time with increasing participant experience. The head mouse was superior to the eye mouse with the highest quality and lowest times for the low and medium experience participants, with the eye mouse achieving parity for the highly experienced participants. Again, of the two elements, the relationship between task time and experience was most interesting (figure 8.5).

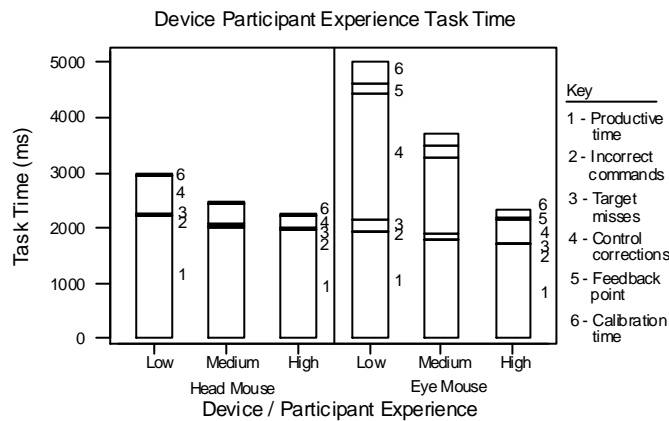


Figure 8.5. Device task time for different participant experience

Here we see dramatic reductions in the non-productive elements of the eye mouse task time for highly experienced participants. Calibration time and time spent at the feedback point has been cut significantly and the time lost in cursor control corrections reduced to near parity with the head mouse. Both sets of data show a levelling out of the productive time component with increasing experience. This suggests that it is possible with much practice to overcome the weaknesses of the eye mouse and use the strengths of the device to match the performance of the head mouse.

8.3.5 How Good Can an Eye Mouse Get?

To investigate just how good the performance of the eye mouse could be, the target size performance of the highly experienced participants was separated from the data for all of the devices, including the baseline hand mouse, and compared (figure 8.6). Here we see that the hand mouse was superior to the assistive technology devices, even when experienced participants use these devices. However as the target size increased, so the performance of the head and eye mice approached that of the hand mouse. Examining the largest target size, the hand mouse had a performance of 83.3% compared to the head mouse at 75.2% and the eye mouse at 78.1%. At this point the eye mouse exceeds the performance of the head mouse, this difference was statistically significant (Mann-Whitney Test, $U=10906$, $n_1=168$, $n_2=168$, $p=0.0003$), and reaches to within a few percentage points of the hand mouse. This showed that if target sizes were larger and users experienced then the performance of an eye mouse might exceed a head mouse, and approach that of a hand mouse.

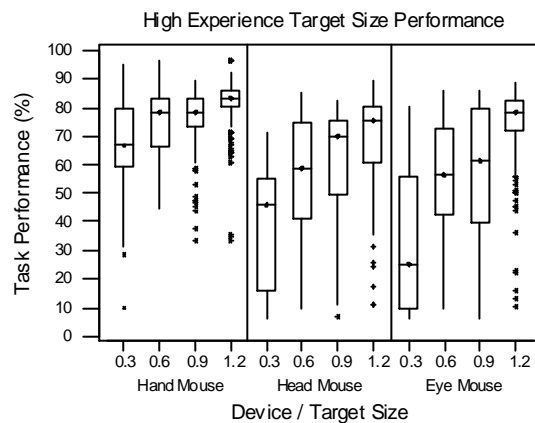


Figure 8.6. High experience participant device performance for different target sizes

8.4 Summary

The question this paper raised was why are eye mice dismissed as unusable when head mice are widely accepted, particularly in view of the advantages of eye mice in terms of speed and ease of use?

There were no differences in task domain performance for a given device, so the nature of the tasks did not distract or influence the performance of the participants with the eye mouse. The overall performance of the eye mouse was found to be poor.

Breaking down the tasks by target size revealed a trend of increasing performance with target size and showed that the eye mouse performance approached the head mouse for larger sizes, as the non-productive elements of

pointing were reduced. It also showed that the actual productive time for the eye mouse was shorter than the head mouse.

Looking at participant experience it was found that highly experienced eye mouse users could exceed their performance with the head mouse in the test as they had learnt how to minimise the non-productive elements of eye mouse pointing. However, this came at a cost of a much longer learning time with the eye mouse compared with the head mouse. It should be noted that it is possible that the performance of the head mouse could improve further than that found in the test if a user had extended head mouse experience up to the 30 hours of the experienced eye mouse users. However, it was observed that the head mouse was extremely quick to learn, as evidenced by performance improvements over time. From this rapid learning time, it was expected that head mouse performance would not improve significantly with extended experience.

Finally, it was shown that an eye mouse could exceed the performance of a head mouse and approach that of a hand mouse if target sizes were large and users sufficiently well practiced.

This experiment used able-bodied participants. It is interesting to speculate as to how the results would change if disabled participants had been used. For example, the range of neck motion and the potential for fatigue may reduce the performance of the head mouse and favour the eye mouse. Conversely any reduction in the gaze performance of the participants, due to nystigmus for example, will degrade eye mouse performance in favour of the head mouse.

To answer the question, why have eye mice not been adopted, is that they require slightly larger target sizes and higher user experience to perform well. For novice users the eye mouse would manifest itself as unusable and the system would be quickly rejected, despite its potential benefits. In comparison, the head mouse has moderate performance on smaller targets and a short learning time, so a new user would be able to use the system to some extent on their first attempt, making the system much more attractive. Methods to allow users to control target sizes through ‘zooming in’ on the screen can overcome this and are already being investigated. This could be coupled with training for users to reach an experienced level quickly. Users should be encouraged to have patience with the eye mouse until they become experienced.

With sufficient control over target size, and sufficient experience, eye mice can be used effectively. For people with motor disabilities who dislike or have difficulties using a head mouse, or cannot use a head mouse, the option to use an eye mouse at least as effectively as a head mouse would greatly enable them to operate the computers and environmental controls around them.

8.5 References

- Bates, R. (1999) Multimodal Eye-based Interaction for Zoomed Target Selection on a Standard Graphical User Interface Proceedings of Interact'99 vol II p7-8. British Computer Society.
- Carpenter, R.H.S. (1991) Eye movements, MacMillan.

- Jacob, R.J.K. (1991) The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get, *ACM Transactions on Information Systems* 9, 3, p152-169.
- Jacob, R.J.K. (1995) Eye Tracking in Advanced Interface Design, in W. Barfield and T. Furness (Ed) *Advanced Interface Design and Virtual Environments*, Oxford University Press, 258-288.
- Macleod, M., Bowden, R., Bevan, N., Curson, I. (1997) The MUSiC performance measurement method. *Behaviour and Information Technology* vol 16, no. 4/5 pages 279-293, Crown.
- Saito, S. (1992) Does Fatigue Exist in a Quantitative Measurement of Eye Movements? *Ergonomics* 35 5/6, 607-615.
- Salvucci, D.D. Anderson, J.R. (2000) Intelligent gaze-added interfaces, *Proceedings of CHI 2000*, 2, 273-280.
- Sibert, L.E. Jacob, R.J.K. (2000) Evaluation of eye gaze interaction *Proceedings of CHI 2000*, 2, 281-288, ACM.
- Szczur, M. (1994) Usability testing - on a budget: a NASA usability test case study. In *Behaviour and Information Technology*, vol 13, no.1 and 2, pp 106-118.
- Velichkovsky, B. Sprenger, A. Unema, P. (1997) Towards Gaze Mediated Interaction: Collecting Solutions of the "Midas Touch" Problem, In S.Howard, J.Hammond & G.Lindgaard (Ed) *Human-Computer Interaction: INTERACT'97*, Chapman & Hall.
- Ware, C. Mikaelian, H.H. (1987) An Evaluation of an Eye Tracker as a Device for Computer Input, *Proceedings of CHI 1987*, ACM.
- Yarbus, A.L. (1967) *Eye Movements and Vision*, Plenum, New York.