Univ Access Inf Soc (2003) 2: 280–290 / Digital Object Identifier (DOI) 10.1007/s10209-003-0053-y

# Why are eye mice unpopular? A detailed comparison of head and eye controlled assistive technology pointing devices

**R. Bates, H.O. Istance**

Human-Computer Interaction Research Group, De Montfort University, The Gateway, Leicester LE1 9BH, UK
E-mail: {rbates, hoi}@dmu.ac.uk

**Abstract.** This paper examines and compares the usability problems associated with eye-based and head-based assistive technology pointing devices when used for direct manipulation on a standard graphical user interface. It discusses and examines the pros and cons of eye-based pointing in comparison to the established assistive technology technique of head-based pointing and illustrates the usability factors responsible for the apparent low usage or 'unpopularity' of eye-based pointing. It shows that user experience and target size on the interface are the predominant factors affecting eye-based pointing and suggests that these could be overcome to enable eye-based pointing to be a viable and available direct manipulation interaction technique for the motor-disabled community.

**Keywords:** Eye tracking – Eye mouse – Head mouse – Assistive technology – Computer input devices

## 1 Introduction

Eye-based and head-based assistive technology computer pointing devices, commonly known as eye mice and head mice, are used to enable computer graphical user interface direct manipulation in a manner similar to standard hand-held desktop mice by simply moving the computer cursor to where a user is pointing with either their eyes or their head. 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 and control computers and communication devices. However, anecdotal evidence from disabled user groups and rehabilitation clinics and the rarity of eye-based devices in the motor-disabled community suggest that eye-based pointing is an unpopular means of assistive technology computer interaction. Typically, eye-based devices are dismissed due to apparent difficulty of use and inaccuracy, making any potential advantages of eye-based pointing difficult to realise. This paper investigates in detail what the limitations of eye-based pointing are and suggests how these may be overcome by making relative performance comparisons with head-based pointing, an established technique of assistive technology pointing and direct manipulation.

### *1.1 Benefits of 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 [12, 23]. Secondly, the speed of eye gaze to locate a target can be very fast when compared to other pointing devices [6, 7, 13, 24, 31]. 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 [22]. Finally, eye trackers can be non-encumbering, as they do not require the user to wear anything in contact with the body. In contrast, head pointing requires conscious movement and steering of the head to point at an object. This form of pointing can be comparatively slow for target acquisition tasks [14, 18] due to the high mass of the head, which restricts rapid movement. Also, it may be difficult, slow and inaccurate due to restrictions in range of neck motion [15, 16] and uncomfortable and fatiguing due to the non-specialised nature of the neck muscles which may tire after repeated head movements [8]. Finally, it is encumbering as it often requires the user to wear a target or device [8, 10, 11, 28]. On the basis of the above considerations, it appears that eye-based pointing has considerable advantages over head-based pointing.

### *1.2 Disadvantages of eye pointing*

Firstly, the eye is not a highly accurate pointing device as it exhibits a positional tolerance [6,13]. The fovea of the eye, which gives clear vision, covers a visual angle of approximately 1° arc of the retina, hence when fixating a target the eye only needs to be within approximately 1° 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 tends to be driven by subconscious interest [32], the eye tends to fixate briefly on targets of interest before jumping to other points of interest. Thus this lack of direct conscious pointing control requires effort by the user to fixate steadily on a target for any extended period of time. This contrasts sharply with the deliberate, if slow, controlled conscious movement and positional accuracy of head-based pointing [14–16,18]. 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 while the eye momentarily looks away to view the results of a user command or feedback from the interface. Such convergence results in unwanted and potentially distracting and unproductive pointing movements at the feedback point on the computer screen as the cursor follows the eye wherever it gazes [13,29]. 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 hence lead to the acceptance or rejection of each device by their intended users.

## 2 Assessing the devices

### *2.1 Test apparatus*

A standard PC was used for the tests. For the eye mouse a Senso-Motoric Instruments [25] infrared videooculography 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 [20] electromagnetic six-degree-of-freedom motion tracking system was used to measure the head position of the test participants for the head mouse, and a second software driver was used to move the cursor in response to head position. Target selection for the eye mouse and head mouse was achieved by a hand-held micro-switch. In addition to the eye and head mice, a standard desktop hand mouse was included in the tests as a baseline for comparison. All text entry was via a WiViK [21] assistive technology on-screen keyboard using the default QWERTY English language layout. Participants were seated with a head and eye to monitor screen distance of 60 cm on a seat with a backrest and head support to help participants maintain their head position and to increase seating comfort.

### *2.2 Test participants*

One important element of the experiment was to investigate how experience with the devices affected performance. The number of available participants with a wide range of experience with both test devices was limited due to the rarity of eye mouse devices and the training time required by the devices. However, large numbers of test participants were not essential for this type of experiment, as the number of participants required to identify the usability problems, and hence performance, of a system can be quite small [30]. Based on [30], only six test participants were required to determine 100% of 'high severity' usability problems, and with this number of participants 95% of 'medium severity' usability problems and 60% of 'low severity' problems could also be found. Hence, six able-bodied test participants were chosen for the experiment. The participants were selected to represent a wide range of experience using the head and eye mice, from very experienced users to novice users with little previous experience with the devices. All participants were experienced hand mouse users. During an initial participant selection process conducted by expert users to determine participant experience with the devices, an observation of the participants while using the devices was performed, and two participants were selected and categorised as 'low' experience users of both devices, two as 'medium' experience users, and two as 'high' experience users. The number of hours of experience using each device for each category is shown later in Table 4.

### *2.3 A 'real world' test*

A real-world experimental test sequence, rather than an abstract target acquisition test, was used to test the usability of the devices. The test consisted of a total of 150 simple tasks using a common direct manipulation graphical user interface, Microsoft Windows. The 150 tasks covered two domains, word processing with Microsoft Word and web browsing with Internet Explorer, and formed a natural flow of interaction. Approximately half of the tasks comprised the word processing sequence and half the web browsing sequence. Two different domains were used so that any performance differences caused by the

**Table 1.** First ten tasks of the test sequence

| Task Number | Task Description | Target | Action | Target Size (smallest VA° at 60cm) |
|---|---|---|---|---|
| 1 | Click the [Start] button on the task bar | Task bar button | Single click | 0.9° |
| 2 | Open the [Programs] menu by clicking the [Programs] icon on the start menu | Start menu entry | Single click | 0.9° |
| 3 | Start Word by clicking the [Microsoft Word] icon from the start menu | Start menu entry | Single click | 0.9° |
| 4 | Click the [WiViK soft keyboard] button on the task bar | Task bar button | Single click | 0.9° |
| 5 | Resize Word by double clicking the window title bar | Window title bar | Double click | 0.6° |
| 6 | Move the Word window to the top left of the screen by dragging the window title bar | Window title bar | Drag | 0.6° |
| 7 | Resize the Word window to fill the top 2/3 of the screen by dragging the bottom right size handle | Window size control | Drag | 0.9° |
| 8 | Click the [File] menu | Menu | Single click | 0.9° |
| 9 | Click the [Open] menu item | Menu | Single click | 0.9° |
| 10 | Double click the [Test File.doc] filename in the list box | List box item | Double click | 0.6° |

different nature of interaction in each domain could be identified. The proportions of interface object usage, interface 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 were involved in 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 60 cm from the screen. To give an example of the tasks comprising the test, a small section of the test sequence is shown in Table 1. 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°) and icons and on-screen 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.

### *2.4 Experimental design*

The experiment was an intra-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. To avoid order effects, these sessions were conducted in a balanced order. Prior to these sessions, each participant was given training and practice to become familiar with all of the test tasks so that they would be undertaken in a smooth flow of interaction, mimicking real-world interaction as closely as possible. 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 to perform the test tasks. Both devices require careful calibration to operate successfully, hence the quality of calibration achieved by the participants with the head mouse and eye mouse was recorded after device calibration and before each test by asking the subjects to point at nine equally spaced targets on the screen, with the overall mean distance of the cursor from the targets recorded. Tests were only conducted with calibrations exceeding 75% of the accuracy typically obtained by expert users with the devices, with calibration repeated until this level was reached. The time taken by repeat calibrations was included in the total calibration time for each test. Maintaining an acceptable level of calibration removed the possibility that an occasional poor calibration would affect the test results. The test was then conducted and the satisfaction questionnaire administered after each test session with a device.

### *2.5 Measuring device usability*

The usability of the two pointing devices was assessed in terms of objective device *performance* and subjective device *satisfaction* based on the European ESPRIT MUSiC performance metrics method [4, 19] and the recommendations outlined in the ISO 9241 Part 11, 'Guidance on Usability' International Standard [26]. From these, performance was defined as *the quality of interaction with the device and the time taken to perform that interaction*, and satisfaction as *the subjective acceptability of the device, expressed in terms of user workload and comfort when using the device and the ease of use of the device.*

#### 2.5.1 Performance

Device performance was calculated by measuring the quality of interaction during the tasks through counting the number of incorrect commands generated (such as selecting the wrong target), the number of target misses (with no command generated), the number of cursor position corrections (unnecessary positional 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). Note that a cursor position correction was defined as a path change of direction or deviation from a near straight line movement from cursor starting point to target end point, or as an unnecessary pause of cursor movement during the task. These variations and pauses indicate a lack of control when compared to an ideally 'perfect' cursor movement. Detailed discussion of cursor movement and positional correction is covered in [17]. Tasks were initially given a quality rating of 5 (perfect) [27], with subsequent errors reducing the quality score until the task was completed or failed and the next task started. To reflect the differing consequences of generating, and later recovering from, each error type, the quality factors were weighted, giving a simple formula for quality (Eq. 1). An error weighting approach to the quality metric was used to reflect the real-world impact differing errors had on interaction quality while still allowing valid relative comparisons between the different devices within the test [27]. Tasks were declared failed when the quality was reduced to 1. Task performance with each device was calculated as a percentage by a simple formula (Eq. 2) such that a task that had the highest level of quality and took no time would give a performance rating of 100%, with any reduction in quality or increase in time degrading the measured performance.

$$\begin{aligned}\text{Quality of interaction} = {} & \\ & 5 - (3 \times \text{count of incorrect commands} \\ & \quad + 2 \times \text{count of target misses} \\ & \quad + 1 \times \text{count of control corrections}) \qquad (1)\end{aligned}$$

$$\text{Task performance} = \frac{\text{Quality of interaction (1–5)}}{5 + \text{Time taken for interaction (secs)}} \qquad (2)$$

#### 2.5.2 Satisfaction

Participant satisfaction with the eye mouse and head mouse was measured using a multidimensional device assessment questionnaire based on the ISO 9241 Part 9, 'Non-keyboard Input Device Requirements' International Standard [26] and the N.A.S.A. 'task load index' workload questionnaire [9]. The test questionnaire consisted of three rating sections: workload, comfort and ease of use, each giving a multidimensional score comprised of ratings from the factors within each section (Table 2). The comfort and ease-of-use factors were chosen specifically to examine issues related to eye and head pointing device user satisfaction. Seven-interval fully labelled scales suitable for input device assessment were used for rating all of the individual questionnaire factors [2, 5].

### *2.6 Test analysis*

All data were obtained by capturing the complete contents of the test computer screen during each session, including the cursor position, with a screen capture program at a rate of five frames per second. The data were 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. Statistical comparisons were made using Mann-Whitney two-sample rank tests, with any significant differences ($p < 0.05$) indicated where appropriate.

**Table 2.** Satisfaction questionnaire factors

| Satisfaction Questionnaire Sections | Section Factors (each rated 1–7) |
|---|---|
| Workload | Physical effort<br>Mental effort<br>Temporal pressure<br>Frustration<br>Performance |
| Comfort | Headache<br>Eye comfort<br>Facial comfort<br>Neck comfort |
| Ease of use | Accuracy of pointing<br>Speed of pointing<br>Accuracy of selection<br>Speed of selection<br>Ease of system control |

## 3 Test results

### *3.1 Task domain and performance*

The test showed that there were no overall differences between the word processing and web browsing domain performance for each device, but there were differences in performance between the devices (Fig. 1). Grouping the word processing and web browser results together for each device gave a task performance of 65.2% for the head mouse and 51.1% for the eye mouse. This difference was statistically significant (Mann-Whitney Test, $U = 301\,228$, $n1 = 900$, $n2 = 900$, $p < 0.0001$), with the head mouse outperforming the eye mouse by 1.28 times. The performance for the hand mouse was the same over the two domains, and grouping the data gave a baseline performance of 83.3%, considerably higher than the assistive technology devices. The similarity of device performance between the word processing and web browsing task domains showed that the context or nature 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, and this likely accounted for the similarities in measured performance much more than the context of the tasks.

### *3.2 Target size and performance*

Examining the device performance by target size shows a distinct relationship between increasing task performance and increasing target size for both devices (Fig. 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 (Eqs. 3 and 4).

$$\text{Task performance for the head mouse} = 38.9 \times \text{target size} + 21.6 \quad (3)$$

$$\text{Task performance for the eye mouse} = 32.2 \times \text{target size} + 16.3 \quad (4)$$

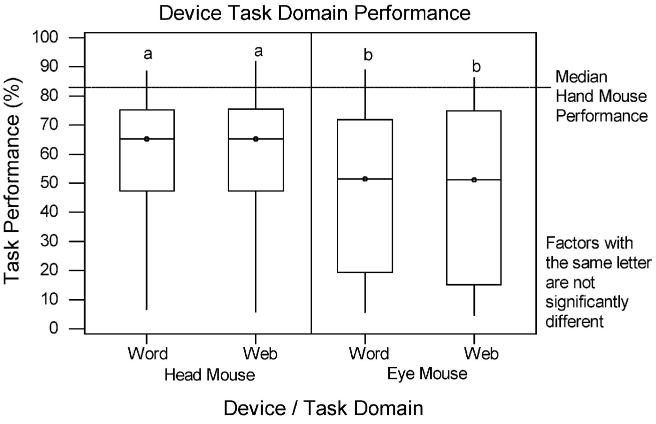

**Fig. 1.** Device task performance in different domains for both devices against hand mouse baseline

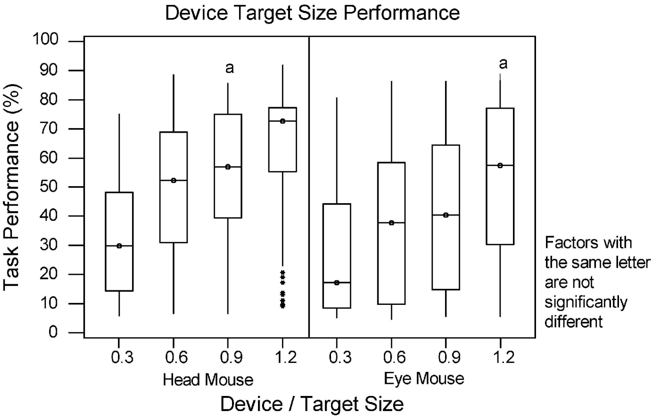

**Fig. 2.** Variation in task performance with target size for both devices

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 device performance will likely 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. For example, for the eye mouse to achieve the same performance as the head mouse at a target size of 0.3°, the target size for the eye mouse would need to be $((38.9 \times 0.3) + 21.6 - 16.3)/32.2 = 0.53°$. The results of these calculations for the target sizes in the test are shown in Table 3.

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, whereby 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.

3.2.1 Elements of target size and performance

Overall, the task quality and task time elements that constituted the performance metric showed steady increases in quality and steady reductions in task time with increas-

**Table 3.** 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% |

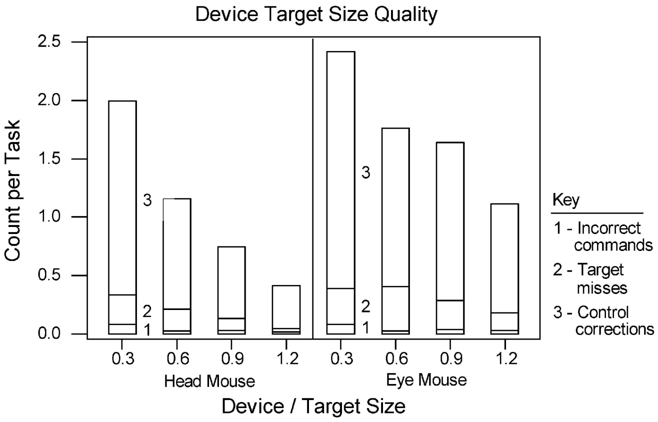

**Fig. 3.** Variation in task quality counts with target size for both devices

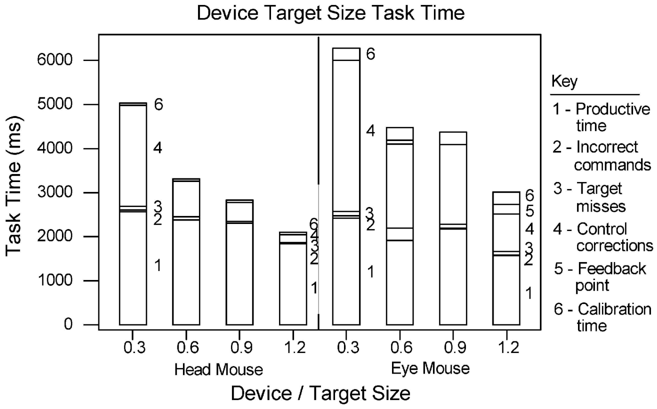

**Fig. 4.** Variation in task time elements with target size for both devices

ing target size. The head mouse was again superior to the eye mouse with the highest quality and lowest times for each of the test target sizes.

Figure 3 shows a breakdown of the individual quality component counts per task occurring for each target size, for tasks across both domains and for all participants. Here the task quality elements of the performance metric showed clear decreases in error counts per task and hence increases in quality, with increasing target size for both devices. It was notable that both devices had low counts of incorrect commands, suggesting that the devices could be accurate when the consequences of error, such as correcting the outcome of an incorrect command, are high. The higher counts of target misses for the eye mouse in comparison to the head mouse indicate some difficulty in maintaining the cursor over the intended target during selection. Review of the test videos confirmed that this effect was due to a ‘machine gun’ approach to selection with multiple selection button presses close to and around the intended target without hitting adjacent targets. The high rate of control corrections for the eye mouse in comparison to the head mouse indicates some considerable difficulty in manoeuvring and positioning the cursor onto a target – the rate is equivalent to more than one control correction per interaction. Of all the quality metrics, it is clear that the number of control corrections generated causes the most impact on the performance of the device.

Of the two elements that comprised performance, the relationship between task time and target size was perhaps more revealing (Fig. 4).

Here the task time was broken down into six time elements: productive time (the time cursor movement contributed to task completion), the time lost generating incorrect commands, target misses and cursor control corrections, the time lost while the eye mouse cursor was displaced looking at the feedback point on the interface, and finally the time lost during the calibration 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. 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.

Examining the elements of eye mouse quality and task time that could be reduced suggested that if target size continued to increase, then the control correction count and time lost during control corrections could be minimised, and the eye mouse would perform as well as the head mouse. It is most notable how these results show the performance dependency of these assistive technology devices on the size of the targets present on the computer interface.

### *3.3 Participant experience and performance*

There was a definite relationship between the experience rating of test participants with the devices and their performance and increasing experience resulting in increasing task performance (Fig. 5). Table 4 shows the numbers

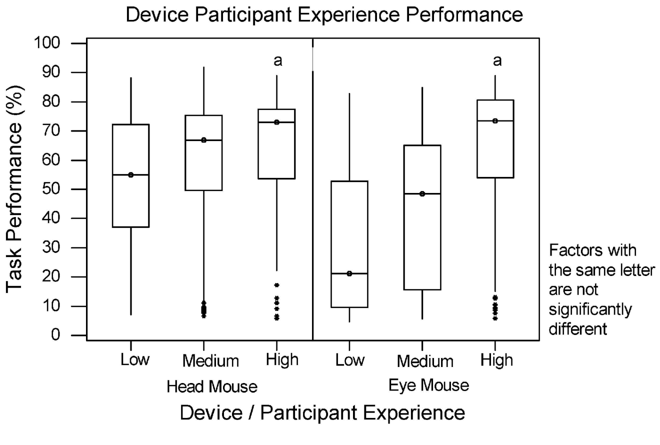

**Fig. 5.** The effect of participant experience on task performance for both devices

**Table 4.** 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 |

of hours of device usage accumulated by participants in each of the three experience categories.

Figure 5 shows that 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 of device experience for these participants is very different between devices, the data suggest that the eye mouse could approach the performance levels of the head mouse if participants are sufficiently practiced. In steady state, performance with either device is not likely to exceed the 83% performance baseline obtained from the hand mouse.

The performance of the medium group of head mouse users after 0.5 to 1 h experience is far higher than the low experience eye mouse group with 1 to 2 h 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.

### 3.3.1 Elements of experience and performance

Overall, 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 near parity for the highly experienced participants.

Examining the task quality and task time elements that comprised the performance metric (Fig. 6) showed steady increases in quality and steady reductions in task time with increasing participant experience.

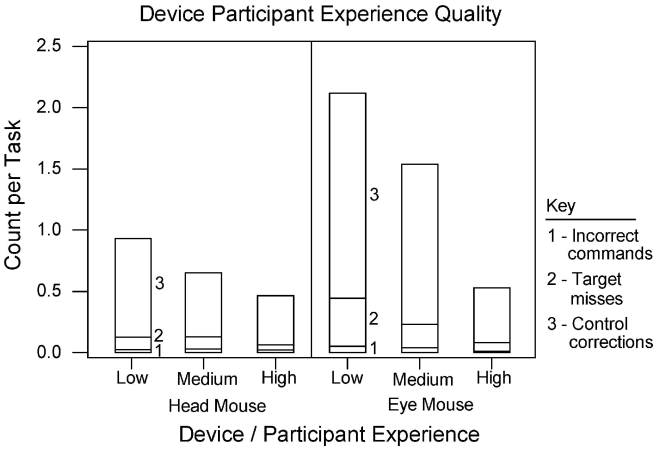

**Fig. 6.** The effect of participant experience on task quality counts for both devices

Examining the breakdown of the individual quality components for increasing participant experience showed steady decreases in error counts per task and hence increases in quality, with increasing experience for both devices. However, the overall rate of decrease in total error counts per task for the eye mouse is much more marked, with a 75.0% reduction from low to high experience compared to a 50.2% reduction for the head mouse. This rapid rate of reduction brought the eye mouse error count to near parity (within 10%) with the head mouse. It was notable that the overall proportions of each of the quality elements remained approximately constant as the overall count decreased with increasing participant experience, with control corrections dominating the quality metric irrespective of experience. Again, of the two performance elements, the relationship between task time and experience was most revealing (Fig. 7).

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

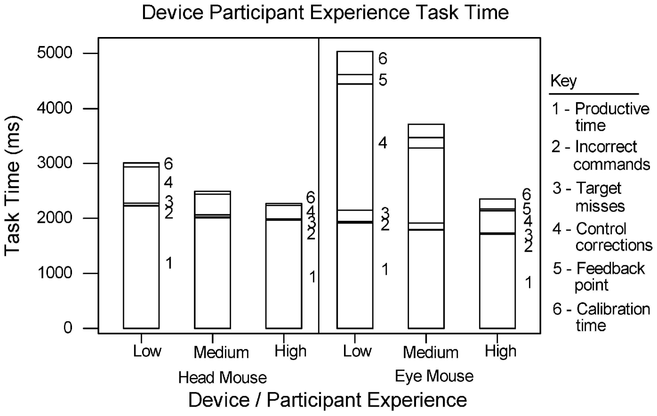

**Fig. 7.** The effect of participant experience on task time elements for both devices

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.

### *3.4 Device satisfaction*

Overall satisfaction ratings for each of the three questionnaire categories of workload, comfort and ease of use (Table 2) were calculated by taking the mean values of the individual ratings of the participants for each of the sections (Fig. 8).

Examining the workload questionnaire results showed that the eye mouse had a significant increase in workload that was rated 38.4% higher than the head mouse (Mann–Whitney Test, $U = 5, n1 = 6, n2 = 6, p < 0.0187$). In addition, the eye mouse was rated 24.2% less comfortable to use than the head mouse (Mann–Whitney Test, $U = 7,\ n1 = 6,\ n2 = 6,\ p < 0.0463$). These results indicated that the head mouse required less work to operate than the eye mouse, was more comfortable to use and was also probably the most sustainable to use over protracted periods of time. Such relatively poor subjective workload and comfort results may indicate some of the reasons why the eye mouse has not been widely accepted as a viable device. However, there was no significant difference in overall ease of use between the devices. This suggested that, although the eye mouse required more workload and was less comfortable, some individual factors of ease of use had brought the eye mouse ease of use to parity with the head mouse.

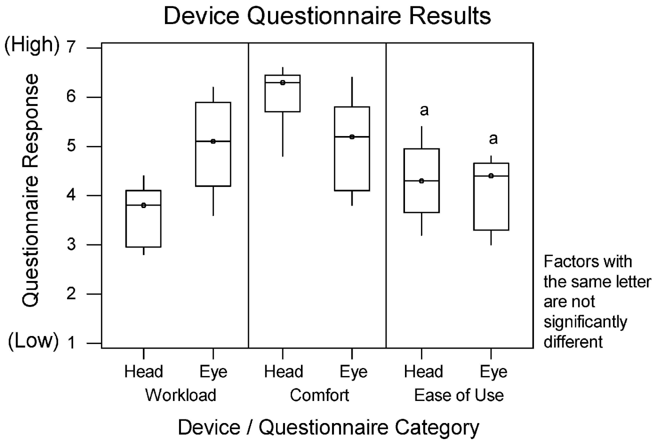

**Fig. 8.** Subjective questionnaire results for both devices

#### 3.4.1 Elements of satisfaction

The results of the individual factors are shown in Table 5 together with any differences between the eye mouse and head mouse subjective ratings.

Examining the individual factors for workload showed that the eye mouse exhibited consistently higher workload than the head mouse for all factors. This was particularly true for mental workload, indicating that a high degree of concentration was required for the eye mouse. However, it was notable that there was a considerably smaller difference between the devices with respect to performance. Examining the individual factors for comfort again showed consistently poor ratings for the eye mouse, and this was particularly notable for eye discomfort. There was a small reduction in difference between the devices for neck discomfort due to a low rating for the head mouse, indicating that the head mouse did cause some discomfort in operation. Taken together the individual workload and comfort ratings confirmed that the eye mouse caused considerable workload and discomfort during operation. Finally, examination of the ease-of-use ratings revealed the operational properties of the two

**Table 5.** Individual satisfaction factors

| Factor/Device | | Eye Mouse | Head Mouse | Difference (eye-head) |
|---|---|---|---|---|
| Workload (low=good) | Physical | 5.5 | 3.8 | +1.7 |
| | Mental | 5.7 | 3.8 | +1.9 |
| | Temporal | 4.3 | 2.7 | +1.6 |
| | Frustration | 5.0 | 3.7 | +1.3 |
| | Performance (inv.) | 4.7 | 4.2 | +0.5 |
| Comfort (high= good) | Headache | 5.5 | 6.5 | −1.0 |
| | Eye | 4.7 | 6.2 | −1.5 |
| | Facial | 5.0 | 6.2 | −1.2 |
| | Neck | 3.8 | 4.7 | −0.9 |
| Ease of Use (high=good) | Pointing Accuracy | 2.2 | 3.8 | −1.6 |
| | Pointing Speed | 4.5 | 3.5 | +1.0 |
| | Clicking Accuracy | 4.5 | 4.5 | 0.0 |
| | Clicking Speed | 5.1 | 4.5 | +0.6 |
| | System Control | 4.1 | 5.2 | −1.1 |

devices, with the eye mouse showing superior pointing speed (due to the rapid movement of the eye) and clicking speed (possibly due to some hand-eye co-ordination effects, even though both devices used the same selection mechanism), and the head mouse showing superior pointing accuracy and overall control (due to the ease of positional correction of the head).

## 4 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 (Fig. 9).

Here we see that the hand mouse was superior to the assistive technology devices, even when experienced participants used the devices. However, as the target size increased, so the performance of the eye mouse and head mouse approached that of the hand mouse. Examining the largest target size, the hand mouse had a performance of 83.3% compared to the eye mouse at 78.1% and the head mouse at 75.2%. At this point the eye mouse exceeds the performance of the head mouse, with a statistically significant difference, (Mann–Whitney Test, $U = 10\,906$, $n1 = 168$, $n2 = 168$, $p = 0.0003$) and reaches within a few percentage points of the performance of the hand mouse. This suggested that if target sizes were larger than those in this test and participants highly experienced, then the performance of an eye mouse might exceed a head mouse and approach that of a hand mouse.

Comparing only the experienced participant satisfaction questionnaire results for the eye mouse and head mouse revealed that the comfort rating remained approximately unchanged from 24.2% difference between the devices for all participants to 19.5% for experienced participants only. Clearly, the eye mouse is uncomfortable to use irrespective of experience. However, there were changes in the workload and ease-of-use ratings. The workload rating for the eye mouse moved from 38.4% higher than the head mouse for all participants to near parity at 5.4% higher than the head mouse for experienced participants only. Finally, the ease-of-use rating for the eye mouse moved from approximate parity with the head mouse for all participants to exceed the head mouse by 10.5% for experienced participants only. Some care should be taken with these results due to the small sample sizes for experienced participants; however, they do offer indications as to the effect of participant experience on the performance and satisfaction of the devices. This effect of experience on usability is currently being further investigated.

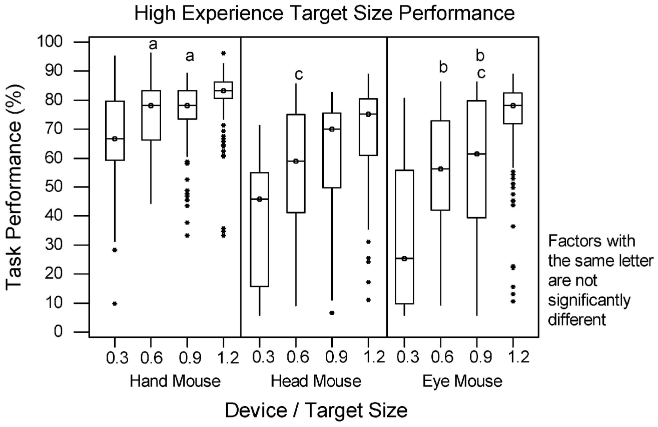

**Fig. 9.** The effect of high experience participants on task performance with different target sizes for all devices

## 5 Summary

The question this paper raised was why eye mice are 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 for the head mouse. This suggests that mechanisms to reduce the impact of target size on eye mouse performance could be used, such as providing the user with the ability to either zoom in on an area of the screen to effectively increase target sizes temporarily during interaction [1, 3] or to ‘snap-on’ to close targets [33]. However, controlling these enhancements could incur additional task time overheads. These mechanisms and the resultant effects on performance are subject to further investigation. 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 only occurred after 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 to a level greater than that found in the test if a user had extended head mouse experience up to the 30 h of the experienced eye mouse users. However, it was observed that the head mouse was extremely quick to learn. From this rapid learning time it was expected that head mouse performance would not improve significantly with extended experience.

The experiment has 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.

Examining participant satisfaction with the devices showed that the eye mouse was consistently more uncomfortable to use, required higher workload from the user and had an ease of use approximately equal to that of the head mouse. However, the effects of participant experience suggested that workload could be reduced to near parity with the head mouse and eye mouse ease of use could exceed the head mouse with user experience. There was little change in comfort with increasing experience.

To date we have found few comparable studies, but relating the results obtained in the experiment presented here to indications in other work [8, 12, 14, 18] seems to support the findings of rapid but inaccurate eye pointing and slower but more accurate head pointing.

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 and satisfaction of the head mouse and favour the eye mouse. Conversely, any reduction in the quality of eye control of the participants, due to nystigmus, for example, will degrade eye mouse performance and satisfaction in favour of the head mouse. The answer to the initial question, namely, why have eye mice not been adopted, is that they initially perform poorly, are uncomfortable to use and require higher workload from the user. An eye mouse requires slightly larger target sizes and higher user experience to perform well and at the same level with a head mouse. For novice users the eye mouse would manifest itself as unusable and unsatisfactory, 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 satisfactorily to some extent on the first attempt, making the system much more attractive. The experiment raises the question of how an interface specifically designed for eye mice would appear and how it would perform. Clearly, it would require larger target sizes to aid target acquisition, but this would have an impact on available screen area, probably giving a trade-off in usability between device performance and interface utility. In addition, such an interface would be non-standard and may not be usable with current commonly used applications. Perhaps methods to allow users to control target size or more easily locate targets on a standard interface could potentially further enhance the eye mouse without the need for specialist interfaces. 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. This paper has shown that with sufficient control over target size and with sufficient experience, eye mice can be used effectively and satisfactorily. For people with motor disabilities who dislike or have difficulties using a head mouse or who 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.

## References

1. Bates R (1999) Multimodal eye-based interaction for zoomed target selection on a standard graphical user interface. Proceedings of Interact'99, vol II. British Computer Society, pp 7–8
2. Bates R, Bierton R (2000) Experimental determination of quantifiers for usability questionnaire design. In: Proceedings of human-computer interaction 2000, vol II. British Computer Society, pp 57–58
3. Bates R, Istance HO (2002) Zooming interfaces! Enhancing the performance of eye controlled pointing devices. In: Proceedings of ASSETS 2002, the fifth international ACM SIGCAPH conference on assistive technologies, 8–10 July 2002, Edinburgh
4. Bevan N, Kirakowski J, Maissel J (1991) What is usability? In: Human aspects in computing: design and use of interactive systems and work with terminals. Elsevier Science, Amsterdam
5. Bierton R, Bates R (2000) Experimental determination of optimal scales for usability questionnaire design. In: Proceedings of human-computer interaction 2000, vol II. British Computer Society, pp 55–56
6. Carpenter RHS (1991) Eye movements. MacMillan
7. Edwards G (1998) New Software makes eye tracking viable: You can control computers with your eyes. In: Proceedings of CSUN conference on technology and persons with disabilities 1998, California State University, Northridge
8. Evans G, Blenkhorn P (1999) A head operated joystick – experience with use. In: Proceedings of the CSUN conference on technology and persons with disabilities 1999, California State University, Northridge
9. Hart SG, Staveland LE (1988) Development of the NASA-tlx (task load index): results of empirical and theoretical research. In: Human mental workload. Elsevier, Amsterdam, pp 139–183
10. Prentke Romich Company, HeadMaster Plus Head Mouse. http://store.prentrom.com/
11. Origin Instruments, HeadMouse. http://www.orin.com/access/
12. Jacob RJK (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):152–169
13. Jacob RJK (1995) Eye tracking in advanced interface design. In: Advanced interface design and virtual environments. Oxford University Press, Oxford, pp 258–288
14. Jagacinski RJ, Monk DL (1985) Fitts' law in two dimensions with hand and head movements. J Mot Behav 17:77–95
15. LoPresti E, Brienza DM, Angelo J, Gilbertson L (2000) Computer head controls: ergonomics and effect of neck movement limitations. In: Proceedings of CSUN 15th annual conference on technology and persons with disabilities, California State University, Northridge
16. LoPresti E, Brienza DM, Angelo J, Gilbertson L, Sakai J (2000) Neck range of motion and use of computer head controls. Proceedings of Assets 2000: the fourth international ACM conference on assistive technologies, pp 121–128
17. MacKenzie IS, Kauppinen T, Silfverberg M (2001) Accuracy measures for evaluating computer pointing devices. In: Proceedings of CHI 2001, ACM Press, New York
18. MacKenzie IS (1992) Fitts' law as a research and design tool in human-computer interaction. Hum-Comput Interact 7:91–139

19. Macleod M, Bowden R, Bevan N, Curson I (1997) The MUSiC performance measurement method. In: Behaviour and information technology 16(4/5):279–293, Crown, New York
20. Polhemus Incorporated, 40 Hercules Drive, P.O. Box 560, Colchester, VT 05446, USA. www.polhemus.com
21. Prentke Romich Company, 1022 Heyl Road, Wooster, OH 44691, USA. www.wivik.com
22. Saito S (1992) Does fatigue exist in a quantitative measurement of eye movements? Ergonomics 35(5/6):607–615
23. Salvucci DD, Anderson JR (2000) Intelligent gaze-added interfaces. In: Proceedings of CHI 2000, vol 2, ACM Press, pp 273–280
24. Sibert LE, Jacob RJK (2000) Evaluation of eye gaze interaction. In: Proceedings of CHI 2000, vol 2. ACM Press, pp 281–288
25. SensoMotoric Instruments GmbH (SMI), Warthestraße 21, D-14513 Teltow/Berlin, Germany. www.smi.de
26. Smith WJ (1996) ISO and ANSI ergonomic standards for computer products: a guide to implementation and compliance. Prentice-Hall, New York
27. Szczur M (1994) Usability testing – on a budget: a NASA usability test case study. Behav Informat Technol 13(1/2): 106–118
28. Boost Technology, Tracer HeadMouse. http://www.boosttechnology.com/
29. Velichkovsky B, Sprenger A, Unema P (1997) Towards gaze mediated interaction: collecting solutions of the “midas touch” problem. In: Human-computer interaction: INTERACT’97, Chapman & Hall, New York
30. Virzi RA (1992) Refining the test phase of usability evaluation: How many subjects is enough? Hum Factors 34(4): 457–468
31. Ware C, Mikaelian HH (1987) An evaluation of an eye tracker as a device for computer input. In: Proceedings of CHI 1987, ACM Press, pp 183–188
32. Yarbus AL (1967) Eye movements and vision. Plenum, New York
33. Zhai S, Morimoto C, Ihde S (1999) Manual and gaze input cascaded (MAGIC) pointing. In: Proceedings of CHI 1999, ACM Press, pp 246–253