Behavior Research Methods, Instruments, & Computers
2002, 34 (4), 509-517

Eye movements of large populations: I. Implementation and performance of an autonomous public eye tracker

DAVID S. WOODING, MARK D. MUGGLESTONE, KEVIN J. PURDY, and ALASTAIR G. GALE
University of Derby, Derby, England

This paper details the design and construction of an autonomous public eye tracker exhibit, which was installed at the National Gallery, London, in 2000/2001. For over 3 months, it functioned both as an informative exhibit and as a controlled eye movement experiment, gathering data from over 5,000 participants. The issues associated with automatic unattended recording of the eye movements of members of the public are discussed. The performance of the exhibit is examined, and its successes and problem areas are highlighted with regard to potential applications and future exhibits. The success of the project proves the viability of autonomous public eye trackers as both data-gatherers and public exhibits.

Background to the Project

This paper describes the design and construction of an eye-tracking exhibit and experiment that was part of a public exhibition entitled "Telling Time" at the National Gallery, London, from October 18, 2000 to January 14, 2001. Designed and built by the Applied Vision Research Unit of the University of Derby, the exhibit functioned as a controlled eye movement experiment on a large number of participants, providing a wealth of data on how individuals look at images in general and at paintings in particular. Its aim also was to illustrate to members of the public that viewing a picture is not an instantaneous process but that our understanding of the image develops as we stand before it.

Although there have been eye-tracking exhibits in museums and science centers around the world (see, e.g., Buquet, Charlier, & Paris, 1988; Glenstrup & Engell-Nielsen, 1995; and see Appendix A), these have focused on the eye tracker as a public science *exhibit*, rather than being used to gather scientific data.

The Demands of the Exhibit

The exhibit as experiment. The exhibit was designed to function as a scientifically rigorous experiment that would record the eye movements of large numbers of individuals as they viewed images of paintings. Traditional eye movement experiments use relatively small numbers of participants. With a large number of participants, the data could be generalized widely, with each stimulus being observed by more than 100 participants.

The experiment as exhibit. To encourage participation, the exhibit was as interactive as possible. A large "public" display alongside the exhibit was controlled by the same computer that was running both the experiment and the eye-tracking system and displayed the live eye movements of the current participant. Each participant's eye movements were also replayed while the next participant was undergoing the necessary calibration process. When no current data were available, the public display replayed prerecorded traces or showed more complex representations of eye movement data.

Participants. The participants were self-selected visitors to the exhibition. Data from 1995/1996 show approximately equal numbers of men and women visitors to the Gallery, peaking in the 15- to 34-year-old group. Recording eye movements is more difficult in older participants, so a preponderance of the young among visitors was favorable.

The project was approved by the University Ethics Committee, and the self-selecting participants were informed of the purpose of the experiment through carefully worded static displays accompanying the exhibit.

Requirements of the design. As exhibit visitors, the participants were typically "naive," with no prior knowledge of eye movements or of the purpose of the experiment. Consequently, in use, the exhibit needed to be intuitive and user-friendly, with minimal per-participant time, while maximizing the experimental and exhibition properties.

The main challenge of the exhibit was to record eye movements automatically from a wide range of people without being intrusive, intimidating, or uncomfortable. Recording eye movements typically requires a skilled human operator, but the autonomous system had to operate unattended and with minimal operator intervention for the 3-month duration of the exhibition.

The exhibit was staged in cooperation with various staff at the National Gallery, London. Alexander Sturgis was the curator of the "Telling Time" exhibition, of which the exhibit formed a part. Applied Science Laboratories (ASL) provided the 504 eye tracker and technical help during development. The project was supported by Derby University Enterprises Limited (DUEL), and the analysis was funded in part by the British Academy (Grant 31757). Lyndsey Cobb contributed to preparatory studies. Correspondence concerning this article should be addressed to D. S. Wooding, Applied Vision Research Unit, Kingsway House (West), University of Derby, Derby DE22 3HL, England (e-mail: d.wooding@derby.ac.uk).

The software provided for totally unattended operation of the exhibit, powered up once daily by the museum staff. The software controlled experimental flow—for example, selecting stimuli that had not been used recently. It also had to control the eye tracker hardware and record reliable point-of-gaze data, as well as being designed robustly to allow for different participant behavior—for example, leaving at any point during the process.

Finally, the exhibit needed to be both functional and reliable, so as to avoid public disappointment and to take advantage of this unique opportunity to obtain data from such a large population.

METHOD

The Eye Tracker

Choice of eye-tracking system. Previous attempts at the public siting of eye trackers generally have used custom-built systems that required the development of an eye-tracking system, in addition to the software and hardware associated with the running of the exhibit. Reports of the performance of such exhibits (Buquet et al., 1988; Glenstrup & Engell-Nielsen, 1995) have highlighted some of the problems with this approach, specifically in relation to system reliability and usability.

The approach adopted here was to integrate a commercially available eye-tracking system with the exhibit requirements. Although such systems may have proven reliability, they require a skilled operator. The challenge was to replace the operator with dedicated software that would allow autonomic operation.

There are a number of eye movement recording techniques (see Young & Sheena, 1975, for a review), but within applied environments video-based systems are often used, since these permit the participant some degree of movement and enhance usability. These systems illuminate the eye with infrared (IR) light that is parallel to the axis of a camera viewing the eye. The incident light reflects both off the retina, making the pupil appear bright, and off the cornea, the front surface of the eye. These reflections are processed to enable the calculation of point of gaze with a typical accuracy of approximately 1° of visual angle.

The requirements of the eye movement system were, therefore, that it be remote (i.e., no direct contact with the participant), controllable by external software (the experiment would require low-level control of the eye movement system functionality that would normally require an operator), robust, and reliable.

An ASL 504 remote system (Applied Science Laboratories [ASL], Waltham, MA) was used. This uses a small video-conferencing camera, modified to include IR illuminators, and is controlled by a base unit linked to a PC via an RS232 serial connection. Point-of-gaze data are calculated at 50 Hz. The base unit incorporates a data buffer that ensures that no data are lost even if the controlling PC cannot maintain an adequate data read speed, which was important since the PC was multitasking. ASL provided a Dynamic Link Library so that software could be developed to access functionality normally available to an operator.

The presence of spectacles can produce additional IR reflections, which can be mistaken for eye features. Since no operator was present, a sign requested wearers to attempt the experiment without spectacles, if their visual acuity allowed it.

Natural light was excluded from the exhibit room, enabling the level of ambient visual and IR illumination to be carefully controlled. In addition, spotlighting was used to ensure that the exhibit was in shadow while maintaining adequate lighting for other exhibits in the room.

Design and Equipment

Height of participants. The system was designed to operate with a participant's eye spatially located at a particular height (1,225 mm) from the floor. This height was the lowest at which a tall adult could reasonably be seated and the lowest to which smaller people could feasibly be raised. A height-adjustable stool (375–690 mm) was constructed, enabling people from age 8 upward to use the exhibit. The stool was removable to allow the exhibit to be accessible to some wheelchair users.

Eyepiece. A key issue was to make the participant's eye accessible to the eye-tracking hardware in a reliable way. After evaluating a number of *locating* devices, a spectacle-shaped eyepiece was chosen (Figure 1), fixed in space relative to the eye movement system. This solution was both intuitive and minimally intrusive to use and was practical in a public setting, since physical contact with the participant was minimized. The eyepiece delivered the participant's eye within a known area and at a set distance from the *participant* display, avoiding the need for the eye tracker to locate the eye spatially. The eyepiece also minimized movement by participants during the experiment, reducing the chance of the recording hardware's losing the image of the eye.

A low powered IR beam across the front of the eyepiece was broken when a participant positioned him/herself at it. This alerted the system that a participant was present and also meant that the participant had to come close to the eyepiece, locating the eyes properly, before the experiment would start.

Computer and displays. The exhibit software was developed in Microsoft Visual Basic (v 6.0) running under the Microsoft Windows 98 operating system (which has the ability to use simultaneous multiple displays connected to one PC). A dual display system allowed onlookers to view live eye movements as they were recorded. A Pentium III 450-MHz PC was used, with 128 MB RAM, 10 GB hard disk, and fast graphics cards (Voodoo 3 and XPERT 98). Two displays were used: a Hitachi SuperScan 813 (Model CM813U, 21-in. diagonal with 20-in. viewable) for presentation of stimuli to the participant and a Fujitsu Plasmavision plasma display (Model 4222, 42-in. diagonal) to present the appropriate images to onlookers. Both displays ran at a resolution of 1,024 × 768 pixels and 32-bit color depth.

Hands-free operation and eye buttons. Hands-free operation of the system was desirable, since it would minimize the possibility of the participants' moving away

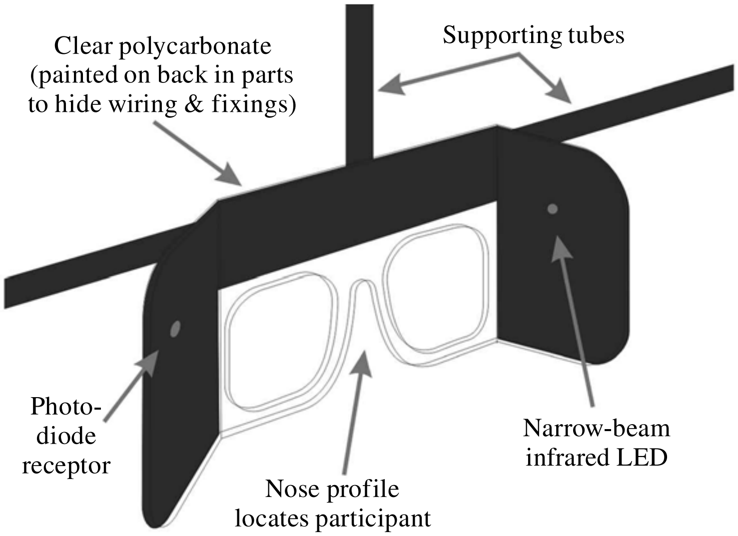


Figure 1. The eyepiece.

from the eyepiece to operate controls, creating difficulties for eye movement recording. However, for experimental reasons, it was useful to obtain information (e.g., gender, age, artistic training, etc.) on participants who had successfully completed the eye movement calibration. This was achieved by using fixation locations to operate *eye buttons* (see also Glenstrup & Engell-Nielsen, 1995; Goldberg & Schryver, 1995; Jacob, 1991; Sibert & Jacob, 2000). A question would be posed near the top of the participant's monitor, with two or three graphical buttons appearing beneath bearing possible answers. The participant simply looked at their preferred button, which changed color. If the participant was still fixating the same button after 2 sec, that answer was accepted, and the next question was displayed.

Stimulus selection and preparation. The Gallery specially prepared 239 stimuli. Images were converted from MARC scan (Cupitt, Martinez, & Saunders, 1996; Saunders, Cupitt, White, & Holt, in press) digital copies of the original paintings, ensuring that resolution and color fidelity were maximized. The stimuli were saved as JPEG format image files in 24-bit color and with a maximum size of 1,024 × 768 pixels (corresponding to 22.5° × 16.2° at the viewing distance of 900 mm). The color and contrast of the participant display were set with feedback from a gallery curator who had detailed knowledge of the paintings from which the stimuli were produced.

The images were divided into eight image sets of around 35 images, and the software used only one set at a time. The presentation order of the images in a set was randomized each time all the images in that set had been presented, so that each image in the set would be seen by the same number of participants. The set size was calculated so that an onlooker would be unlikely to see the same image twice unless he/she spent over half an hour at the exhibit. Extensive pilot studies determined that a presentation time of 20 sec per image was sufficient to allow participants to view the images without losing interest.

Identification of fixations and saccades. The eye tracker delivered point of gaze as a set of coordinates every 20 msec, and these raw data were stored to disk for later analysis. These data were simplified into fixations and saccades in order to communicate the nature of eye movements to onlookers via the public display. Few algorithms for the identification of fixations and saccades from raw eye position are documented in the literature (Jacobs, 1986; Karsh & Breitenbach, 1983; Stampe, 1993), and an algorithm was developed for this exhibit that was based on the method used by ASL. This algorithm groups data points into fixations primarily on their spatial distribution, using a criterion based on the allowable standard deviation of the distribution (see Appendix B). The locations of fixations were represented on the public display by circles around the point of fixation, of diameter equivalent to 1° on the participant display. Saccades were represented as straight lines connecting the fixations.

Mode of Operation

Appearance. Visitors to the exhibit were faced with a large open-fronted "booth" containing the spectacle-shaped eyepiece suspended at the opening of a rectangular recess (Figure 2). At the rear of the recess was the participant display. There was an adjustable stool underneath the eyepiece.

The experimental sequence began when a participant presented at the eyepiece. The participant display presented a welcome message and then instructions on how to take part in the experiment. While the participant read the instructions, the system carried out the eye location and discrimination. This was one of a number of parallel processes designed to minimize the experimental time.

Eye location. As soon as the participant arrived at the eyepiece, the system began the calibration procedure. The initial stage was to center the participant's eye in the field of view of the eye movement camera. This procedure

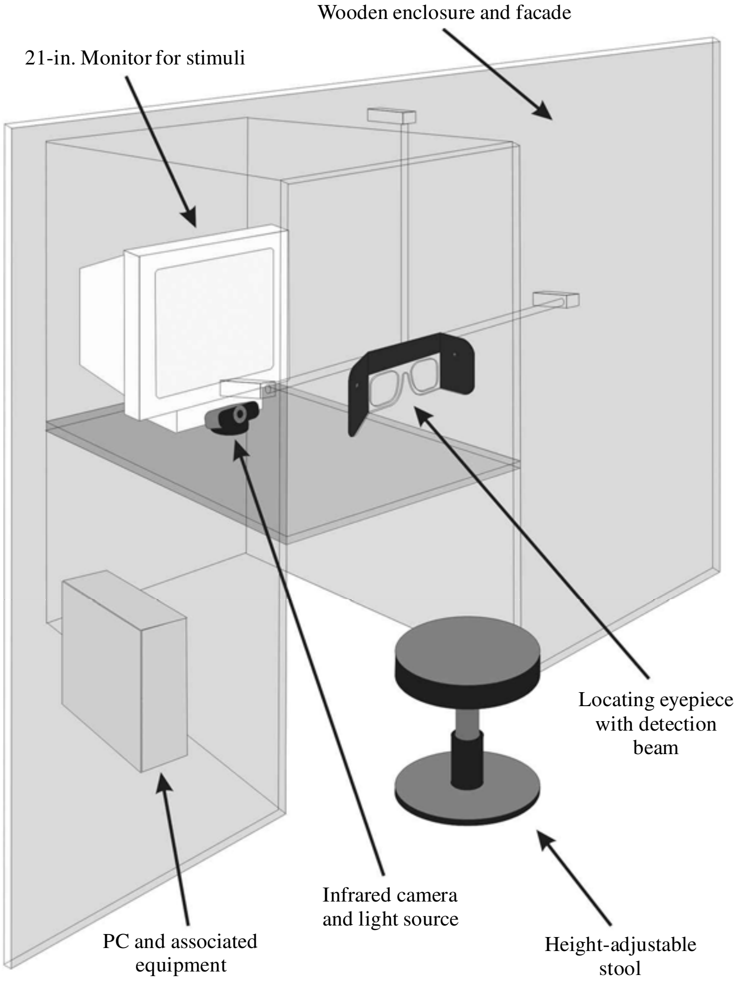


Figure 2. The physical layout of the exhibit.

meant that even with some head movement, the system would still have a usable image of the eye. The possible range of the centering procedure was limited to the dimensions of the eyepiece, to ensure that the system could not center on a spurious external reflection.

Eye discrimination. After the centering procedure, the next stage was to recognize the pupil image and corneal reflection. The luminance of each person's pupil image varies according to pupil size, which is related to age (Winn, Whitaker, Elliott, & Phillips, 1994) and other factors. The corneal reflection may not be discernible if it occurs within a very bright pupil image, and if the pupil is too dark the pupil may not be recognized efficiently. In the exhibit, the aperture of the camera in the eye movement system was adjusted, using a staircase method (Cornsweet, 1962), until the brightness of the pupil image was within an optimum range (determined from pilot studies). Once the brightness of the image had been adjusted, the system then attempted to recognize the pupil and corneal reflections within the image. A staircase method was used to identify the range over which the system could recognize both pupil and corneal reflection. The midpoint of this range was used as the final value.

Eye calibration. The participant was then requested to undertake a simple calibration procedure that mapped the relationship of the corneal reflection and the pupil (Merchant, 1974). The 9-point calibration technique (Sheena & Borah, 1981; Stampe, 1993; Yamada, Fukuda, & Hirota, 1990) was made more user-friendly by replacing the usual calibration points with a small image of a "bee," which, flapping its wings, took a circuitous route around the nine locations. The participants were simply instructed to "follow the bee."

When the calibration bee was stationary over one of the calibration points, 10 samples of eye tracker data were accessed by the control software. These data were checked to ensure the presence of pupil and corneal reflections and the integrity of the data. If the data were valid, the mean values of the 10 points (for the relationship between pupil and corneal reflections) were used as calibration data for that point. The bee then moved to the next point until all 9 points had been completed. All of the raw eye tracker data were also saved, allowing post hoc examination of the efficacy of the calibration and point-of-gaze algorithms.

Accuracy test. If the calibration was successful, the participant followed the bee to an additional four points to provide a measure of calibration accuracy. Data from these points were stored to allow rejection or post hoc adjustment of the participant's data in the event of poor calibration. In the event of failure to localize, discriminate, or calibrate, the software informed the participant that normal variation in eye characteristics meant that eye movements could not be recorded from everyone.

Participant questions. After the accuracy check, demographic information was collected by participants responding to three simple questions, using eye buttons. Each image set had its own three questions associated with it, to ensure both that a sufficient number of participants answered each question and that the same questions were associated with each image.

Display of stimuli. Three images were then presented to the participant for 20 sec each, preceded by a central cross that the participant fixated and that was used in the same manner as the eye buttons to activate display of the stimulus.

On some occasions, a prompt was displayed before the fixation cross, which consisted of text containing certain information about the subsequent image. This was used to investigate how the provision of information about a stimulus before its presentation could affect the eye movements observed when the stimulus was subsequently viewed (Yarbus, 1967).

The participant was then requested to examine the images, and while he or she did so, his or her eye movements were relayed live to the public display on the gallery wall. After the third image, the participant was thanked and asked to leave so that the next person could take part. The participant had to leave the exhibit in order to view his/her own eye movements replayed on the public display, encouraging throughput of participants. The number of images and the characteristics of their presentation, together with the eye movement calibration and the questions put to each participant, were all designed so that the total time a participant was seated at the exhibit should be between 3 and 4 min, with processes occurring simultaneously wherever possible.

RESULTS AND DISCUSSION

The Experiment

Performance. The experiment ran successfully for 89 days. During this time, 120,000 people visited the exhibition, and 9,884 participants took part in the experiment, of which 5,446 completed the full experimental run, viewing all three images. The maximum number of different participants attempting the experiment in a day was 179.

In the event, the software restricted the number of image sets to four in order to ensure that each stimulus was presented to a minimum of 100 participants. As a result, approximately half the images (140) were presented.

The median participant time was 232 sec, of which the median time for instructions (and hidden pupil thresholding) was 42 sec, the time for calibration was 60 sec, and the time for questions and presentation of stimuli was 130 sec. This approximates to 60 days and, together with time spent seating and reading instructions, indicates that the exhibit was in constant use during the hours in which the Gallery was open.

Figure 3 shows how many participants were "lost" in the different parts of the experimental process. Of the total participants, 3,890 (39%) did not complete the calibration process. Of those who did calibrate successfully, 379 (4%) did not have a sufficiently good calibration to operate the eye buttons or the fixation cross.

Some 5,615 participants (57%) viewed at least one of the images, with 5,524 (56%) viewing two and 5,446 (55%) completing the full experimental run of three images. The system recorded that the 169 participants who failed to view all three images left the eyepiece during or between the presentation of stimuli (at a rate of 1% per image).

Participant calibration. No difference was found in the accuracy of fixation for the four accuracy points, with an overall median discrepancy between recorded position and actual position of 0.7°. Accuracy point data for all the participants were examined to determine whether there was a consistent coordinate shift in either a vertical or a horizontal direction. Nineteen percent had an overall horizontal shift, and 21% had an overall vertical shift. In the course of a normal eye movement experiment, a human operator would routinely introduce an offset to account for such shifts. An adjustment could be made to the exhibit data to account for the translation in measured point of gaze and, thereby, improve accuracy. This will be the subject of further work.

The participants fixated a central cross before each stimulus presentation, and the overall median difference between recorded eye position and the cross center was found to be 1.1° in the fixation at image presentation. The difference between the error in fixation of accuracy points and the fixation of the central cross may be explained by the larger central cross's offering a larger target for fixations (resulting in a greater spread of recorded point of gaze) and by the limitations of calculating accuracy from a single fixation. There was no difference between results for the three stimulus presentations, since *drift* throughout the course of an experiment does not occur with this type of eye movement recording technique.

The total given in this paper for the number of different participants failing a calibration is not the same as the total failed number of calibrations, which was 5,184. The analysis in this paper is concerned with the number of participants and so excludes those failed calibrations for which the time between participants was less than 1 sec. It was impossible for 2 participants to swap places in a second, so shorter intervals must represent a second attempt at calibration by a participant who has already failed calibration. The participants demonstrated a surprising degree of determination to become calibrated, since this accounted for 1,294 (25%) of the total failed calibrations and is certainly an underestimate. With these cases excluded, the number of *different* participants failing calibration was 3,890.

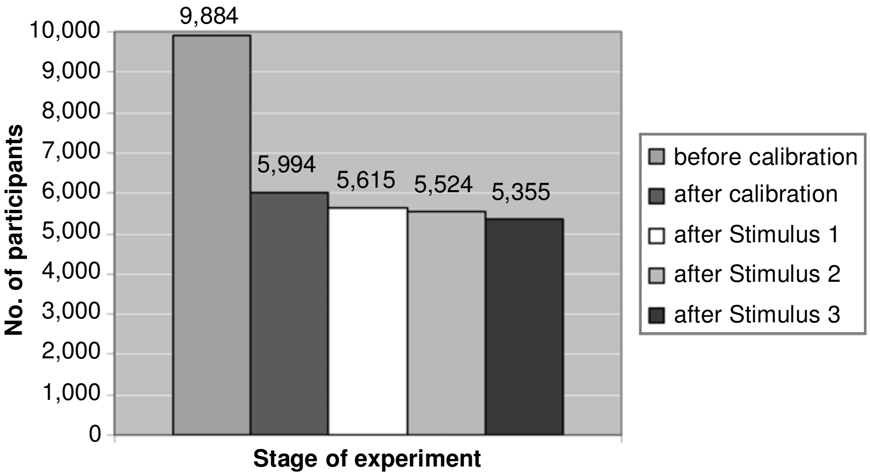


Figure 3. The number of participants at each stage of the experiment. The figure shows (reading from left) the number of participants remaining before and after calibration and, then, after each of the three stimuli had been presented.

The maximum number of participants calibrated in a day was 122. The median daily success rate was 65%, with the maximum daily success rate being 84% (85 participants calibrated and 16 failures). The proportion of successful to failed calibrations varied more widely than was expected on a daily basis, and we conclude that there was an unidentified intermittent problem with part of the system.

The participants may have failed to calibrate for a number of reasons—for example, the system's failing to recognize small pupils or problems with contact lenses and spectacles. The main reason was poor recognition of either the pupil or the corneal reflection. This is supported by the data in Figure 4, which shows that the majority of failed calibrations happened within the first 30 sec of calibration. Inappropriate values for the initial recognition criteria would mean that the pupil or the corneal reflection was not recognized when the participant was fixating on one of the nine points of the calibration pattern. Another problem is that the brightness of the pupil image (i.e., the amount of the reflected light) can vary as a participant looks at different parts of the participant display. This brightness change will lead again to an inability of the system to recognize one of the required elements. Limited observation of the exhibit suggested that this was an increasingly likely problem for young participants (whose generally larger pupils produce a brighter image), although over 50% of the successfully calibrated participants were in the 15- to 35-year age range. The recognition thresholds used were developed from extensive pilot testing, but since these tests were on a different population than the exhibit participants, this may also have contributed to the number of unsuccessful calibrations.

To increase the proportion of successful calibrations would require a more sophisticated recognition system. The system may have either hardware or software components but would probably have to contain a more intelligent process that could deal with changes in the image of the participant's eye.

The Exhibit

Public perception and usability. The novelty of the exhibit proved a great attraction, with the public display drawing people to the exhibit and people queuing to take part. This highlighted the importance of having clearly visible instructions sited with regard to possible queues, so that those waiting could familiarize themselves with the procedure *before* sitting down in front of the eye tracker.

People appeared to find the exhibit easy to use, with no real difficulty in understanding the instructions. Even novel concepts such as the eye buttons worked well, and people found them surprisingly enjoyable to operate. In retrospect, more questions could have been included.

The IR beam built into the eyepiece proved extremely effective in detecting the presence of a participant's head. It was not uncommon, though, for participants to break the beam several times while positioning themselves. Fortunately, the majority of these *new participant* signals could be filtered out, but the function of the beam remained a compromise between correctly positioning participants in the eyepiece (ensuring that the eye tracker received a good and reliable image of the eye) and determining their presence. If the head moved back a little during the session, the system would register that they had left the eyepiece, and the session would end, so the beam was placed conservatively, offering less optimum positional results. A more complex arrangement might contain two beams: one for detection and one for location.

In our observations of the exhibit in operation, it was apparent that participants often failed to use the adjustable stool as requested—that is, to change its height to one more suitable to them. This may well have caused poor positioning of the eye at the eyepiece or positioning that

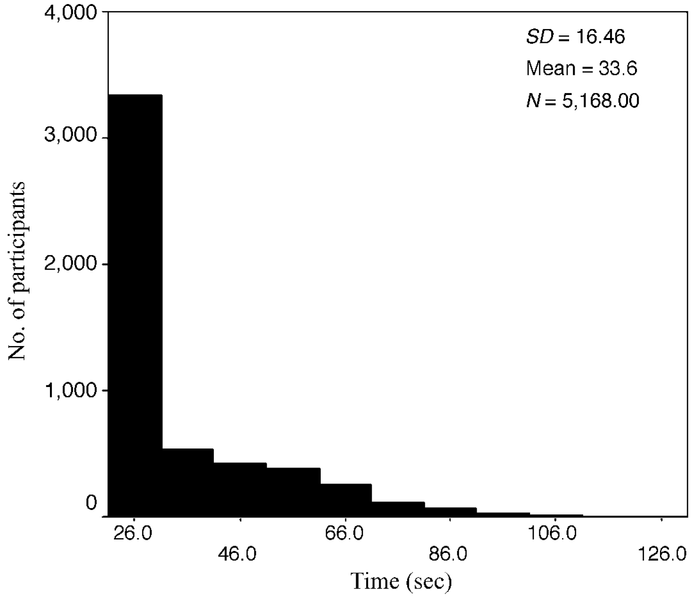


Figure 4. Calibration time to failure. For all participants who failed the calibration process, the figure shows the distribution of times taken to fail.

was uncomfortable for the participant and so could not be maintained over the duration of the experimental run.

Operation of software. A lot of the novel solutions to technical problems worked very well (e.g., using a dual display system). Pseudo-multitasking in the software (e.g., public replay of a previous participant's eye movements during the next participant's calibration) succeeded in increasing participant throughput and maintaining levels of interest.

CONCLUSIONS AND FUTURE WORK

The exhibit was successful, both as a large-scale eye movement experiment on the general public and as an interesting exhibit. It also demonstrated that such automatic recording of eye movements is possible.

Glenstrup and Engell-Nielsen (1995) have reported data on the performance of other public eye trackers. Buquet et al. (1988), in Paris, attracted a lower rate of participants (9,000 in around 21 months), but early trials claimed a success rate of 84.3%. It is not clear whether this performance was maintained or what accuracy of data was obtained. Glenstrup and Engell-Nielsen also found difficulties with the calibration of the EyeCatcher exhibit in Hellerup (Appendix A).

Current eye-tracking hardware will never reach 100% performance in terms of participant calibration, particularly if high standards of spatial resolution are demanded (as in this experiment). For now, researchers will have to accept that a certain proportion of participants will be lost in the calibration process. How this is handled, in terms of messages to the unfortunate participant and accompanying displays, will make a huge difference to how the exhibit as a whole is perceived.

There is clearly the potential for a full-time exhibit in a public place, such as a gallery or museum, which might gather data for any number of eye movement experiments on any range of stimuli. We have shown that it is possible to use and adapt commercially available systems to this purpose, although it may be necessary to add components in order to extend the use of the system or minimize problems.

Since eye-tracking equipment has traditionally required skilled operators, eye movement research has relied on small numbers of participants. Publicly located, autonomous eye trackers may be the answer to obtaining the large data sets from which new findings will arise. Analysis of the large amount of data from the National Gallery exhibit is underway.

REFERENCES

Buquet, C., Charlier, J. R., & Paris, V. (1988). Museum application of an eye-tracker. *Medical & Biological Engineering & Computing*, **26**, 277-281.

Cornsweet, T. N. (1962). The staircase method in psychophysics. *American Journal of Psychology*, **75**, 485-491.

Cupitt, J., Martinez, K., & Saunders, D. (1996). A methodology for art reproduction in color: The MARC project. *Computers & the History of Art*, **6**, 1-20.

Glenstrup, A., & Engell-Nielsen, T. (1995). *Eye controlled media: Present and future state.* Unpublished minor subject thesis, Laboratory of Psychology, University of Copenhagen. (Available at http://www.diku.dk/users/panic/eyegaze/)

Goldberg, J. H., & Schryver, J. C. (1995). Eye-gaze determination of user intent at the computer interface. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 491-502). New York: Elsevier.

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**, 152-169.

Jacobs, A. M. (1986). Eye-movement control in visual search: How direct is visual span control? *Perception & Psychophysics*, **39**, 47-58.

Karsh, R., & Breitenbach, F. W. (1983). Looking at looking: The amorphous fixation measure. In R. Groner, C. Menz, D. F. Fisher, & R. Monty (Eds.), *Eye movements and psychological functions: International views* (pp. 53-64). Hillsdale, NJ: Erlbaum.

Merchant, J. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Transactions on Biomedical Engineering*, **21**, 309-317.

Saunders, D., Cupitt, J., White, C., & Holt, S. (in press). The MARC II camera and the scanning initiative at the National Gallery. *National Gallery Technical Bulletin*, **23**, 76-82.

Sheena, D., & Borah, B. (1981). Compensation for some second-order effects to improve eye position measurements. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 257-268). Hillsdale, NJ: Erlbaum.

Sibert, L. E., & Jacob, R. J. K. (2000). Evaluation of eye gaze interaction. In *Proceedings of ACM CHI 2000 Human Factors in Computing Systems Conference* (pp. 281-288). New York: Addison-Wesley/ACM Press.

Stampe, D. M. (1993). Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems. *Behavior Research Methods, Instruments, & Computers*, **25**, 137-142.

Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science*, **35**,1132-1137.

Yamada, M., Fukuda T., & Hirota, M. (1990). Development of an eye movement analyzer possessing functions for wireless transmission and autocalibration. *Medical & Biological Engineering & Computing*, **28**, 317-324.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum.

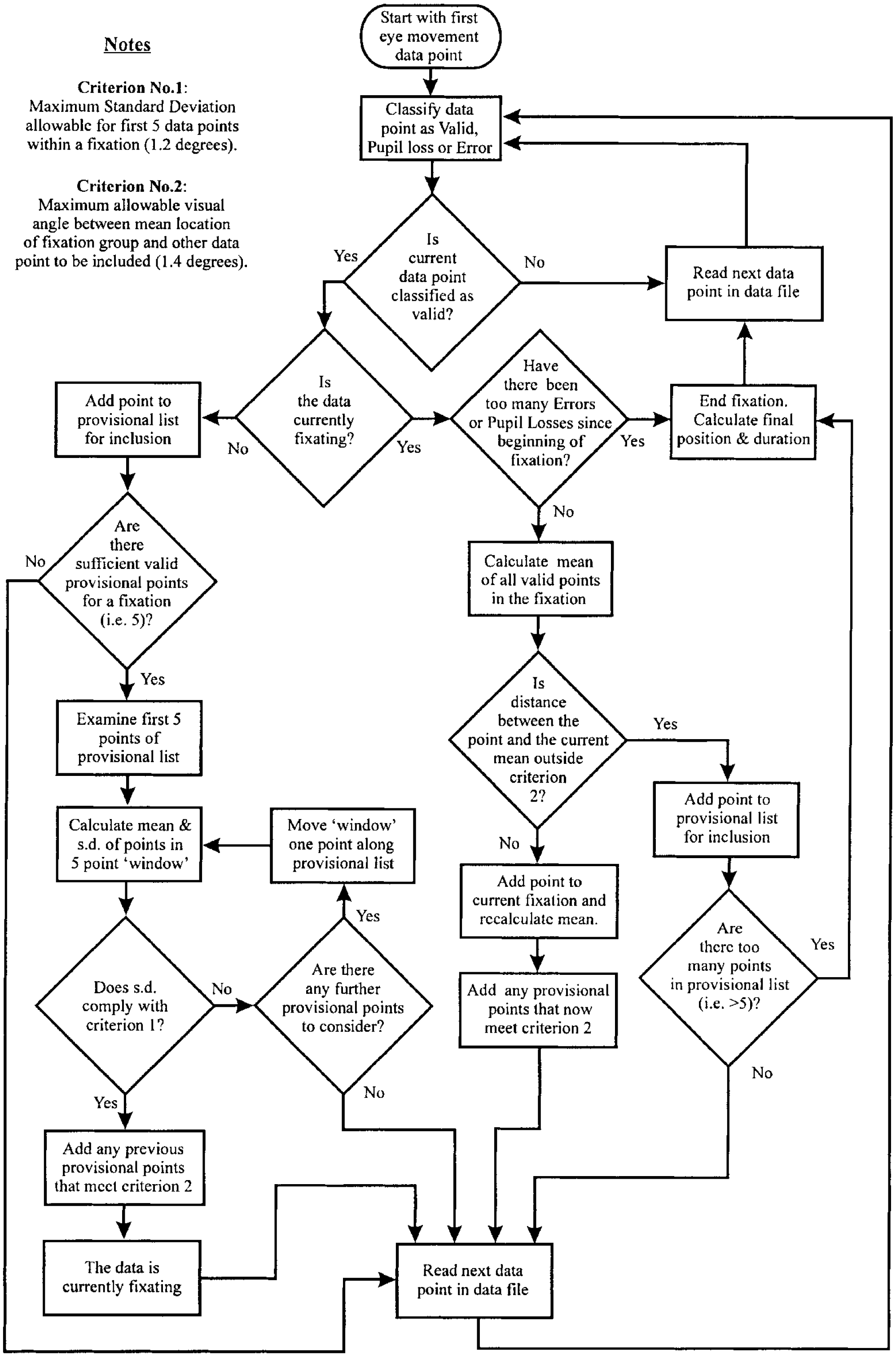
Young, L., R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, **7**, 397-429.

APPENDIX A
Current Examples of Publicly Sited Eye Trackers

The Eye-Follower (Cité des Sciences et de l'Industrie 30, avenue Corentin-Cariou, 75930 Paris cedex 19). The original eye-tracker was installed in 1986 by Buquet et al. (1988). The current tracker, running for more than 10 years, is based around a Metrovision (http://www.metrovision.fr) eye tracker. http://www.cite-sciences.fr/english/ala_cite/expo/explora/expressions_comportements/expr_10.htm.

The EyeCatcher (Experimentarium, Tuborg Havnevej 7, DK-2900, Hellerup, Denmark). Constructed by the Risø National Laboratory at Roskilde in Denmark and running since 1995. http://www.experimentarium.dk/uk/udstillinger/se_paa_lyset/opstilling.298.5.html

APPENDIX B
The Algorithm Used to Identify Fixations and Saccades



Notes

Criterion No.1:
Maximum Standard Deviation allowable for first 5 data points within a fixation (1.2 degrees).

Criterion No.2:
Maximum allowable visual angle between mean location of fixation group and other data point to be included (1.4 degrees).

(Manuscript received March 20, 2002;
revision accepted for publication October 14, 2002.)