Routledge
Taylor & Francis Group

# Eye tracking research and technology: Towards objective measurement of data quality

## Eyal M. Reingold

Department of Psychology, University of Toronto, Mississauga, ON, Canada

Two methods for objectively measuring eye tracking data quality are explored. The first method works by tricking the eye tracker to detect an abrupt change in the gaze position of an artificial eye that in actuality does not move. Such a device, referred to as an artificial saccade generator, is shown to be extremely useful for measuring the temporal accuracy and precision of eye tracking systems and for validating the latency to display change in gaze contingent display paradigms. The second method involves an artificial pupil that is mounted on a computer controlled moving platform. This device is designed to be able to provide the eye tracker with motion sequences that closely resemble biological eye movements. The main advantage of using artificial motion for testing eye tracking data quality is the fact that the spatiotemporal signal is fully specified in a manner independent of the eye tracker that is being evaluated and that nearly identical motion sequence can be reproduced multiple times with great precision. The results of the present study demonstrate that the equipment described has the potential to become an important tool in the comprehensive evaluation of data quality.

Over the past several decades there has been a dramatic increase in the quantity and variety of studies employing eye tracking technology. This is illustrated in Figure 1 that displays the number of peer reviewed articles over the past 50 years (grouped into successive 5-year bins) containing the phrase "eye tracking" and/or "eye movements". The rapid advancement in this field occurred in part due to the development of an increasing number of dedicated software and hardware tools that greatly facilitated eye tracking research activity by both expert and novice users. Currently, eye tracking systems are routinely employed across an incredibly broad and interdisciplinary spectrum of both basic and applied research paradigms (for reviews, see Duchowski, 2003; Holmqvist et al., 2011; Liversedge, Gilchrist, & Everling, 2011).

Despite the popularity of eye tracking technology, there has been surprisingly little emphasis in the literature on the development of tools and methods for evaluating the fidelity with which the continuous variation in the eye movement signal is reflected in the values measured and reported by the eye tracker (henceforth referred to as *data quality*). Unfortunately, although experts in the field have conducted their own rigorous evaluations of data quality requirements relevant to specific eye movement paradigms, these typically remain unpublished (see Inhoff & Radach, 1998, for a related discussion). A notable exception to this trend is reflected in a seminal paper by McConkie (1981) that provided a comprehensive introduction to the complexity of the data quality topic. More recently, Holmqvist and his colleagues (2011; Holmqvist, Nyström, & Mulvey, 2012) provided an excellent description of key concepts related to eye tracking data quality and strongly advocated renewed focus on this topic. The goal of the present paper is to further explore and illustrate tools that might permit individual researchers to objectively evaluate the aspects of data quality that
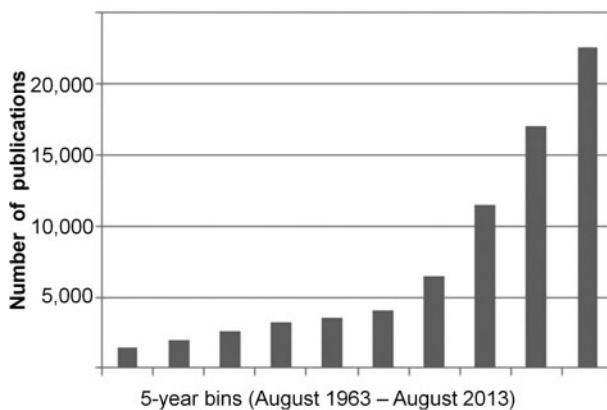


Figure 1. The number of peer reviewed articles in ProQuest databases over the past 50 years containing the phrase "eye tracking" and/or "eye movements" (grouped into successive 5-year bins).

are relevant to their scenarios of use (i.e., within their specific paradigms, and hardware and software setup). Accordingly, a few basic concepts related to data quality are briefly reviewed before a rationale for the proposed approach is presented. Finally, two methods for objectively evaluating data quality are described and illustrated.

With respect to the evaluation of data quality, it is important to distinguish between *accuracy* and *precision*, two concepts that are often mistakenly treated as equivalent in the eye tracking literature. Given repeated measurement of an actual (true) value, accuracy is often defined as the mean difference between the measured and true value. In contrast, precision, also referred to as reproducibility or repeatability, is the degree to which the repeated measurement of a set of true values produces the same or similar set of measured values regardless of the accuracy of these values. Consequently, measurement can be accurate but not precise; precise but not accurate, both accurate and precise or neither accurate nor precise. Figure 2 illustrates the difference between accuracy and precision by plotting the results of repeated measurement of a single true value by two different instruments (labelled A and B). As can be seen by an inspection of this figure, there are two fundamentally different types of error that characterize
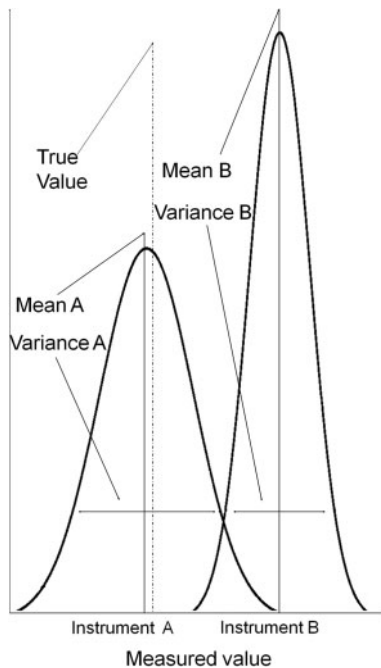


**Figure 2**. Illustration of the repeated measurement of a single true value by two different instruments: Instrument A (on the left) reflects better accuracy but poorer precision than Instrument B (on the right).

measured values relative to the true value. Specifically, normally distributed *random* error produces larger variance and poorer precision in the distribution of values measured with Instrument A than in the distribution of values measured with Instrument B. However, when comparing the mean of the two distributions to the true value it is clear that the distribution corresponding to Instrument B displays poorer accuracy than the distribution corresponding to Instrument A. This is shown in the figure by a consistent overestimation bias in the distribution on the right, referred to as *systematic* error that is larger in magnitude than the slight underestimation characteristic of the distribution on the left.

How can we apply these concepts to the evaluation of eye tracking accuracy and precision? Eye tracking is the process of acquiring the spatial information corresponding to the movement of the eye as it unfolds in time. The continuous spatiotemporal eye movement signal is discretely sampled at a given rate (henceforth referred to as the *sampling rate*). At a minimum, each sample data should contain both spatial information (e.g., gaze coordinates) and temporal information (e.g., timestamp). The eye tracking accuracy is dependent on the extent to which the eye motion pattern that is reconstructed from the eye tracker data matches the actual (true) eye motion signal. In addition, the variability in tracked motion data given multiple repetitions of the same true eye motion signal can be used to quantify eye tracking precision. Thus, a prerequisite for a rigorous and objective evaluation of the accuracy and precision of an eye tracking system (i.e., the measurement of data quality) would seem to be the requirement for an exact specification of the true eye movement signal in a manner that is independent of the eye tracker that is being evaluated.

Two strategies could be followed to confront this formidable challenge. The first approach involves the selection of an extremely accurate and precise eye tracking system as a kind of *gold standard* to be used for simultaneous recording together with the eye tracker that is being evaluated. The data from such a system would then serve as a proxy for the true eye movement signal. It has been proposed that the *search coil* system might serve as proper benchmark in simultaneous recordings with video-based eye trackers (see Kimmel, Mammo, & Newsome, 2012, for such a comparison). However, every known type of eye tracking technology is susceptible to artifacts and the search coil system is no exception. The search coil has been shown distort the kinematics of eye movements (e.g., Frens & van der Geest, 2002; Traisk, Bolzani, & Ygge, 2005; van der Geest & Frens, 2002) and in human subjects the slippage of the contact coils is known to produce errors in the reported spatial position (e.g., Collewijn, van der Mark, & Jansen, 1975). The second approach to the requirement for an exact specification of the true eye movement signal is based on the proposal by McConkie (1981) that data quality could be evaluated "by monitoring the movements of an artificial eye that can be accurately moved at different rates" (p. 99). While stationary artificial eyes are routinely employed in the evaluation of eye tracker precision (i.e., by measuring the variability in reported spatial

position given a stationary true eye motion signal), we are yet to realize McConkie's proposal of a fully controllable moving artificial eye capable of mimicking the spatiotemporal properties of biological eye motion. The main goal of the present investigation was to explore the feasibility of using such a computer controlled physical model of the moving eye in order to evaluate eye tracking data quality.

It is important to consider the use of artificial motion for eye tracking data quality evaluation in the context of the current practice of using actual biological motion for that purpose. Prior to a typical evaluation of eye tracker spatial accuracy and precision with human observers a conversion of raw eye tracker data to screen coordinates must be performed by using a mapping function. This function is derived from a process known as calibration, in which a set of targets in known positions on the screen are displayed, and subjects are asked to fixate (i.e., direct and hold their gaze) on the centre of the target. The corresponding eye tracker position data is recorded for each calibration point and the mapping function is derived. Following calibration, the eye tracker data typically contains gaze position in screen coordinates. Accuracy is then computed as the average distance in degrees of visual angle between the gaze position reported by the eye tracker and the centre of the fixated targets. In addition, precision is computed based on the variance in eye tracker gaze position data (e.g., standard deviation [*SD*] and root mean square [RMS]; see Holmqvist et al., 2011, for a review) when subjects repeatedly fixate the same points. Calibration schemes reported in the literature use anywhere from 3 points to 25 points. Importantly, the choices of a mapping function and the calibration process have a powerful impact on the nature and magnitude of systematic errors that are introduced, thereby influencing accuracy.

One major problem with using human observers to evaluate accuracy and precision is the fact that the entire procedure is based on the assumption that eye tracker data points are collected when observers are looking at the centre of fixation targets. This assumption is clearly false as observers have limited control in accurately directing and holding their gaze and limited awareness for their actual gaze position. In one sense the procedure as described earlier shares some of the classic problems of any method which primarily depends on subjective report. Another problem with attempting to measure eye tracking data quality with human observers is that the procedure used is typically limited to gathering data while subjects are fixating static targets. Ironically, devices purported to be eye movement monitoring systems are evaluated under conditions in which the eye is assumed to be motionless. In fact, even during fixations, the normal eye is never still due to miniature eye movements (i.e., microsaccades, drift, and tremor). The emphasis on the tracking of "stationary" biological or artificial eyes as a primary method for eye tracker data quality evaluation has the unfortunate consequence of creating an incentive for manufacturers to produce systems that use heavy filtering (i.e., denoising algorithms) that, while making the eye look

stable during fixations, severely distort the eye movement signal in terms of the velocity profile of the motion. Although appearing to improve static accuracy, it is often not appreciated that such filtering destroys important aspects of data quality including the temporal accuracy of identifying the beginning and end of fixations, the number of fixations detected (see Holmqvist et al., 2012), the kinematics of saccadic eye movements, the ability to detect small saccades, and eye movements produced while looking at dynamic stimuli (e.g., smooth pursuit).

Thus, the main potential advantage of using artificial motion for testing eye tracker data quality is the fact that it permits tight control over the input to the tracker and an ability to repeat the same input multiple times. The broader framework for such an approach is the functional testing method (commonly referred to as black box testing) that is used to evaluate software and hardware systems. Functional testing requires no knowledge of the inner working of the system that is being evaluated. Rather, given specification of system require-ments by users, relevant and well-specified inputs are used and the system's output is recorded and analysed. The focus of the present paper is on evaluating the potential usefulness of two devices for producing artificial pupil motion. The first device is designed without any attempt to mimic the nature of biological eye motion and is only meant to trick the eye tracker to detect an abrupt change in the gaze position of the artificial pupil (henceforth referred to as an *artificial saccade generator*). In contrast, the second device is designed to be able to provide the eye tracker with motion sequences that closely resemble biological motion. In the remainder of this paper each of these devices are described and their potential utility is illustrated by examining the data produced by a video-based eye tracker (an EyeLink 1000 Plus system, SR Research Ltd) using these two methods to generate various inputs.
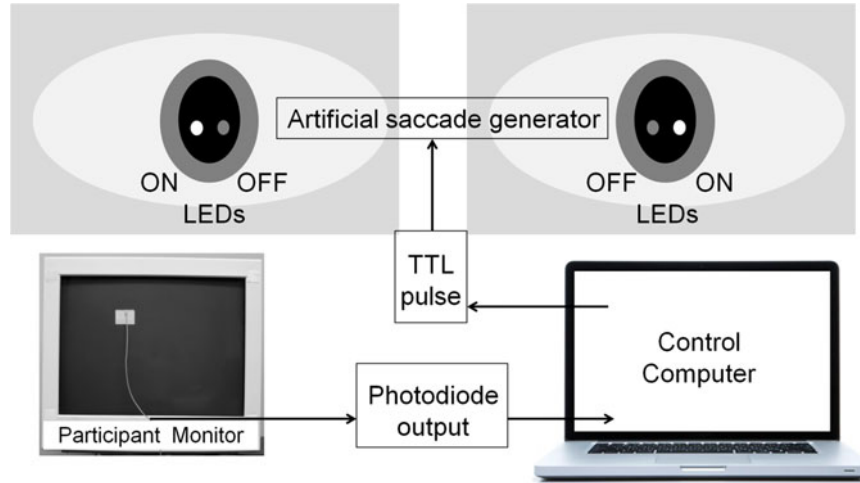
## ARTIFICIAL SACCADE GENERATOR

An artificial saccade generator is typically designed to test a critical functional requirement for gaze contingent paradigms. Specifically, gaze contingent methodology (see Rayner, 1998, 2009, for reviews) often requires minimal latency between the onset of an eye movement (e.g., a saccade) and onset of a display change. For example, in a gaze contingent moving window technique (McConkie & Rayner, 1975), the stimulus is modified (typically degraded) except in an experimenter-defined area (i.e., the window). The position of the window is continuously updated as a function of the observer's gaze position (e.g., the window could be constantly moved to be centred on the last known gaze position). A long latency between eye movements and display change would create a noticeable lag in the motion of the window resulting in undesirable consequences for both theoretical and applied implementations of this paradigm

(see McConkie, Wolverton, & Zola, 1984; McConkie, Zola, Wolverton, & Burns, 1978; Reingold, Loschky, McConkie, & Stampe, 2003, for related discussions). Versions of artificial saccade generators have been used frequently by eye movements researchers that extensively employ gaze contingent methodology in their laboratories (often without any published mention of their use; but see Bernard, Scherlen, & Castet, 2007; Holmqvist et al., 2012; Richlan et al., 2013, for descriptions of such devices). Although narrow in focus, an artificial saccade generator is an excellent example of functional testing of data quality that is motivated by a clear and influential scenario of use (e.g., gaze contingent paradigms).

The specific saccade generator used in the present paper is shown in Figure 3. In most video-based eye trackers, gaze position is computed based on processing the position of the pupil in the camera image (e.g., by determining the image coordinates corresponding to the pupil centre) as well as the position of a corneal reflection (CR) which is produced by an infrared (IR) illumination source. As illustrated in Figure 3, the artificial saccade generator used in the present study employs a printed pattern of an eye to provide a stable pupil target for the eye tracker. In addition, two light emitting diodes (LEDs) are embedded behind the pupil. Only one of the two LEDs is powered on at any time, providing the eye tracker with a CR target. Given the combination of pupil and CR inputs, the eye tracker data should reflect a stable spatial position (i.e., an artificial fixation). A control computer switches between the on/off LEDs producing an abrupt change in the position of the CR target and providing an artificial saccade input for the eye tracker. In such a setup, unlike with a human observer, the exact timing of the physical eye movement (i.e., saccade) can be ascertained with submillisecond precision independently of the eye tracker. The average latency or delay between the onset of the artificial saccade and the availability of the first eye tracker data sample which reflect the change in spatial position can be used as a measure of temporal accuracy (sometime referred to as the system *end-to-end delay*), and variability in this latency constitutes a measure of temporal precision.

However, given that the functional testing scenario was defined based on the requirements of gaze contingent techniques, it is far more important to evaluate the latency to display change than the latency to eye tracker data availability. In other words, in this scenario, as is the case in most experimental setups, the eye tracker interacts with a variety of other software and hardware systems (e.g., commodity computers and operating systems, device drivers, response/input devices, graphics engine and stimulus generation and display software and hardware, other intrusive software such as antivirus software, and physiological or neuroimaging recordings; see Plant & Quinlan, 2013, for a review). These additional factors often have an influence on the temporal precision and accuracy of the entire experimental setup that far outweighs the properties of the eye tracker in isolation.

**Figure 3**. Illustration of an artificial saccade generator: a printed pattern of an eye with 2 LEDs placed behind pinholes in the pupil (the black inner circle in the pattern which could also be white if a bright pupil rather than a dark pupil eye tracker is evaluated). The eye tracker acquires the pupil and detects the LED that is powered on as a corneal reflection (CR). The control computer produces a TTL pulse that is used to switch between the on/off LEDs, moving the CR position and generating an instantaneous artificial saccade. A photodiode is attached to the surface of the participant's monitor and a display change (increasing the luminance of the area underneath the photodiode) is initiated as soon as the eye tracker output reflects the change in gaze position due to the artificial saccade. The output from the photodiode is then used by the control computer to calculate the latency between the onset of the saccade and the physical display change.

To illustrate this issue, the computer that controls the artificial saccade generator using a parallel port signal (henceforth, output TTL) also received input from a photo diode that was attached to the surface of the participant's monitor (21-inch Viewsonic G225f CRT monitor). A custom program written in C language using the GDI graphics engine turned a small area underneath the photo diode (50 × 50 pixels) from black to white pixels. This display update was initiated as soon as a change in eye tracker data occurred in response to the artificial saccade. The photo diode signal was then logged via a second parallel port installed in the control computer (input TTL). Latency to display change was computed as the difference between the timing of the input TTL and the output TTL. Two conditions were used during the testing of latencies. In the first condition the frame rate of the CRT monitor was set to 160 Hz (160 frames/s) and in the second condition the standard 60 Hz setting was used. In each condition, a thousand artificial saccades were used and the eye tracker was configured to provide unfiltered data at a sampling rate of 2000 Hz.

Figure 4 shows the distribution of latencies from the onset of the artificial saccade to the onset of the display change. As shown in Figure 4, the frame rate of the monitor had a dramatic effect on the temporal accuracy and precision of the overall setup. Specifically, for the 160 Hz condition average latency was 4.82 ms, approximately half of the average latency in the 60 Hz condition (mean = 9.69 ms). In addition, there was a substantially larger variance in the latter than
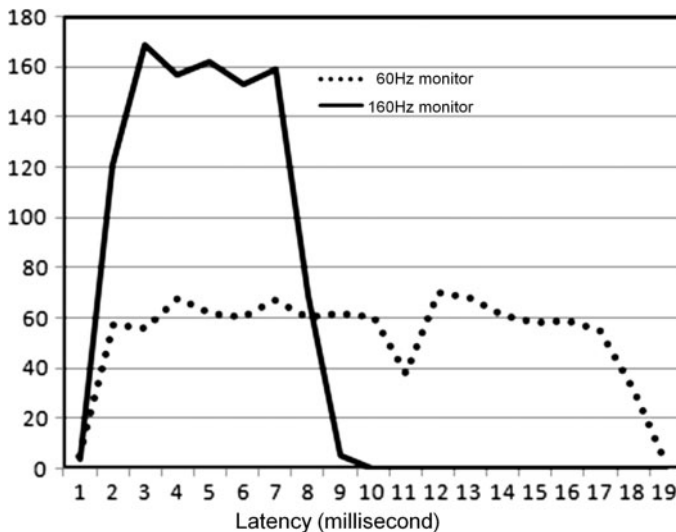


Figure 4. The distribution of latencies (in milliseconds) from the onset of the artificial saccade to the detection of the display change by the photodiode for 160 Hz monitor (solid line, mean = 4.82, $SD$ = 1.86) and a 60 Hz monitor (dotted line, mean = 9.69, $SD$ = 4.79). Y-axis values represent the number of saccades (out of a 1000) in any given one millisecond time bin (shown on the x-axis).

the former condition (60 Hz, $SD$ = 4.79; 160 Hz, $SD$ = 1.86). Importantly, the eye tracker end-to-end delay for the both conditions was measured at only 1.62 ms ($SD$ = 0.28). Consequently, given the enormous variability in setups across different laboratories (and even between different setups in the same lab), an important advantage of using a simple artificial saccade generator such as the one used in the present paper is that it can be employed by an individual researcher to test the delay in their specific setup, eye tracker type and configuration, and experimental paradigm (see Plant, Hammond, & Turner (2004), for a similar functional testing approach). For example, when choosing a monitor for a gaze contingent experiment it is not enough to examine the advertised frame rate. This is because high frame rate does not always translate into lower latency as some displays or projectors internally use a buffer of one or more frames thereby increasing the update delay.

## ARTIFICIAL MOVING PUPIL THAT CAN MIMIC BIOLOGICAL MOTION

To be useful for mimicking biological eye motion, an artificial eye must be mounted on a moving platform capable of both horizontal and vertical rotation with maximum angular velocity greater than 500 deg/s and maximum angular acceleration greater than 25,000 deg/s$^2$. In addition, the moving platform must exhibit minimal inertia and drift, and be capable of reproducing motion sequences with low latency and high spatial resolution and precision. Schneider and et al. (2009) developed a potentially suitable computer controlled moving platform and demonstrated that, by driving the platform based on real time input from an eye tracker, the moving platform can be successfully used to rotate a small camera or robotic eyes (Kohlbecher, Bartl, Bardins, & Schneider, 2010) to mimic the eye movements of a human observer. In the present application, rather than using eye tracking data as an input to drive the moving platform, an artificial eye was mounted on the moving platform in order to provide well-specified and reproducible motion sequences, which closely mimic biological eye motion, to serve as inputs for the evaluation of eye tracking data quality.

The setup used in the current paper is shown in Figure 5. The moving platform used was the EyeSeeCam platform developed by Schneider and colleagues (see Schneider et al., 2009, for a detailed description of the EyeSeeCam device). An artificial pupil was mounted in the centre of the platform and a laser diode was affixed above the pupil for the purpose of projecting a visible trace during platform motion (see Figure 5). In order to produce the desired motion sequences by the artificial pupil, a custom program written in C language was used to send commands via a serial link (RS-232) using a fast binary communication protocol. In addition, in order to provide a precise measure of the onset of the platform's motion, an accelerometer was attached to the back of the moving plate, and once
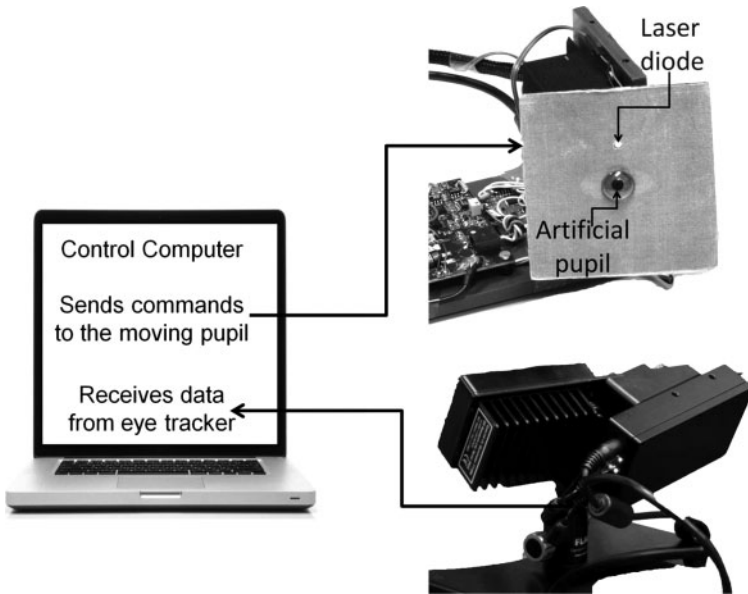
**Figure 5**. Illustration of the moving artificial pupil setup: The computer controls in real-time a modified EyeSeeCam platform (Schneider et al., 2009). An artificial pupil and a laser diode were attached to the front surface and an accelerometer (not shown) is attached to the back surface of the platform. This computer controlled platform can be used to move the pupil and/or the laser projection in a manner that closely mimics biological eye motion. The eye movement monitoring system (EyeLink 1000 Plus, SR Research Ltd) tracks the moving artificial pupil and sends real-time data to the control computer. Both the eye tracker and the moving platform were clamped to an optical bench in order to minimize the impact of environmental or mechanical vibrations.

motion was detected a signal was sent to the control computer via a parallel port. With this addition of the accelerometer, the moving artificial pupil could be used to evaluate temporal data quality in the same manner as described previously with the artificial saccade generator.

The primary focus of the present exploration was on the unique potential of using the EyeSeeCam moving platform to test the spatial precision of eye tracker data by repeating a variety of motion sequences that mimic biological motion. Accordingly, the moving artificial pupil was programmed to produce (1) 175 repetitions of a 2.5 seconds motion sequence taken from eye movement recording during reading, and (2) 245 repetitions of a 2.5 seconds smooth pursuit sequence of a target executing horizontal sinusoidal motion at a frequency of 0.4 Hz. In addition, the impact of filtering on spatial precision was assessed by configuring the eye tracker (EyeLink 1000 Plus, SR Research Ltd) to produce either unfiltered gaze position data or filtered data (with two-sample delay) at 1000 Hz sampling rate. Figure 6 displays the results from these recordings. A visual inspection of plots superimposing 175 reading traces (Panel
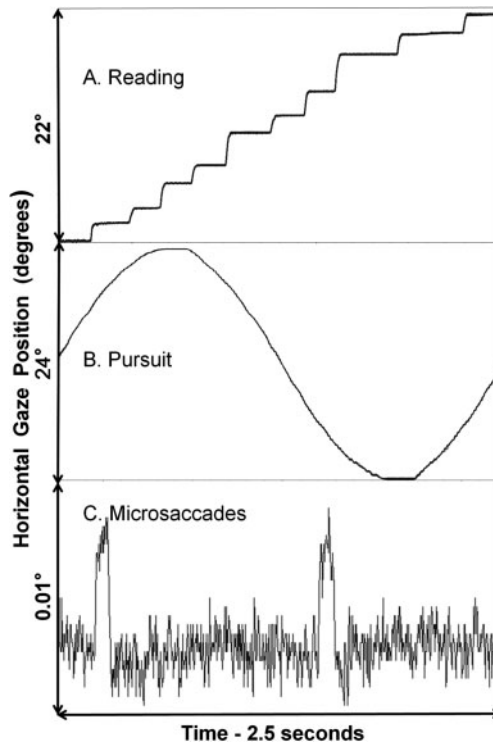
**Figure 6**. Horizontal gaze position traces obtained from recordings of the moving artificial pupil: (A) An overlay of 175 traces with the artificial pupil repeating a reading sequence; (B) an overlay of 245 traces with the artificial pupil repeating a 0.4Hz sinusoidal motion; (C) a single trace with the artificial pupil producing twice a 1.25 seconds motion sequence (based on data from a study of microsaccades by Engbert & Kliegl, 2003).

A) and 245 smooth pursuit traces (Panel B) indicate excellent spatial precision (the overlaid plots almost look like single traces).

In order to derive numeric estimates of spatial precision, for each of the 2500 gaze position samples in both the reading and pursuit tasks, the standard deviation was computed across all the repetitions of each sequence (175 in the reading task and 245 in the pursuit task) and then the average standard deviation was calculated for each task. These spatial precision values are shown in Table 1, which also documents the results from two recording conditions with a stationary pupil. Specifically, in the power off condition the EyeSeeCam moving platform was powered off during recording, while in the power on condition the moving platform was powered, but was programmed to be stationary. Consistent with the traces shown in Panels A and B of Figure 6, the level of spatial noise in all conditions was very low (see Table 1). As expected filtering improved precision

TABLE 1
Spatial precision in degrees of visual angle by recording condition and filtering level
(unfiltered vs. filtered)

| Recording condition | Unfiltered data | Filtered data |
|---|---|---|
| Stationary pupil (power off) | 0.008 | 0.005 |
| Stationary pupil (power on) | 0.020 | 0.011 |
| Reading task | 0.033 | 0.022 |
| Pursuit task | 0.037 | 0.025 |

The values in the table were obtained by computing the average standard deviation of horizontal gaze position in each condition.

in all recording conditions. In addition, there was approximately twice as much noise in the power on than power off conditions, likely due to small amount of motor feedback loop noise. Most importantly, as indicated by the low noise across all conditions, it is clear that the EyeSeeCam moving platform has sufficient spatial precision to mimic most types of biological eye motion. As a case in point, it was interesting to examine if this device could also be used to execute miniature eye movements such as microsaccades. To that end, the moving artificial pupil was programmed to produce twice a 1.25 seconds motion sequence based on data from a study of microsaccades by Engbert and Kliegl (2003). An inspection of the corresponding gaze position trace in Panel C of Figure 6 appears to suggest that in fact microsaccade motion sequences could be executed with the moving artificial pupil. However, more systematic exploration would be needed in order to fully determine the system's capabilities in this scenario.

Notwithstanding the earlier demonstration of the feasibility of using a moving artificial pupil to test the spatial precision of eye trackers, some researchers whose primary focus requires determining the location and duration of fixations often downplay the importance of minimizing spatial noise. The argument is that such noise would be "averaged out" given that fixations are composed of multiple gaze position samples. To illustrate the problem with this argument, consider the idealized saccade and fixation sequence shown in Figure 7. An inspection of this figure reveals that spatial noise added to a few gaze position samples can cause very substantial velocity noise resulting in dramatic distortion in the number, timing, and duration of fixation and saccade events (see Holmqvist et al., 2012, for an excellent illustration and discussion of this issue). Interestingly, the method demonstrated in the present paper can also be utilized for evaluating the precision of various eye movement measures and the quality of the algorithms employed for computing them. To illustrate this, several fixation and saccade measures were derived for each repetition of the reading sequence shown in Panel A of Figure 6. For each measure the standard deviation was then computed across repetitions to produce an estimate of precision. The precision
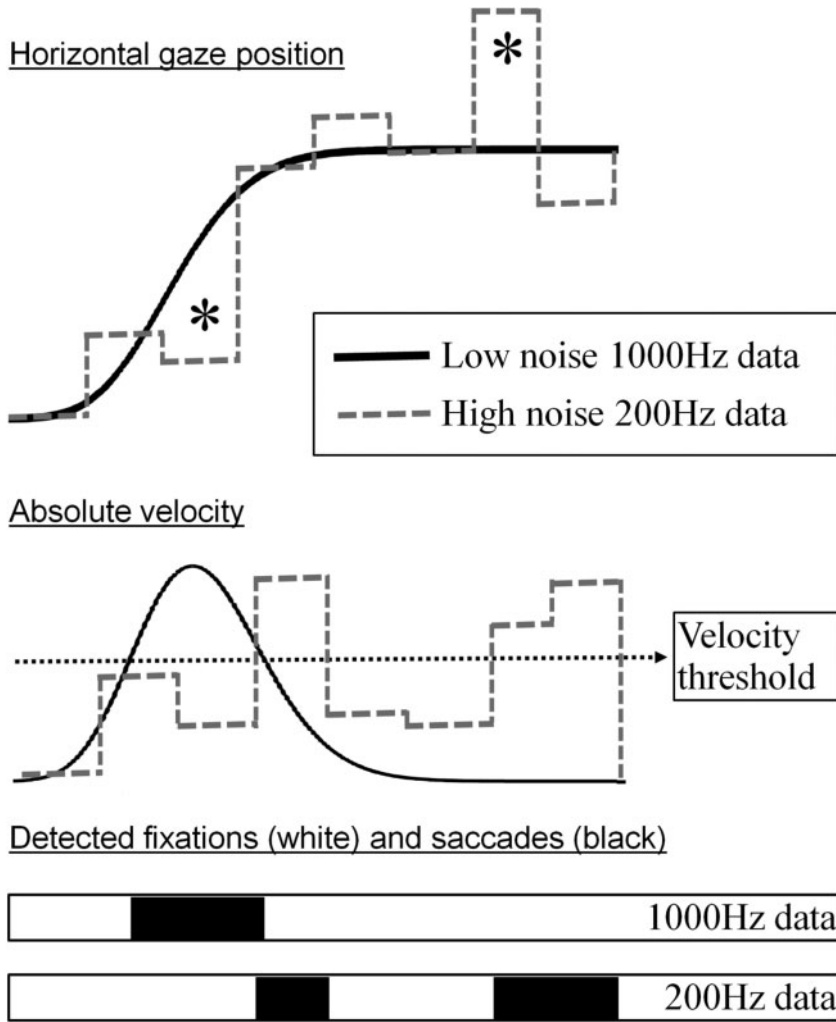
**Figure 7**. The importance of minimizing spatial noise for the valid measurement of fixation and saccades is illustrated by contrasting low noise 1000 Hz data and high noise 200 Hz data depicting the recording of an idealized fixation–saccade–fixation sequence. Horizontal gaze position traces reveal that two of the eight samples produced by the 200 Hz eye tracker displayed moderate levels of spatial noise (marked by asterisks), whereas the other six samples were in fairly good agreement with the low noise 1000 Hz data. The absolute velocity data corresponding to these traces and the fixations and saccades detected using a simple velocity threshold are shown below. Note the dramatic errors in the number, timing, and duration of fixations and saccades that were detected with the 200 Hz data.

values for fixation and saccade measures are shown in Table 2. As was the case with the spatial precision values of individual gaze position samples (see Table 1), exceptionally high precision values were obtained for the various

TABLE 2
The precision of fixation and saccade measures during reading by filtering level (unfiltered vs. filtered)

| Measure | Unfiltered data | Filtered data |
|---|---|---|
| Fixation measures | | |
| Duration (ms) | 0.959 | 0.823 |
| Begin time (ms) | 0.753 | 0.756 |
| End time (ms) | 0.833 | 0.666 |
| Mean horizontal gaze position (degrees) | 0.020 | 0.012 |
| Saccade measures | | |
| Duration (ms) | 0.699 | 0.734 |
| Amplitude (degrees) | 0.048 | 0.035 |
| Mean velocity (degrees/s) | 2.30 | 2.53 |
| Peek velocity (degrees/s) | 2.36 | 2.32 |

The values in the table were obtained by computing the average standard deviation for each measure.

fixation and saccade measures. Filtering improved position related measures (average fixation position, saccade amplitude), but did not substantially impact velocity and/or duration measures.

Finally, it is also interesting to consider the potential utility of the moving artificial pupil for accuracy testing. Based on preliminary evaluation it appears that the functions provided for aiming the platform are currently not accurate enough for that purpose (static accuracy was estimated to be ~0.85 of a degree). Nevertheless, using an extremely laborious manual tweaking procedure it would be possible to direct the projection of the laser diode to visit successive locations on a grid that is superimposed on the screen in a fixed order. Note that any change in the position of any part of the setup (e.g., camera, artificial pupil platform, and screen) would render this effort useless. In the present study, such a procedure was used to aim the laser projection to be perfectly aligned with the centre of the 9-point calibration targets. This permitted the artificial moving pupil to be calibrated. Following calibration, when the same targets were "fixated" again in the same order, gaze position accuracy varied between 0.02 to 0.05 of a degree, a result that is an order of magnitude better than the findings typically obtained with human observers using the same procedure (~0.5 of a degree). However, unlike the clear promise of the present system for testing eye tracker precision, much more work is required before the utility of the moving artificial pupil for accuracy testing could be determined.

## CONCLUDING COMMENTS

The present paper reported on a preliminary exploration of the feasibility of using artificial motion for the purpose of objective testing of eye tracking data

quality. The artificial saccade generator represents a very simple and inexpensive solution for the testing of temporal accuracy (i.e., latency) and temporal precision. Importantly, this solution is useful not only for testing an eye tracker in isolation, but also for testing the entire experimental setup within a particular scenario of use. Specifically, gaze contingent paradigms and saccadic reaction time paradigms would greatly benefit from the availability of such a method of timing validation. The second method that was examined was based on a recent innovative device (EyeSeeCam; Schneider et al., 2009) that permitted the implementation of a moving artificial pupil capable of motion sequences based on biological motion. Overall, the present testing demonstrated that, unlike human observers, this system is capable of reproducing spatiotemporal motion with great precision. Thus, such a device has the potential of fulfilling the visionary proposal by McConkie (1981), and becoming an essential tool for the evaluation of eye tracking data quality. Furthermore, such a device might be useful for evaluating the precision of different eye movement measures (e.g., fixation and saccade measures) and for developing and testing related algorithms (see Inhoff & Radach, 1998, for a review).

It is by no means suggested here that using moving artificial eyes would completely eliminate the need for testing human participants as part of a comprehensive assessment of eye tracking data quality (see also Holmqvist et al., 2012, for a related discussion). Instead, based on the present exploration, it is clear that future studies using progressively more realistic artificial models of biological eye movements could be very instrumental for interpreting the data obtained with human subjects, and as such might facilitate the development of better eye tracking systems. In trying to accomplish this goal, there are many possible interesting extensions to the present approach. For example, Villgrattner, Schneider, Andersch, and Ulbrich (2011) developed a version of the EyeSeeCam platform capable of orienting a small camera around its pan, tilt, and roll axes. Such a platform could possibly be adapted to study data quality of torsional eye tracking systems. In addition, artificial moving eyes could be made more realistic by adding mechanisms for producing artificial blinks, squints, or even for dynamically controlling pupil size. Finally, it is hoped that the broader availability of moving artificial eyes would promote greater focus on the study of eye tracking data quality by empowering eye movement researchers to investigate the reliability and validity of the hardware and software tools that they employ in their studies.

## REFERENCES

Bernard, J. B., Scherlen, A. C., & Castet, E. (2007). Page mode reading with simulated scotomas: A modest effect of interline spacing on reading speed. *Vision Research, 47*, 3447–3459.

Collewijn, H., van der Mark, F., & Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research, 15*, 447–450.

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. New York, NY: Springer-Verlag.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, *43*, 1035–1045. 10.1016/S0042-6989(03)00084-1

Frens, M. A., & van der Geest, J. N. (2002). Scleral search coils influence saccade dynamics. *Journal of Neurophysiology*, *88*, 692–698.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka,H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. *Proceedings of the 2012 Symposium on Eye Tracking Research and Applications* (pp. 45–52). New York, NY: ACM.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–54). Oxford: Elsevier.

Kimmel, D. L., Mammo, D., & Newsome, W. T. (2012). Tracking the eye non-invasively: Simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers in Behavioral Neuroscience*, *6*(49), 1–17.

Kohlbecher, S., Bartl, K., Bardins, S., & Schneider, E. (2010). Low-latency combined eye and head tracking system for teleoperating a robotic head in real-time. *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications* (pp. 117–120). New York, NY: ACM.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *The Oxford handbook of eye movements*. Oxford: Oxford University Press.

McConkie, G. W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods and Instrumentation*, *13*, 97–106.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception and Psychophysics*, *17*, 578–586.

McConkie, G. W., Wolverton, G. S., & Zola, D. (1984). Instrumentation considerations in research involving eye-movement contingent stimulus control. In A. G. Gale & F. Johnson (Eds.), *Theoretical and applied aspects of eye movement research* (pp. 39–47). Amsterdam: North-Holland.

McConkie, G. W., Zola, D., Wolverton, G. S., & Burns, D. D. (1978). Eye movement contingent display control in studying reading. *Behavior Research Methods and Instrumentation*, *10*, 154–166.

Plant, R. R., Hammond, N., & Turner, G. (2004). Self-validating presentation and response timing in cognitive paradigms: How and why? *Behavior Research Methods, Instruments, & Computers*, *36*, 291–303. 10.3758/BF03195575

Plant, R. R., & Quinlan, P. T. (2013). Could millisecond timing errors in commonly used equipment be a cause of replication failure in some neuroscience studies? *Cognitive, Affective & Behavioral Neuroscience*, *13*, 598–614. 10.3758/s13415-013-0166-6

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009). The thirty-fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multi-resolutional displays: An integrative review. *Human Factors*, *45*, 307–328.

Richlan, F., Gagl, B.Schuster, S., Hawelka, S., Humenberger, J., & Hutzler, F. (2013). A new high-speed visual stimulation method for gaze-contingent eye movement and brain activity studies. *Frontiers in Systems Neuroscience*, *7*, 24.

Schneider, E., Villgrattner, T., Vockeroth, J., Bartl, K., Kohlbecher, S., Bardins, S., Ulbrich, H., & Brandt, T. (2009). EyeSeeCam: An eye movement-driven head camera for the examination of

natural visual exploration. *Annals of the New York Academy of Sciences*, *1164*, 461–467. 10.1111/j.1749-6632.2009.03858.x

Traisk, F., Bolzani, R., & Ygge, J. (2005). A comparison between the magnetic scleral search coil and infrared reflection methods for saccadic eye movement analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *243*, 791–797.

van der Geest, J. N., & Frens, M. A. (2002). Recording eye movements with video-oculography and sclera search coils: A direct comparison of two methods. *Journal of Neuroscience Methods*, *114*, 185–195.

Villgrattner, T., Schneider, E., Andersch, P., & Ulbrich, H. (2011). Compact high dynamic 3 DoF camera orientation system: Development and control. *Journal of System Design and Dynamics*, *5*, 819–828.