



OPEN

Full gaze contingency provides better reading performance than head steering alone in a simulation of prosthetic vision

Nadia Paraskevoudi^{1,2} & John S. Pezaris^{3,4}✉

The visual pathway is retinotopically organized and sensitive to gaze position, leading us to hypothesize that subjects using visual prostheses incorporating eye position would perform better on perceptual tasks than with devices that are merely head-steered. We had sighted subjects read sentences from the MNREAD corpus through a simulation of artificial vision under conditions of full gaze compensation, and head-steered viewing. With 2000 simulated phosphenes, subjects ($n = 23$) were immediately able to read under full gaze compensation and were assessed at an equivalent visual acuity of 1.0 logMAR, but were nearly unable to perform the task under head-steered viewing. At the largest font size tested, 1.4 logMAR, subjects read at 59 WPM (50% of normal speed) with 100% accuracy under the full-gaze condition, but at 0.7 WPM (under 1% of normal) with below 15% accuracy under head-steering. We conclude that gaze-compensated prostheses are likely to produce considerably better patient outcomes than those not incorporating eye movements.

Visual scanning is an inherently automatic aspect of perceiving our external world, and supporting these motions in artificial vision devices has been one of many challenges for the field of visual prostheses. Visual prosthesis devices to restore function to the blind have been under investigation and development for nearly a century (for reviews see^{1–6}), with the most recent major milestones being approval of devices for clinical use in the US and Europe (Alpha IMS and AMS, Retina Implant AG, Reutlingen, Germany; Argus II, Second Sight Medical Products Inc, Sylmar, CA; PRIMA and IRIS II, Pixium Vision, Paris, France^{1,7–13}). To-date, most visual prostheses provide a modest level of visual function through electrical stimulation applied to a collection of electrode contacts at one of the stages of the early visual pathway (e.g., retina, optic nerve, lateral geniculate nucleus or primary visual cortex). Each of these contacts is used to stimulate a small volume of neural tissue and evoke a focal visual percept called a *phosphene*, colloquially known as the pixels of artificial vision. Coordinated stimulation across electrodes creates patterns of phosphenes through which images are conveyed to the subject, providing restoration of function.

Visual prostheses can be classified into two broad categories, those that use the optics of the eye as a critical part of their design, typically placing photodiode arrays on the retina that turn light focused by the eye into patterns of neural activity (Alpha IMS and AMS, PRIMA), and those that deliver stimulation without employing the optics of the eye (Argus II, Orion, EPIRET3, IRIS II, Boston Retinal Implant, Australian bionic eye, CORTIVIS, etc.^{1,14–20}; see reviews by Lews and colleagues²¹ and Yue and colleagues²²). For subjects to explore the visual scene, these two classes of devices present very different interfaces. The first group employs the natural optics and is *gaze steered*, that is the direction of interest is controlled by the subject through both eye and head movements, while the second group does not employ the natural optics and is *head steered*, that is, the direction of interest is controlled only by head position. Here, we will be concerned only with the latter group.

Head-steered visual prostheses use a camera to image the external world, translating signals from that camera into patterns of neural stimulation that are delivered to the early visual pathway without engaging the optics of the eyes^{5,6}. Most often the cameras in these devices are mounted on a set of goggles, roughly aligned with the optical axis of the eyes. To emphasize the distinction from gaze-steered prostheses, head-steered prostheses have

¹Brainlab - Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain. ²Institute of Neurosciences, University of Barcelona, Barcelona, Spain. ³Department of Neurosurgery, Massachusetts General Hospital, Boston, MA, USA. ⁴Department of Neurosurgery, Harvard Medical School, Boston, MA, USA. ✉email: pezaris.john@mgh.harvard.edu

an important limitation that they are controlled only through motions of the head and do not incorporate eye position information. It is this specific issue we address with the present report.

The influence of oculomotor processes such as saccadic movements on normal visual perception is substantial (see review by Paraskevoudi and Pezaris²³). Eye movements used to scan our environment cause the visual presentation of external objects to move across our retinas^{24–26}. Despite these frequent displacements, the brain constructs a stable representation of our surroundings by compensating each moment's retinal snapshot with information about the corresponding gaze position^{5, 26, 27} through a process called *spatial updating*. In diseases such as oscillopsia²⁸ where the brain no longer has an accurate representation of the position of the eye²⁹, the visual and cognitive effects can be debilitating³⁰, because spatial updating can no longer work properly. In normal visual perception, even small eye movements are important as they modulate visual processing^{31–33}, optimize eye position during fixation^{32, 34}, and prevent retinal fading³⁵. Thus, eye movements and spatial updating contribute to the correct localization and perception of objects, making them critical for visuomotor coordination, navigation, and interaction with the external world.

Visual prostheses create images by stimulating patterns of phosphenes. Eye movements that could be used to drive spatial updating in visual prostheses cause phosphenes to shift with each new position, just as eye movements in normal vision cause images on the retina to shift. This observation has a cause with easily-overlooked consequences: phosphenes are retinotopically locked, that is, they appear in a fixed position relative to the current gaze location. Brindley and Lewin recognized this effect in their seminal 1968 report on cortical stimulation³⁶, as was confirmed by later studies^{5, 37–45}. In normal vision, each time the eye moves to a new position, a different part of the visual scene falls on the retina and the brain uses its knowledge of gaze position to compensate the new input to correctly assemble a stable perception of the external world. An important limitation of head-steered prosthetic devices is that they do not update visual information as patients move their eyes to scan the environment, but the brain still is aware of eye position, creating a conflict (Fig. 1). By using cameras that are fixed in position relative to the head, typically on a set of goggles, such systems fail to maintain alignment of the camera and pupillary axes as the eyes move, and require implanted patients to direct their field of view solely with head scanning, while suppressing eye movements^{14, 15}, an unnatural behavior that requires substantial training^{46, 47}. Misalignments between the camera position and the patients' gaze position impair perception and decrease performance in simple object localization tasks⁴⁸; even minor misalignments between the eye and camera optical axes are known to affect performance in prosthesis user⁴⁹.

Presenting images in eye-centric coordinates through a *gaze contingent* mechanism would potentially increase the usability of these devices⁵⁰. Ideally, the imaging system (e.g., a goggles-mounted camera) should reproduce the view based on instantaneous eye position as if the camera had moved like the eye. Perhaps the most direct method to achieve this compensation would be to electronically shift a region of interest from a wide field of view so as to track eye-driven changes in gaze position⁵.

Recent attempts to introduce gaze contingency in simulation studies of artificial vision have opted for the incorporation of an eye tracker into the simulated prosthetic device^{5, 40, 51–55}. Simulation studies with sighted individuals have found improved performance in gaze contingent conditions as compared to uncompensated ones^{52, 54}. McIntosh⁵² employed a simulated artificial vision paradigm to simulate a retinal prosthesis with residual natural vision where sighted participants performed three visually-guided tasks, a search task to locate an object among distractors, a reach-and-grasp task, and a navigational task. For the search task, the time to locate the target object was reduced by up to a factor of 2 for gaze-contingent as compared to head-steered conditions. For the reach-and-grasp task, the improvement in completion time was up to 50%, with up to five times fewer touches on erroneous targets. For the navigational task, walking speed was improved by 10% and obstacle bumps reduced by up to 40%. More recently, Titchener and colleagues⁵⁴ examined the effect of gaze eccentricity on hand-eye coordination by employing a simulated artificial visual paradigm, where the phosphene layout was modeled after Bionic Vision Australia's second-generation retinal implant. Sighted participants performed a target localization task, where subjects reported the observed location of large (5° diameter) targets with their finger on the touch-sensitive screen. The gaze-compensation condition led to a significant reduction of pointing error when compared with the head-steered condition in six out of seven subjects. Importantly, in the head-steered condition, larger gaze eccentricity led to larger pointing error that was biased in the direction of gaze location, an effect that did not appear in the gaze-compensated condition.

In line with simulation studies with sighted volunteers, the effects of gaze-contingent versus head-steered stimulation have been described in studies with Argus II users^{15, 48}. In a study by Sabbah and colleagues⁴⁸, patients performed a target localization task where they were instructed to shift their gaze toward different locations on a screen by moving their eyes alone and maintaining their head still before pointing at the target. Rather than pointing at the actual target locations, subjects indicated positions on the touch screen that were biased in the direction the eyes were pointing. As there was no concurrent change in head position, and therefore no change in camera output, their result indicates that camera direction alone was insufficient for correct spatial localization when eye movements caused a misalignment between gaze and camera directions. In a study by Caspi and colleagues¹⁵, Argus II users were fitted with an experimental eye tracker which was enabled or disabled on a similar target localization touch task. When gaze contingency was enabled, head motions were significantly reduced and often nearly eliminated with mean head speed reduced by over half, and the extent of eye movement increased by a mean of 70%. While such changes might be expected with the introduction of an additional degree of freedom, importantly, no training was provided for the new mode, and pointing accuracy on the task improved by an average of 30% (see Discussion).

Research conducted so far to directly answer questions of the importance of gaze compensation suggests that it is indeed a critical aspect of visual prosthetics. Thus far, studies have either employed simple target detection tasks that do not capture the complexity of activities performed in everyday life^{15, 48, 54}, or used relatively

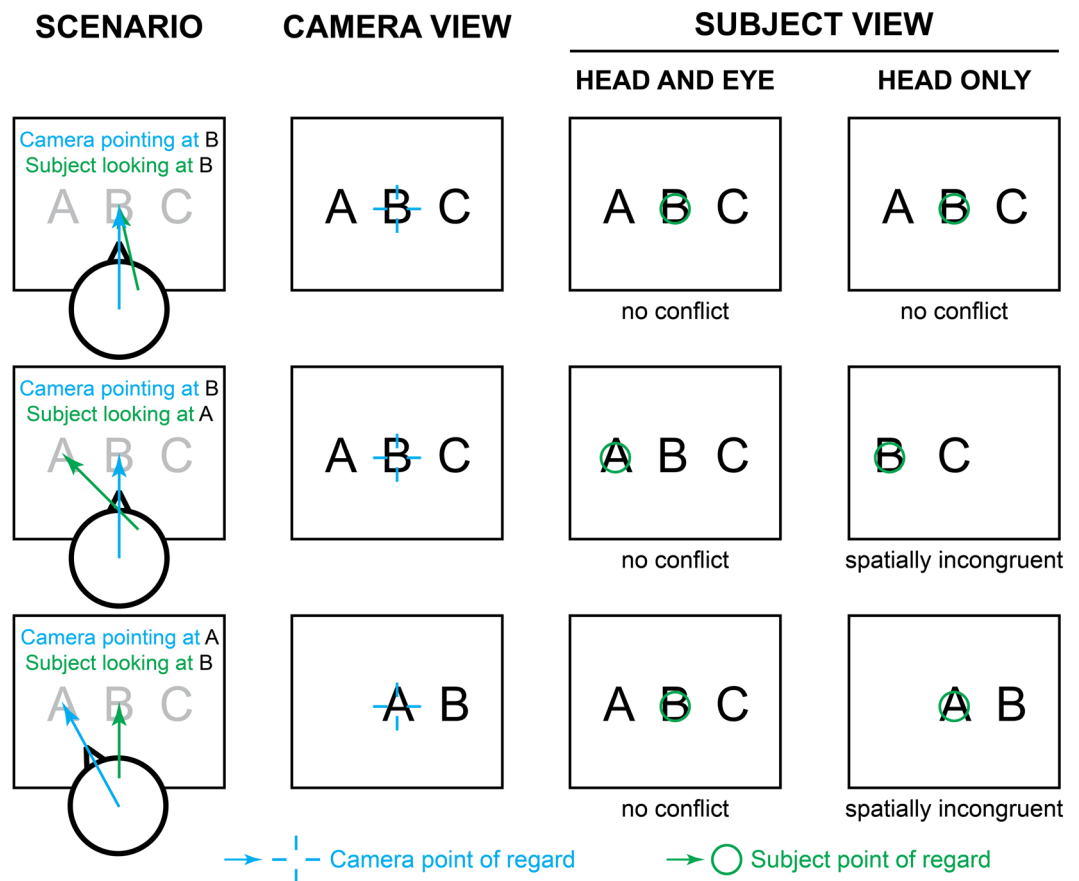


Figure 1. Gaze conditions and the problem created by ignoring eye position. In the general case where the eyes are not used to image the external world in a visual prosthesis, there is potential for the camera axis, typically steered by head position, and the gaze location, determined by the combination of head position and eye position within the head, to be out of agreement. This figure depicts three conditions, by row, the first where the gaze (blue) and camera (green) are aligned, pointing forward (top row), the second where the head and camera are in the same positions, but the eyes have deviated to the left (middle row), and the third where the head and camera have been turned to the left, but the eyes have counter-rotated to the right such that gaze is once again straight ahead (bottom row). The second column from the left shows the image that the camera sees. The third column depicts what is perceived when the prosthesis includes compensation for gaze location (Head and Eye, or Full Gaze), with congruent stimulation resulting in a perception that is spatially stable. The fourth column depicts what is perceived by the subject where the prosthesis delivers the image from the uncompensated camera (as in head-steered devices), but the brain continues to apply its knowledge about eye position, creating spatially incongruent conditions whenever the camera and eyes are not pointing in the same direction.

large and simple optotypes⁵⁶, leaving open the question of the impact of gaze compensation in more cognitively engaging, naturalistic tasks.

Here, we address this gap by exploring the effect of gaze contingency with a reading task. Specifically, we created a virtual reality simulation of prosthetic vision that could be switched between modes where the simulated head-mounted camera produced a video stream that was either (a) corrected for gaze position, a mode that we call *Full Gaze*, or (b) steered only by head position, a mode that we call *Head Only*. As a metric of system utility, we used a reading task derived from the MNREAD test of visual acuity⁵⁷, and compared reading accuracy and speed under the two experimental conditions. In our simulations, the Full Gaze condition, with inclusion of gaze position compensation, resulted in a marked increase in reading performance over the Head Only condition. This finding has strong implications for future visual prosthesis designs in order to maximize performance in tasks of daily living.

Results

Subjects ($n = 23$, 9 male and 14 female, mean age of 26 years old, range 18–44 years old) were recruited from students at the University of Athens and the general population around the university and required to have proficiency in English (see Methods). They were verified to have normal vision with an informal Snellen chart evaluation, with binocular results of 20/13, or -0.2 ± 0.1 logMAR (mean, s.d.; range of -0.3 to 0.0 logMAR). Importantly, the assessed visual acuities were far better than what was required for the experimental task.

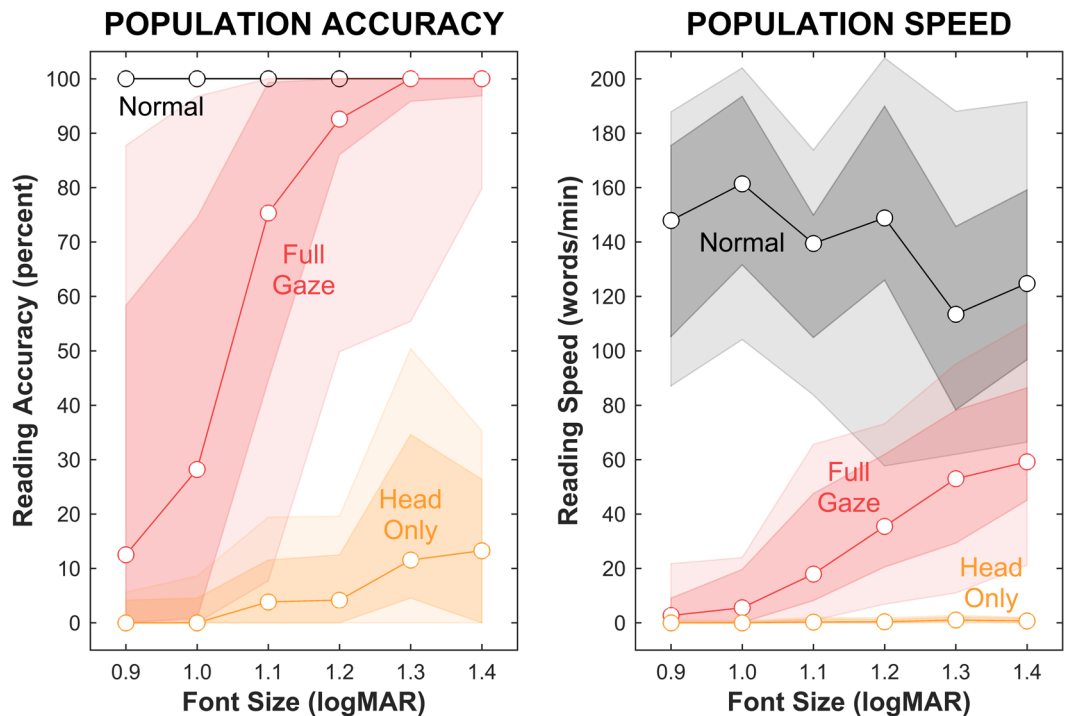


Figure 2. Population reading accuracy and speed. Reading accuracy (left) and speed (right) are shown for the three viewing conditions with population median as lines, 16/84 percentile range as dark color, 5/95 percentile range as light color. Performance is shown across the range of font sizes and measured in percent of correctly read words for accuracy, and number of correctly read words per minute for speed. The three conditions are Normal, where text is shown unadulterated on the screen (blacks), Full Gaze, where text is shown in a simulation of phosphene vision with full gaze compensation of the scene camera (reds), and Head Only, where text is shown in phosphene vision where the scene camera is steered only by head motions (oranges). In both phosphene view cases, the phosphene locations are locked to retinotopic location based on the instantaneous measured gaze position; we expect phosphenes for all visual prostheses to be stable in the retinotopic coordinate system (see Introduction). In the Full Gaze condition, the simulated visual prosthesis is assumed to incorporate a gaze tracker that provides rapid compensation for eye movements, performing a translation of the scene camera image as if it were being steered by the eyes. In the Head Only condition, gaze compensation of the scene camera is disabled, rendering it steerable only through head motion; this reflects the operation of many contemporary visual prosthesis designs.

Subjects were presented a series of simple, three-line sentences to read on a computer monitor that simulated artificial vision from a device with 2000 phosphenes under the Full Gaze and Head Only conditions. The sentences were taken from the MNREAD corpus^{57, 58} and used to develop performance curves versus font size through the measurement of reading accuracy (percentage of words read correctly) and reading speed (number of correctly read words per minute). *Normal* mode trials without the artificial vision simulation were included as a control condition. The experiment was broken into a series of six-trial mini-blocks where each mini-block presented text for a given viewing condition at the six font sizes (see Discussion). Mini-blocks were ordered so that Normal (A), Full Gaze (B), and Head Only (C) conditions were each presented twice (A₁, B₁, C₁, B₂, C₂, A₂). Full details are provided in the Methods section.

Reading accuracy and speed. Our primary finding is that both reading accuracy and speed were substantially higher for the Full Gaze condition than for Head Only for the font sizes measured (Fig. 2; Table 1). Performances under the three conditions are distinct by inspection for both metrics (Wilcoxon rank sum tests between Full Gaze and Head Only conditions pooled over all font sizes produce $p = 10^{-31}$ and 10^{-35} , $z = 11$ and 12 , $r = 1.0$ and 1.1 , on accuracy and speed respectively; individual font-size-by-font-size comparisons produce similar results with a maximum of $p < 0.001$, some of which are detailed below).

For reading accuracy, under the Full Gaze condition, 19/23 (83%) subjects displayed 100% accuracy at the largest font size (1.4 logMAR); the population accuracy was 100^{+0}_{-3} % (median and differences to 16th and 84th percentiles) with performance for the population tailing off over decreasing font size to 13^{+46}_{-13} % at 0.9 logMAR, following a sigmoidal curve. In comparison, under the Head Only condition, subjects were only able to attain 13^{+13}_{-13} % accuracy at the largest font size, tailing off to 0^{+4}_{-0} % at the smallest size. For the largest font size, effect size between Full Gaze and Head Only conditions for Wilcoxon rank sum was $r = 1.3$ (with $z = 6.0$, $p = 10^{-9}$), trailing off to 0.9 ($z = 4.4$, $p = 10^{-5}$) at the smallest. The Normal condition of reading ordinary text confirmed there were no fundamental issues with 100^{+0}_{-0} % reading accuracy at all font sizes. Sigmoidal fits of the population data for

Measurement	Viewing condition	logMAR					
		0.9	1.0	1.1	1.2	1.3	1.4
Reading accuracy (percent)	Normal	100.0 ⁺⁰ ₋₀	100.0 ⁺⁰ ₋₀	100.0 ⁺⁰ ₋₀	100.0 ⁺⁰ ₋₀	100.0 ⁺⁰ ₋₀	100.0 ⁺⁰ ₋₀
	Full gaze	13 ⁺⁴⁶ ₋₁₃	28 ⁺⁴⁶ ₋₂₇	75 ⁺²⁵ ₋₃₁	96 ⁺⁷ ₋₇	100 ⁺⁰ ₋₄	100 ⁺⁰ ₋₃
	Head only	0.0 ^{+4.2} _{-0.0}	0.0 ^{+4.5} _{-0.0}	3.8 ^{+7.7} _{-3.8}	4.2 ^{+8.3} _{-4.2}	12 ⁺²³ ₋₇	13 ⁺¹³ ₋₁₃
Reading speed (WPM)	Normal	148 ⁺²⁸ ₋₄₃	161 ⁺³² ₋₃₀	140 ⁺¹⁰ ₋₂₅	149 ⁺⁴¹ ₋₂₃	113 ⁺³² ₋₃₅	124 ⁺³⁴ ₋₂₈
	Full gaze	2.7 ^{+6.4} _{-2.7}	5.5 ^{+14.1} _{-5.4}	18 ⁺³⁰ ₋₁₀	35 ⁺²⁷ ₋₁₅	53 ⁺²⁵ ₋₂₄	60 ⁺²⁷ ₋₁₄
	Head only	0.0 ^{+0.3} _{-0.0}	0.0 ^{+0.4} _{-0.0}	0.3 ^{+1.1} _{-0.3}	0.4 ^{+0.8} _{-0.4}	1.0 ^{+1.1} _{-0.6}	0.7 ^{+1.0} _{-0.7}

Table 1. Reading accuracy and speed results. Values for reading accuracy in percent correct and reading speed in words per minute (WPM) are given for the three viewing conditions over the population of 23 subjects. Data are presented by font size and given as median values with differences to 16th and 84th percentiles. See Fig. 2 for a graphical presentation of these data.

reading accuracy to estimate visual acuity at the 50% level showed the Full-Gaze condition provided an equivalent acuity of 1.02 ± 0.12 logMAR (95% confidence interval) and the Head-Only condition an equivalent acuity of 1.66 ± 0.37 logMAR (95% confidence interval; rank sum $p = 10^{-8}$).

For reading speed, there was a larger contrast between the two phosphene view conditions than for reading accuracy. For the Full Gaze condition, speed at the largest font size was 60^{+27}_{-14} WPM, decreasing in a sigmoidal fashion to $2.7^{+6.4}_{-2.7}$ WPM at the smallest font size (Fig. 2; Table 1). The Head Only condition revealed a profound problem, with reading speed at the largest font size of $0.7^{+1.0}_{-0.7}$ WPM (one word about every $1 \frac{1}{2}$ min, or two orders of magnitude slower than the Full Gaze condition) and with subjects barely able to read any words at the smallest font size ($0.0^{+0.3}_{-0.0}$ WPM). For the largest font size, effect size was $r = 1.2$ ($z = 5.8$, $p = 10^{-9}$), trailing off to 0.9 at the smallest ($z = 4.5$, $p = 10^{-6}$). Reading in the Normal condition that served as a control increased from 124^{+34}_{-28} WPM at the largest font size to 148^{+28}_{-43} WPM at the smallest (the smaller font sizes require shorter gaze shifts, so reading is somewhat faster), consistent with a previous report from a highly similar population⁵³, and verifying that normal reading was possible.

We examined performance between the first and second presentation of each class of mini-blocks to look for learning or other longitudinal effects (Fig. 3). For reading accuracy, there was no significance found for the Normal condition (Wilcoxon rank sum test, $p = 1.0$), with the median values being 100^{+0}_{-0} % for both A_1 and A_2 mini-blocks across all font sizes. For the Full Gaze condition, there was, again no significance found ($p = 0.2$) with the median values being 92^{+8}_{-84} % for B_1 , and 100^{+0}_{-88} % for B_2 , suggestive of an improvement and consistent with what was found in an earlier report^[53]. For the Head Only condition, a significant difference was found ($z = 2.5$, $r = 0.2$, $p = 0.01$) with median values of $0.0^{+17}_{-0.0}$ % for C_1 and $7.7^{+13}_{-7.7}$ % for C_2 .

For reading speed, there was a significant but small ($z = 2.8$, $r = 0.2$, $p = 0.005$) slow-down in the Normal condition, from 148^{+43}_{-48} to 130^{+41}_{-35} WPM for A_1 to A_2 . For the Full Gaze condition, there was no significance ($p = 0.9$) between B_1 and B_2 at 29^{+38}_{-28} and 26^{+24}_{-38} WPM, respectively. For the Head Only condition, there was a significant difference ($z = 3.0$, $r = 0.3$, $p = 0.003$) between C_1 and C_2 at $0.0^{+1.0}_{-0.0}$ and $0.3^{+1.4}_{-0.3}$ WPM, respectively. The C_1 to C_2 increase corresponds to one additional correctly read word (from zero to one, or from one to two) in the three largest font sizes. This small difference may reflect a leftward movement of the sigmoid curve of performance versus font size expected from a learning process⁵³.

We compared average performance under the two experimental conditions on a per-subject basis to see if there were correlations (did being better at one viewing condition mean subjects were better at another?). No strong correlations were found. Comparing mean reading accuracy averaged over all font sizes for each subject, Spearman's rank correlation was $r_s = 0.36$ with $p = 0.09$ between Full Gaze and Head Only conditions: good overall performance on one condition had limited predictive power for the other. For mean reading speed, similar values were found at $r_s = 0.37$ with $p = 0.09$, again showing a lack of strong relationship. Lower levels of correlation that were also not significant were found for mean reading speed in the Normal condition against Full Gaze ($r_s = 0.08$, $p = 0.7$) and Head Only conditions ($r_s = -0.11$, $p = 0.6$).

Subject behavior for gaze and head position in the control and two experimental conditions showed a high degree of similarity between Normal and Full Gaze conditions, with gaze location tracing out the three lines of text and head position relatively still, but no similarity with the Head Only condition that exhibited little-to-no structure in gaze location or head location despite higher levels of activity (Fig. 4). The total time spent reading with text at 1.4 logMAR was $6.1^{+2.0}_{-1.3}$ seconds in the Normal condition, $13.2^{+3.9}_{-4.3}$ seconds in the Full Gaze condition, reflective of a reading speed about half as fast (at the same 100% accuracy), and 138^{+104}_{-106} seconds in the Head Only condition, reflective of the substantially increased difficulty.

Discussion

In this study, we assessed the importance of gaze contingency by employing a naturalistic reading task through a simulation of prosthetic vision that could be switched between two modes where the visual stimuli were either (a) corrected for eye position (*Full Gaze*), or (b) steered only by head position (*Head Only*). Performance was measured using the metrics of reading accuracy and reading speed that are well-studied in the low-vision literature, built upon the standard MNREAD test⁵⁷, allowing the ready comparison and replication of our observations.

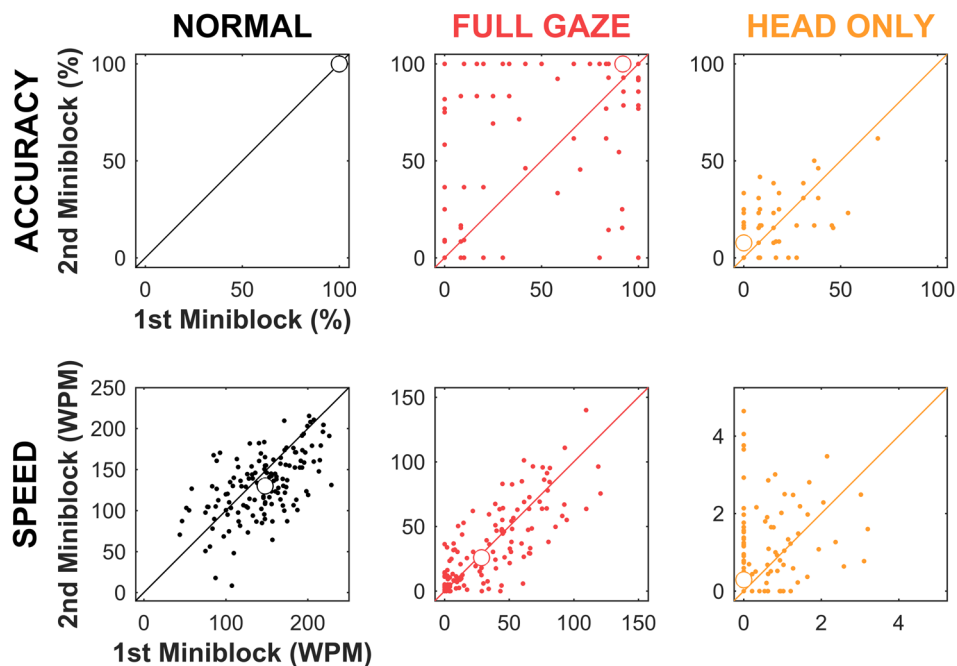


Figure 3. Performance in first versus second mini-block. Performance on reading accuracy (top row) and reading speed (bottom row) for the three viewing conditions, Normal (black), Full Gaze compensation (red), and Head Only steering (orange) is shown for the first versus second mini-block in each condition. All conditions had two mini-block presentations, each of which included one trial for all font sizes. Dots show the performance of matched stimulus conditions on the two mini-blocks. Unfilled circles show the population mean of the scattergrams. For the Normal condition, reading accuracy did not produce meaningful data as accuracies were all at 100%.

Our primary finding is that under equivalent conditions, a simulated prosthesis that includes gaze compensation to determine how to activate phosphenes has substantial performance advantage over one that is limited to head steering of a camera. Although the differences in performance were most striking for the easiest conditions with large font sizes that activated many phosphenes, differences were still observable in the more difficult conditions with small font sizes that activated fewer phosphenes (e.g., the 1.0 logMAR text typically activated only 175 phosphenes for the straight-ahead gaze position). We suggest that our findings can be applicable to contemporary clinical head-steered devices with limited resolution and highlight the need to introduce gaze-contingent information to visual prosthetics by the incorporation of an eye-tracker to current head-steered devices.

Reading performance is enhanced under full gaze compensation. Our findings show that reading performance in the Full Gaze condition can be relatively fluid and comparable to reading in the control condition, but becomes profoundly affected when participants were required to adopt a head-scanning strategy in the Head Only condition. This effect was evident in both accuracy and speed, as discussed in the following two subsections.

Reading accuracy. We obtained higher reading accuracy in the Full Gaze compared to the Head Only condition, replicating and extending previous work that has reported better performance under gaze-contingent versus non-gaze-contingent conditions^{15, 52, 54}. A recent study by Caspi and colleagues¹⁵ had Argus II patients perform a pointing task to compare eye-and-head or head-only conditions analogous to our Full Gaze and Head Only conditions, respectively. They reported a pointing error of 3.7 ± 1.0 degrees with eye-and-head steering versus 5.1 ± 1.4 degrees with head-only (we compute $p=0.016$ using a Wilcoxon signed rank test from their published data), for an immediate reduction in error of over 30% when engaging full gaze tracking with subjects who were already highly trained on head-only steering. Although that study and the present one are not directly comparable due to differences in the experimental design (localization versus reading task) and the subjects used in each study (blind implanted patients versus healthy, normal-sighted participants), our results are consistent with theirs, albeit our observations are that the effects are larger.

Reading speed. The differences between the Full Gaze and Head Only conditions were even more prominent for reading speed performance, which was deeply affected under the Head Only condition. This finding is in line with previous simulation work with sighted individuals showing faster completion times at a visual search task under a gaze compensated condition compared to a head-steered one⁵², as well as with reports that Argus

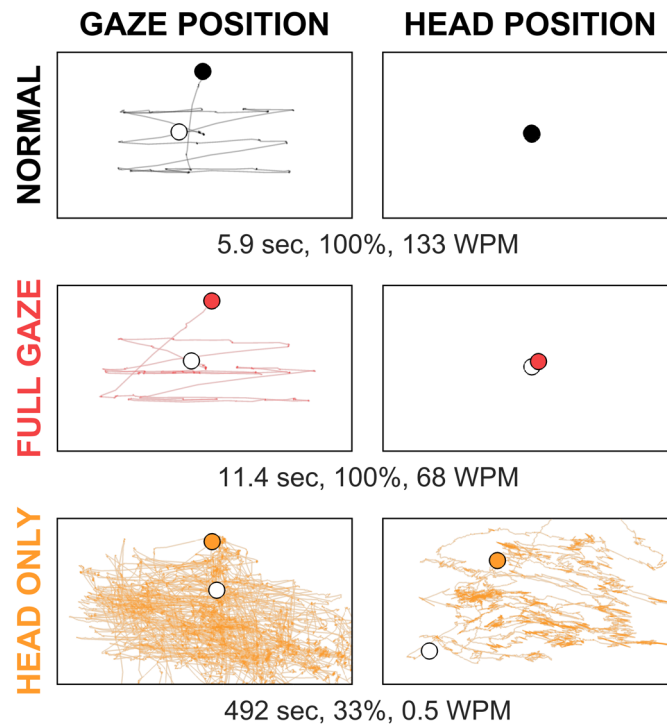


Figure 4. Example gaze and head traces from one subject. Gaze (left column) and head (right column) position traces are shown for one of the most overall-capable subjects reading text at 1.4 logMAR (the largest size) under the three viewing conditions, Normal (top row), Full Gaze (middle row), and Head Only (bottom row). For each trace, the start is shown by an unfilled circle, and the end by a filled circle. Trial length, reading accuracy, and reading speed are shown underneath each pair of plots: the first two trials (top two rows) took some few seconds to complete while the third trial (bottom row) took over 8 min; for Head Only, the values are substantially better than the population means, but a similar dichotomy in trial completion time was typical across the population. Eye movements during the Normal and Full Gaze conditions here follow a typical scan path for reading the three lines in the MNREAD sentences; head position is nearly motionless, also typical, without having instructed the subject to hold their head still. Eye movements in the Head Only condition reflect stereotypical looping center-out-and-back movements, while head position reflects attempts to scan the scene. This looping behavior is triggered by eye movements to foveate a portion of text the subject wishes to view, followed by a realization that the text tracks their gaze location (Fig. 9), and a subsequent saccade back to the center of the screen.

II users adopt time-consuming compensatory head scanning strategies when presented with camera-gaze misalignments⁴⁸. Our observation of the highest reading speed during Head Only condition being below 1 WPM (at 1.4 logMAR) is also consistent with reports from the literature for reading speed with actual visual prostheses that lack gaze contingency⁵⁶. This consistency provides not only confirmation of our results, but by extension, an estimate for the level of performance increase that might be expected with the addition of gaze contingency to a clinical device. High-functioning Argus II users are reported to have single word reading accuracies of 71% with font sizes between 2.4 and 2.6 logMAR (see Table II in the 2013 report by da Cruz and colleagues⁵⁶). Combining the results from that report with the range of completion times given elsewhere⁵⁹, we estimate that these patients had a reading speed of between 0.2 and 0.9 WPM. While the experimental conditions were different from those presented here (larger font sizes used by da Cruz and colleagues⁵⁶, single words versus full sentences, different phosphene counts and distribution), and their figures cannot be reliably translated to MNREAD measurements, the similar, low range of reading speed suggests that adding gaze contingency to the Argus II device may substantially improve its utility. While various factors might prevent the gain of two orders of magnitude in reading speed we reported above (for example, MNREAD sentences at 2.4 logMAR do not fit in a normal full visual field), it would not seem unreasonable to expect reading speeds in the range of multiple words per minute.

Performance under the uncompensated Head Only condition can be improved. Although our task was not designed to test for possible learning effects, we observed performance enhancements between C_1 and C_2 , the two Head Only mini-blocks. Learning effects have been already observed in gaze-contingent simulations of artificial vision^{53, 60}, suggesting that training could create similar improvements in patients even with uncompensated devices. Given the case report of an individual who lacks the ability to move their eyes having spontaneously developed head-steering⁶¹, a natural follow-on question would be to see if training would benefit one mode more than the other, since we did not find a significant improvement for B_1 to B_2 , the two Full Gaze mini-blocks. Discouragingly, training-driven improvements for camera-gaze misalignments have been found to

be small and to progress at a slow rate⁶². We, thus, speculate that trained reading performance under the Head Only condition might not reach the levels we observed for the Full Gaze condition (see [Post-implant rehabilitation and training](#), below).

No training required for full gaze, invasive coaching required for head only condition. In addition to the striking differences in performance, we found that subjects did not require introduction or coaching to perform the first mini-block with phosphene view, B₁, which employed the Full Gaze condition and resulted in naturalistic gaze shifts (Fig. 3). Subjects adapted to it quickly within the first trial and performed well, as expected from earlier work^{53, 55, 60}. On the other hand, the Head Only mini-block C₁ that followed required frequent coaching to, “keep your eyes looking forward and use your head to scan.” Without such coaching, subjects were, in general, unable to perform the task. The difference was not only striking, it was sobering for its implications for actual implants, as it touched the critical shortcomings that underlie approaches that do not support gaze steering, either intrinsically or through a compensatory mechanism. We are driven to the conclusion that the reports of seemingly poor performance in the literature that are in conflict with our previous work both in simulations with sighted humans⁵¹ and non-human primates⁶³, and with implants in non-human primates in ongoing work are the results of the disadvantage faced by systems that lack gaze compensation. We further speculate that even with subject training to suppress gaze movements and employ head scanning, it will be difficult for prostheses lacking gaze compensation to match the utility of those that have incorporated it. Indeed, even if we were to further suppose that parity might be obtainable with the two approaches should non-gaze-contingent recipients undergo post-implant rehabilitation training, we have seen that an equivalent training effort with both simulated^{53, 60} and real gaze-contingent prostheses⁶⁴, would result in substantial improvement in utility that would once again propel gaze-contingent systems to advantage. It is difficult to express how compelling the difference is between the two systems for naive subjects from our experiment: one is natural and fluid, and the other appears to require extensive training to establish even minimal utility.

With informal reports from our subjects that the Head Only condition induces mild nausea and vertigo, we can further speculate that one factor causing the relatively low retention rate of retinal prostheses in implanted patients^{65, 66} may be a lack of gaze contingency for certain devices. According to a recent report, two thirds of the Argus II patients used the device less than 5 h per week⁶⁶. Fortunately, it is straightforward to augment nearly every current visual prosthesis design that lacks gaze contingency with a gaze tracker and upgraded external image processing to improve the experience for recipients, such as done on an experimental basis for the Argus II¹⁵. Taken together, based on our results and our subjects’ informal reports, we strongly encourage all visual prosthesis designs that currently lack gaze contingency to incorporate the feature as quickly as is practical.

Post-implant rehabilitation and training. We expect training to be an important part of post-implant rehabilitation for foreseeable future prostheses as artificial vision is widely understood to be a crude approximation to normal vision, even for the most advanced, naturalistic designs. Our previous study on how reading ability changed with experience for a simulated thalamic visual prosthesis⁵³ showed that active training provided substantial benefit, even in the case of phosphene counts not far from what are currently in clinical use. Although beyond the scope of this study, training effects would certainly be expected in the Head-Only condition with additional experience. We provide a speculative prediction immediately below.

Based on the amount of improvement seen in earlier work from our laboratory⁵³, we can estimate the amount of improvement that would be expected here after extensive training. For the Full Gaze condition here, we have a close match to the conditions in the previous study—both included the same 2000-phosphene pattern, the same gaze tracker, the same gaze compensation algorithm, and observed highly similar initial performance—and thus we would expect reading accuracy averaged over the font sizes used here to reach 100% and equivalent acuity to improve from 1.0 to 0.7 logMAR in 10–15 sessions. For the Head Only condition, the closest match performance-wise is the lowest resolution pattern in that previous work, P₅₀₀/Low. Understanding that the correspondence between the two conditions is poor (beyond one being gaze contingent and the other head steered), we can speculate that the Head Only acuity here might improve from 1.6 to 1.3 logMAR after 40 sessions and the mean reading accuracy might rise from 8 to 50%, based on the improvement in P₅₀₀/Low acuity there. Under this highly speculative model, while performance under the Head Only condition here would improve with extensive training, it still would not match performance in the Full Gaze condition observed with naive subjects.

A recently published report⁴⁶ of interviews of patients undergoing the training necessary for the Argus II device, which lacks gaze contingency, suggests that it is a difficult and unpleasant experience. An earlier case report⁴⁷ describes similar training frustrations in one patient. Our results point to the lack of gaze contingency being a potentially important factor in these observations.

Phosphene pattern, field extent, and theoretical acuity. The phosphene pattern used in our simulations (Fig. 8) is generated by an electrode array that is like many from current clinical prostheses in that it has a constant density in the tissue where electrodes are implanted. For constant tissue density of electrodes, the spacing of phosphenes in the visual field varies from area to area: in the retina, it results in an even spacing of phosphenes, but in LGN and visual cortex, it results in a pattern that includes effects of tissue magnification to create a distribution that is center-weighted and representative of the endogenous profile of acuity versus eccentricity^{5, 67}.

Contemporary visual prostheses typically stimulate a limited extent of the visual field, but here, we have simulated a device that has electrodes that span the entirety of LGN, and thus creates phosphenes spanning the entirety of the visual field (Fig. 8). In this way, our simulation is intended to be forward-looking, rather than representative of devices in the clinic, and will not reflect the experience of a current prosthesis recipient, except insofar as to predict that the addition of gaze compensation to head-steered prostheses will result in an

improvement in patient experience and device utility. Combining observations that have been made by others in such devices^{15,52,54} and the results presented here, we believe this prediction to be well-founded.

The pattern used here, with 2000 phosphenes, has a theoretical acuity of 1.1 logMAR within the foveal area (the central two degrees) that falls off with eccentricity and is 2.7 logMAR at a point corresponding to the left and right edges of the monitor in our simulation when the subject's gaze is at the center. The central acuity in our pattern was therefore higher than the best reported values for Alpha AMS and IMS patients of 1.4 logMAR⁶⁸, and for Argus II patients of 1.6 logMAR⁶⁹, although of the two, only the latter device lacks gaze contingency.

Linguistic ability was not a factor. While the subjects were non-native English speakers, they were required to present certification of college-level ability, and the MNREAD sentences use a 3rd grade elementary school vocabulary. The observed high reading speeds and 100% accuracy in the Normal condition demonstrate that linguistic issues were not a limiting factor in our results.

Eye tracking and calibration in blind individuals. The inclusion of an eye tracking device in a visual prosthesis design that does not inherently include gaze compensation brings the question of tracker calibration. Normally, calibration of eye tracking devices is done in sighted subjects by activating targets at known locations, an option that is not generally available with blind patients. Nevertheless, there are advanced approaches that allow basic calibration using only external measurements of the eye that do not depend on subject perception^{70–72}. We expect these approaches combined with a refinement thorough a perceptual process such as the pointing paradigms extensively described in the literature^{15,36,45,73,74}, or staircase saccades, vestibulo-ocular reflex (VOR), smooth pursuit, or other continuous observation methods (see review by Kar and Corcoran⁷⁵) will prove sufficient to reap significant benefits from incorporating gaze-contingency into visual prosthesis devices. While there is some concern about abnormal gaze shifts in blind patients, Alpha IMS implant recipients have been observed to have relatively normal eye movements, suggesting a re-provisioning of visual input has a normalizing effect³².

High quality gaze tracking was critical to support the accuracy of our simulations, but it will also play a role in implementing gaze contingency in prosthetic devices that would otherwise be head-steered. Current commercial goggles-mounted eye trackers (e.g., Tobii Pro Glasses 3, Tobii, AB; Pupil Core, Pupil Labs, GmbH) are approaching the accuracies of the hardware used in this work, suggesting that designers of future gaze-compensated prostheses will have capable tools at their disposal.

Conclusion

Visual prosthesis devices that deliver a replacement for the visual information that is normally provided by the retina are constrained by the design of the visual pathway. The brain combines retinotopic information with information about the location of the eyes and the head to integrate stimuli from different gaze locations into a unified representation of the environment. As retinotopic information is perforce locked to gaze direction, artificial visual devices should optimally compensate their regions of interest to track instantaneous gaze position. We have demonstrated that in a simulation of artificial vision, including real-time gaze compensation makes the simulation easy to use, and reducing the simulation to a head-steered device that lacks awareness of eye position reduces the utility to near zero for naive subjects in a reading task. We conclude that adding eye tracking to existing devices should immediately improve usability and may impact patient retention.

Materials and methods

Participants. Twenty-three subjects took part in the study. The sample size was selected based on previous simulation studies from our laboratory with sighted individuals^{51,55}. Subjects were required to have self-reported normal or corrected-to-normal vision, no history of psychiatric or neurological impairment, and be able to read English text out loud. One participant was a native English speaker, while the rest ($n=22$) were native Greek speakers with sufficient English language skills for the modest requirements of the task (Level B2 according to the Common European Framework of Reference for Languages). One additional subject (a 24th, data from whom are not otherwise reported), was disqualified during the consenting process due to strabismus. Subjects were assigned pseudonyms for the purpose of anonymizing data collection and they either volunteered or received course credits for their participation in the study.

Ethics statement. The research in this report was performed according to a protocol approved by the Ethics Committee of the Department of History and Philosophy of Science at the University of Athens and the Institute Review Board at Massachusetts General Hospital and in compliance with the guidelines of the Declaration of Helsinki. This study was determined to contain minimal risk and informed consent was provided by each participant.

Apparatus. The system used in this experiment has substantial commonality with one that has been described in a previous publication⁵³ using techniques that have been described in detail⁵⁵. Subjects were seated in front of an apparatus that included a heads-free gaze tracker (EyeLink 1000+, SR Research, Inc.), a head tracking system (TrackClip Pro, Natural Point, Inc., as infrared target; *opentrack* software, TrackHat, Ltd., to perform tracking; PS3eye, Sony, Inc., as imaging camera), and associated computers for processing gaze and head position, along with stimulus generation and experiment control (Fig. 5).

Visual stimuli were presented to the subjects on an LCD computer monitor (PG279Q, Asus, Inc.) operating at 1600 × 900 resolution and 100 Hz video refresh, placed 70 cm away from the subject. At that distance, the 59.9 cm

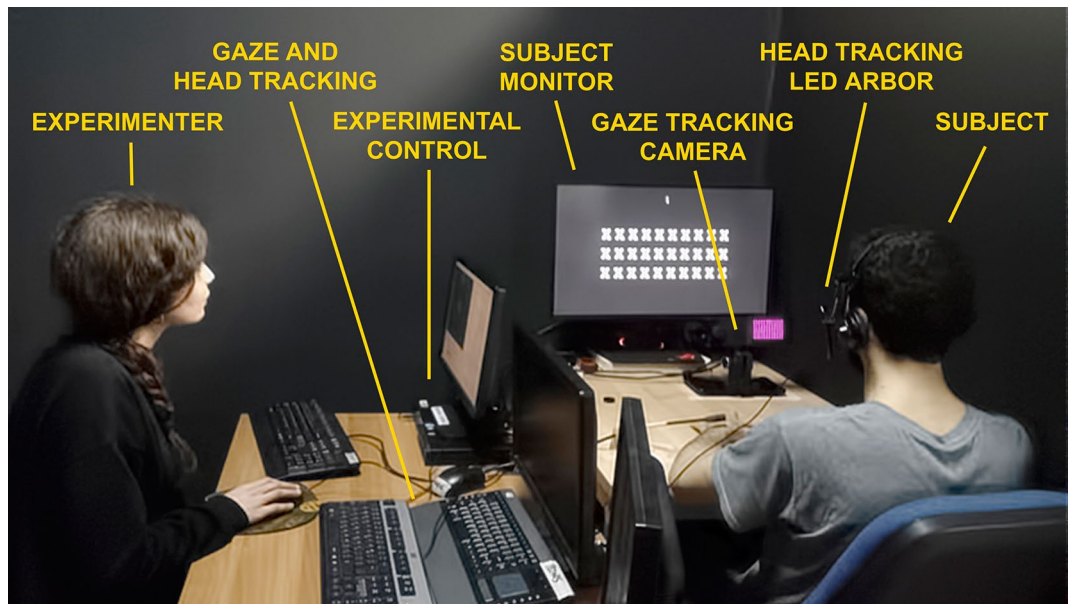


Figure 5. Apparatus. An apparatus was constructed to drive the simulation under the various viewing conditions. Subjects are seated at a table on which rests a high-refresh rate LCD monitor (ASUS, PG279Q, running at 100 Hz refresh), a gaze tracking camera (SR Research, EyeLink 1000+), a head position tracking camera (TrackHat, opentrack), and a microphone. As part of the head tracker, subjects wear a set of headphones that are used only to hold a three-LED arbor so that it faces the head-tracking camera (Natural Point, TrackClip Pro). Simple sentences are presented on the subject screen in a simulation of artificial vision using phosphenes that are stabilized on the subject's retina based on the instantaneous report from the gaze tracker. A virtual scene camera is steered either through gaze location (Full Gaze condition) or head position (Head Only condition) in the simulation. As the subjects read the sentences aloud, scores are kept by the experimenter on sheets that are not visible to the subjects. The vision simulation and experimental control are run from a single computer (Lenovo Tiny M710p).

by 33.6 cm extent of the monitor's display subtended 41 by 24 degrees in the subject's visual field. As described below, during the presentation of simulated phosphene vision, the display was updated on a frame-by-frame basis.

System latency was not directly measured, but is estimated to have been under 20 ms. The eye tracker was operating in 1000 Hz mode with 2-point averaging, for an effective 2 ms latency. The gaze position was sampled at the start of each 10 ms video refresh cycle and used to create a stimulus frame that was sent to the video monitor immediately prior to the next refresh for a total of 12 ms signal latency. The monitor itself was measured to have an additional 6 ms latency (Video Signal Input Lag Tester, Leo Bodnar Electronics, Ltd), which has been reported to be primarily pixel response time (TFT Central, http://www.tftcentral.co.uk/reviews/asus_rog_swift_pg279q.htm). Adding these three sources of latency together (2 + 10 + 6), gives us an estimate that we conservatively rate at 20 ms.

Design and stimuli. *Experimental conditions: primary experimental variable.* To investigate the effects of gaze contingency on reading performance in a simulation of artificial vision, we manipulated the gaze compensation mode. Specifically, we introduced three viewing conditions that were presented in mini-blocks (Fig. 6): (A) a *Normal* condition, where clear text was presented on the screen as a reference for normal reading behavior, (B) a *Full Gaze* artificial vision condition, where the simulated phosphenes were updated based on the position and direction of both the head and the eye, thus stabilizing the phosphene locations on the retina, and (C) a *Head Only* condition, where phosphenes were updated based on the position and direction of the head only. The conditions were presented in a mini-block fashion, as pilot data showed that participants encountered major difficulties in switching between Full Gaze and Head Only conditions on a trial-to-trial basis.

Letter size and text sentences: secondary experimental variables. Secondary experimental variables that were used to compute subject performance for each mini-block mode included font size of the text, carefully calibrated to be equivalent to 0.9 to 1.4 logMAR in steps of 0.1, and the text content of each sentence (Fig. 7). These secondary variables were used to generate psychometric curves and validate observations against previous reports^{53,55}.

Sentences were taken from the MNREAD corpus⁵⁷ and presented in the Times New Roman font. Sentences in the MNREAD corpus have 60 characters, including spaces but not including a final period, include no punctuation, and no proper nouns. They are presented across three lines (Fig. 9), and have been extensively studied to ensure they have little variation in readability⁵³.

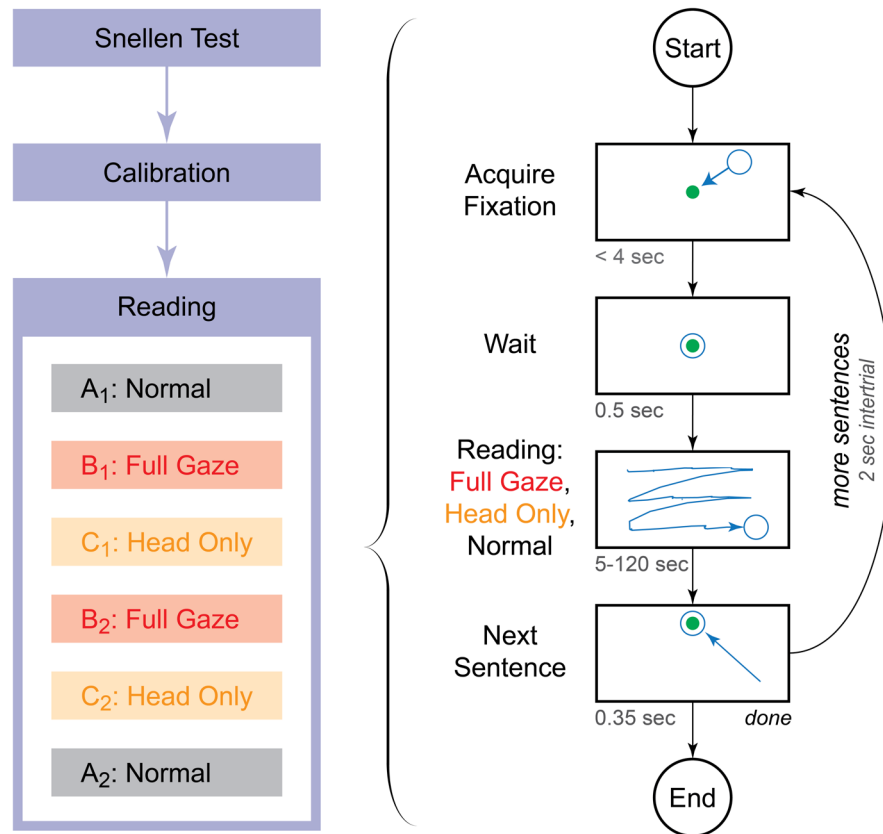


Figure 6. Overview of experimental design and procedure. (Left) Each subject's session started with a coarse assessment of visual acuity using a traditional Snellen chart. This test was to verify that subjects had largely normal vision, rather than to precisely measure their visual acuity, as the experimental requirements were well below ordinary acuity. Then, after subjects were seated in front of the apparatus, a quick calibration of the gaze position system was performed, typically lasting 2–3 min. Finally, the experiment itself proceeded with the presentation of the six mini-blocks in sequence, starting with a control condition mini-block, followed by four experimental mini-blocks, and then closing with a second control mini-block. (Right) The reading task used a trial-based structure where each trial began with the subject fixating a central point using normal vision. After a brief pause, a simple, three-line sentence was presented on the screen in one of the three conditions, depending on the mini-block: the Normal condition using natural vision, and the Full Gaze and Head Only conditions using phosphene-view simulation to present the stimulus. In both phosphene view conditions, the phosphenes were stabilized on the retina; in the Full Gaze condition, the simulated scene camera was steered by the combination of eye and head positions, whereas in the Head Only condition, the scene camera was steered only by head position. For all three conditions, the subjects read the sentence aloud as best they could, and then fixated the Next Sentence dot to proceed to the next trial.

Phosphene pattern. As the intent was to simulate the style of devices under investigation in our laboratory, the pattern of phosphenes used were heavily center-weighted, but fully span the visual field (Fig. 8), as has been previously discussed^{5, 67}. In particular, the phosphene pattern used here is based on a visual prosthesis design that places stimulating electrodes in the dorsal lateral geniculate nucleus (LGN). This thalamic approach affords advantages to retinal or cortical approaches that have been discussed in the literature^{4, 5, 23, 40, 51, 55, 63}, and include applicability to a wide range of disease conditions and causes of blindness. Importantly for the present work, the distribution of phosphenes in a given implant is expected to reflect the endogenous distribution of visual acuity across the visual field with resolution highest at the point of regard with small, densely packed phosphenes, falling away toward the periphery with large, sparse phosphenes (see Discussion).

Procedure. The experimental procedure was divided into three parts: (1) Snellen acuity assessment, (2) subject seating and calibration of the gaze tracker, and (3) the main experiment which was subdivided into the six mini-blocks (Fig. 6). A full session with a given subject took about 30 min.

Snellen acuity task. Each participant was administered a standard Snellen chart task to verify that they had approximately normal vision. Subjects stood at a measured mark, 3 m away from a wall-mounted Snellen chart. Normal overhead office lights were used to generate ordinary levels of illumination. The test was administered

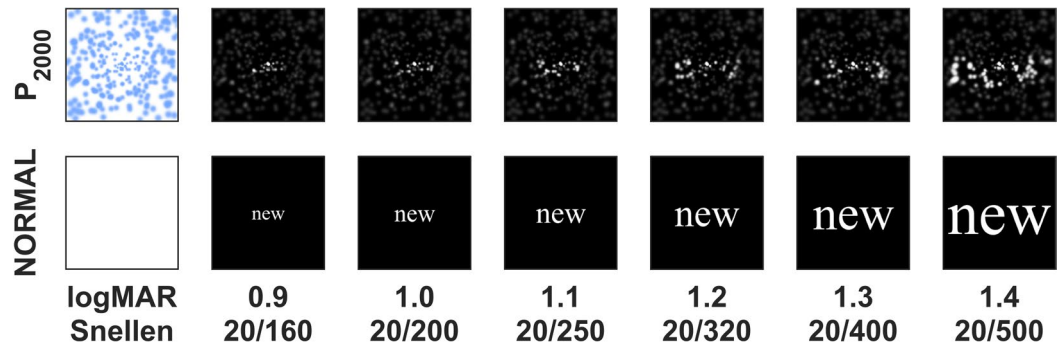


Figure 7. Stimulus conditions and font sizes. Example renderings of text used during the simulations is shown for the Normal condition (bottom row) where text is rendered at the native resolution of the screen, and the phosphene view conditions (top row) where text is shown as seen through a collection of 2000 phosphenes. The central part of the screen, approximately 10 degrees of visual angle across, is depicted for the six different font sizes used. Font sizes were calibrated to correspond to 0.9 logMAR (smallest) through 1.4 logMAR (largest). The central part of the phosphene pattern corresponding to the screen areas rendered here is shown in the upper left square, with each phosphene appearing as a light blue Gaussian on a white background. When this pattern is used as a filter on the white text/black background, it results in the images in the top row, for gaze and camera positions both at the center of the screen. As the gaze location moves around in either of the phosphene view conditions, the phosphene pattern is stabilized on the retina, but the scene camera is steered according to the testing condition of the mini-block (Full Gaze or Head Only).

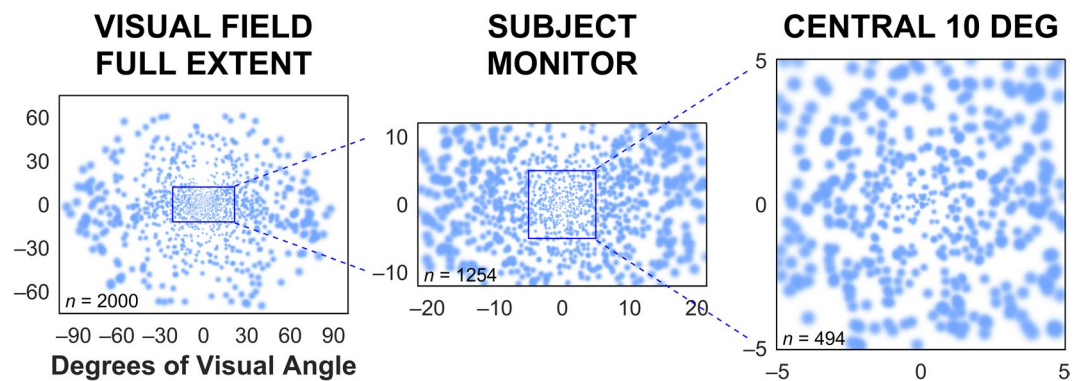


Figure 8. Phosphene pattern. The phosphene pattern used in this experiment contained 2000 phosphenes total, spread over the entire visual field in a center-weighted pattern that follows the natural profile of visual acuity from the center of gaze to the periphery^{5,67}, as seen in the left image. During the task, the phosphenes that fall on the subject monitor would be simulated. As the pattern is gaze-locked in retinotopic coordinates, the pattern is shifted with gaze movements during the simulation, and the number of phosphenes that fall on the monitor accordingly varies; that number would be maximum at about 1200 phosphenes when the gaze position was straight ahead, as seen in the middle image. In phosphene view mode, those phosphenes would be used as a filter on the image to be presented to the subject (see Figs. 7, 9). The central part of the visual field that carries the highest density of phosphenes is shown in the right image. For phosphene patterns in this class that model thalamic visual prostheses under development in our laboratory, the central 10 degrees typically has about one-quarter of the total phosphenes⁵¹.

binocularly with corrective lenses if the subject normally used them, and at a pace determined by the subject. Responses were converted to logMAR units.

Initial seating. After the Snellen screening, participants were brought to the experimental room, and seated comfortably in front of the subject monitor at a distance of about 70 cm from the display. Subjects wore both a bulls-eye sticker on their forehead to support frame-free gaze tracking, and a set of non-functional headphones onto which the head-tracking LED arbor was mounted. To improve the quality of tracking results, subjects were seated with their chest against the table and arms on the table to create a natural centering of head position without undue limits on head rotation. Instructions were given for the calibration task which was then performed, followed by instructions for the reading task which was then in turn performed.

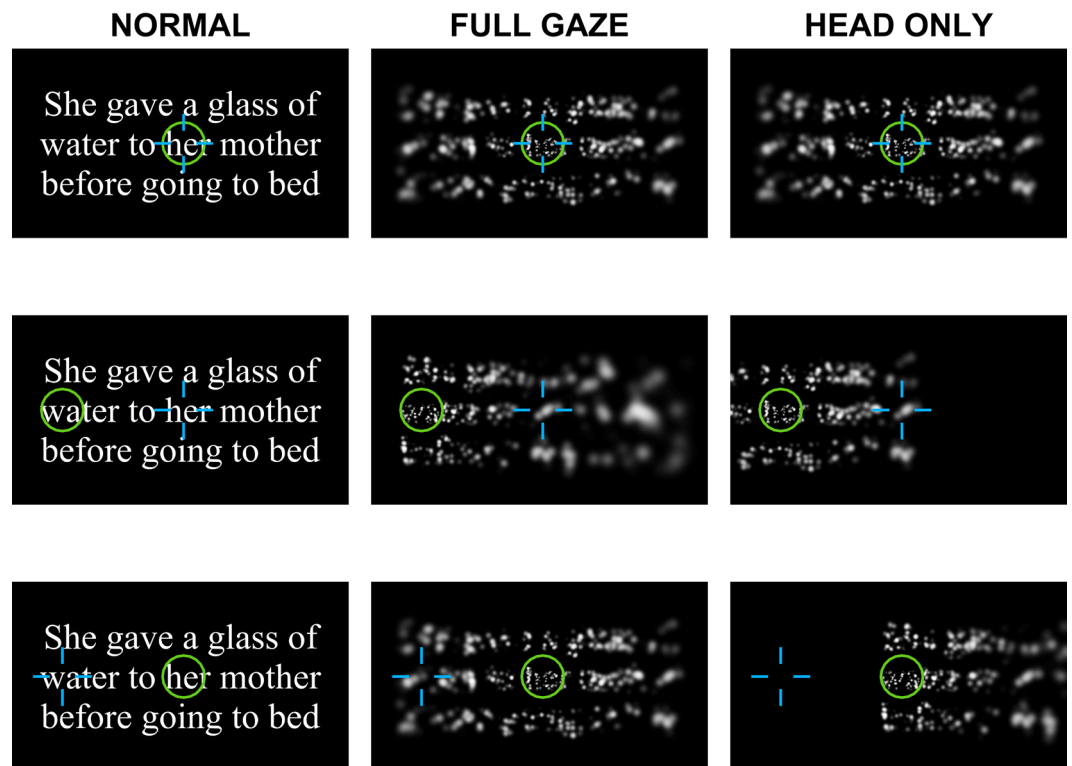


Figure 9. Screen captures from experimental conditions. Example screen shots are shown from the three conditions for an example sentence at the largest font size, 1.4 logMAR. The three columns correspond to the Normal (left column), Full Gaze (middle), and Head Only (right) conditions. The three rows correspond to three different alignment conditions for gaze (green circle) and head (blue cross) positions: the first (top row) contains the stimuli displayed under the three conditions when both the gaze and head positions are straight ahead, at the center of the screen where the subject is looking at the word /her/; the second (center) contains stimuli when the gaze is left (eyes rotated left), but the head is straight such that the subject's gaze is on /water/, but the camera remains directed at /her/; the third (bottom) is for the gaze straight (eyes counter-rotated right) and head left such that the gaze is again on /her/, but the camera is on /water/. When the eyes are deviated from straight forward in the head, the resulting disagreement between gaze and head position creates spatially incongruent stimuli for Head Only conditions that was found to be highly disconcerting to subjects and resulted in very poor performance compared to the Full Gaze conditions that are sensitive to eye position.

Calibration task. Calibration of the gaze tracking system has been described previously⁵³ and is summarized here. Participants were presented a series of small dots one at a time in an array of locations that spanned the screen and were instructed to fixate each as closely and as accurately as possible, maintaining fixation for the duration of each dot. The first thirteen presentations (one for each location) were used to trigger EyeLink 1000+ calibration while the subsequent 27 presentations (three instances for each of nine locations) were used to calibrate a secondary non-linear correction in the experimental software. The head position tracker did not require per-subject calibration as it used a fixed LED arbor of known geometry.

MNREAD-based reading task. The MNREAD task^{57, 58} assesses visual acuity by presenting a series of simple three-line sentences that are read aloud. Both reading speed and accuracy are tracked as the size of the font used to display the text is varied, typically resulting in logistic-like psychometrics where decreasing text size results in decreasing performance. The MNREAD task can be considered an alternative to single letter acuity tasks such as the familiar Snellen or Sloan charts that are common in clinical use. In the laboratory setting, we find MNREAD sentences are an efficient and effective way to collect acuity-related data with human subjects; the variant we have used here has been validated previously⁵⁵.

The main experiment consisted of three mini-blocks of 6 trials each (Fig. 6, left panel). Each mini-block was presented twice and the order of presentation was kept constant across participants. Each participant completed 36 trials in total. We used a mini-block structure, rather than a fully interleaved method, due to the unanticipated difficulty of the Head Only condition discovered during development. Within each mini-block, each of the six font sizes was presented once in pseudo-random order that varied from mini-block to mini-block, but was conserved across subjects. The mini-block sequence presented viewing conditions in order of increasing difficulty (Natural Vision, Full-Gaze, Head-Only) so as to ensure any bias introduced by the sequence would diminish, rather than enhance, any differences from condition to condition.

Each trial was subdivided into a series of four phases (Fig. 6, right panel), *Start*, *Pre-Stimulus*, *Reading*, and *End*. During the Start Phase, a fixation point appeared in the middle of the screen that the subject was required to foveate in order to engage the experiment. The color of the fixation point served as a visual cue informing participants about the viewing condition of the trial (white dot indicating the Normal condition; red dot, Full Gaze condition; orange dot, Head Only condition). Once foveated for the duration of the Pre-Stimulus Phase, the fixation point disappeared. After the fixation offset, the Reading Phase started with one of the sentences displayed with either no adulteration for the Normal condition (ordinary, plain text shown on the monitor), phosphene view with gaze-compensation for the Full Gaze condition, and phosphene view without sensitivity to eye position for the Head Only condition, each at one of the font sizes (Figs. 7, 9). An additional *Next Sentence* dot, which matched the color of the trial's fixation point, was displayed near the top center of the screen. Subjects were instructed to read each sentence out loud as best as they could while skipping words they were unable to discern, or to state that they were unable to read the text at all. Subjects were instructed to look at the Next Sentence dot in order to advance to the next sentence. Subjects could take as long as they wanted, consistent with reading quickly and accurately. Once subjects foveated the Next Sentence dot for 350 ms, the trial entered the End Phase, the screen was blanked, and a 2000 ms pause provided an intertrial interval before the onset of next trial. Each trial used a different MNREAD sentence, but the sequence of sentences and conditions was maintained across subjects. An audio recording was made for the entirety of each experiment.

To ensure an accurate simulation and reading performance assessment, the text for each trial was rendered off-screen with compensation for the head-to-screen distance as measured during the pre-stimulus portion of the trial. An ideal simulation might re-render the reference image of the text with each frame, but with the hardware at hand, that would have unacceptably slowed the simulation.

However, there was sufficient computational power available to adjust the diameter of phosphenes as drawn on the screen to the per-frame viewing distance so that the sizes on the retina were approximately constant. This compensation was computationally inexpensive, and thus readily incorporated without impact on the simulation.

Accurate measurements of reading speed required an accurate assessment of the time spent in reading for each trial. During the developmental phase of the experiment it became clear that if viewing conditions were interleaved on a trial-by-trial basis, a substantial amount of time was being spent at the start of each trial for the subject to figure out the current experimental mode. For Full-Gaze trials, this probing did not add as much of an artifactual latency as for Head-Only trials; thus, to reduce bias, a mini-block structure was selected for collection of the primary data reported here.

Data analysis. *Reading accuracy.* Each subject's orally provided responses were compared to a printed version of the MNREAD sentences in order to score the number of correctly read words for each sentence. Counts were then normalized by sentence length and multiplied by 100 to find the reading accuracy, or percentage of correctly read words for each combination of gaze compensation mode and font size. Each combination was presented twice during a session (once per mini-block), but with two different sentences, and the mean value over the two presentations was used for the subject's performance on that combination. Control comparisons were performed between the first and second mini-block for each condition to rule out longitudinal effects. An equivalent visual acuity was computed by measuring the 50% point along a logistic curve fitted to reading accuracy versus font size. As reading accuracy has a bounded range, median values with 16/84 percentile ranges are reported in lieu of means and standard deviations.

Reading speed. The reading speed was computed as the number of correctly read words per minute, for each combination, again averaged over the two presentations. The length of time for a given trial was measured from the time the sentence text appeared on the screen until the subject foveated the Next Trial dot. In Head Only trials, detection of Next Trial foveation was based on normal gaze position on the screen, rather than head-steered aiming, and was computed independently from stimulus presentation. As with Reading Accuracy, reported values are medians with 16/84 percentile ranges.

Expectations. For both reading accuracy and reading speed, we expected to observe significantly higher performance for the Full Gaze compared to the Head Only condition. Reading accuracy is not normally distributed because of its limited range, an effect that becomes especially apparent near 0% or 100%. Reading speed has a similar lack of normality for very low values like the ones we observed. To test our hypotheses, we determined statistical significance using Wilcoxon rank sum tests as the assumption of normality was not met using a Kolmogorov–Smirnov test for normality in many cases. For statistical comparisons, a *p* value below 0.05 was considered to be significant.

Received: 21 August 2020; Accepted: 23 March 2021

Published online: 27 May 2021

References

1. Chuang, A. T., Margo, C. E. & Greenberg, P. B. Retinal implants: a systematic review—Table 1. *Br. J. Ophthalmol.* **98**(7), 852–856. <https://doi.org/10.1136/bjophthalmol-2013-303708> (2014).
2. Donaldson, N. & Brindley, G. S. The historical foundations of bionics. In *Neurobionics: The Biomedical Engineering of Neural Prostheses* (ed. Shepherd, R. K.) 1–37 (Wiley, 2016). <https://doi.org/10.1002/9781118816028.ch1>.
3. Goetz, G. A. & Palanker, D. V. Electronic approaches to restoration of sight. *Rep. Prog. Phys.* **79**(9), 096701. <https://doi.org/10.1088/0034-4885/79/9/096701> (2016).
4. Mirochnik, R. M. & Pezaris, J. S. Contemporary approaches to visual prostheses. *Mil. Med. Res.* **6**(1), 19. <https://doi.org/10.1186/s40779-019-0206-9> (2019).

5. Pezaris, J. S. & Eskandar, E. N. Getting signals into the brain: visual prosthetics through thalamic microstimulation. *Neurosurg. Focus* **27**(1), E6. <https://doi.org/10.3171/2009.4.FOCUS0986> (2009).
6. Schiller, P. H. & Tehovnik, E. J. Visual prosthesis. *Perception* **37**(10), 1529–1559. <https://doi.org/10.1068/p6100> (2008).
7. Edwards, T. L. *et al.* Assessment of the electronic retinal implant Alpha AMS in restoring vision to blind patients with end-stage retinitis pigmentosa. *Ophthalmology* **125**(3), 432–443. <https://doi.org/10.1016/j.ophtha.2017.09.019> (2018).
8. Stingl, K. *et al.* Artificial vision with wirelessly powered subretinal electronic implant Alpha-IMS. *Proc. R. Soc. B Biol. Sci.* **280**(1757), 20130077. <https://doi.org/10.1098/rspb.2013.0077> (2013).
9. Stingl, K. *et al.* Subretinal visual implant Alpha IMS: clinical trial interim report. *Vis. Res.* **111**, 149–160. <https://doi.org/10.1016/j.visres.2015.03.001> (2015).
10. Stingl, K. *et al.* Functional outcome in subretinal electronic implants depends on foveal eccentricity. *Investig. Ophthalmol. Vis. Sci.* **54**(12), 7658. <https://doi.org/10.1167/iovs.13-12835> (2013).
11. Zrenner, E. Will retinal implants restore vision? *Science* **295**(5557), 1022–1025. <https://doi.org/10.1126/science.1067996> (2002).
12. Zrenner, E. *et al.* Subretinal electronic chips allow blind patients to read letters and combine them to words. *Proc. R. Soc. B Biol. Sci.* **278**(1711), 1489–1497. <https://doi.org/10.1098/rspb.2010.1747> (2011).
13. Zrenner, E. *et al.* The subretinal implant ALPHA: implantation and functional results. In *Artificial Vision* (ed. Gabel, V. P.) 65–83 (Springer, 2017). https://doi.org/10.1007/978-3-319-41876-6_6.
14. Ahuja, A. K. *et al.* Blind subjects implanted with the Argus II retinal prosthesis are able to improve performance in a spatial-motor task. *Br. J. Ophthalmol.* **95**(4), 539–543. <https://doi.org/10.1136/bjo.2010.179622> (2011).
15. Caspi, A. *et al.* Eye movement control in the Argus II retinal-prosthesis enables reduced head movement and better localization precision. *Invest. Ophthalmol. Vis. Sci.* **59**(2), 792–802 (2018).
16. Fernández, E. & Normann, R. A. CORTIVIS approach for an intracortical visual prostheses. In *Artificial Vision* (ed. Gabel, V. P.) 191–201 (Springer, 2017). https://doi.org/10.1007/978-3-319-41876-6_15.
17. Hornig, R. *et al.* Pixium vision: first clinical results and innovative developments. In *Artificial Vision* (ed. Gabel, V. P.) 99–113 (Springer, 2017). https://doi.org/10.1007/978-3-319-41876-6_8.
18. Kelly, S. K., Shire, D. B., Chen, J., Gingerich, M. D., Cogan, S. F., Drohan, W. A., Ellersick, W., Krishnan, A., Behan, S., Wyatt, J. L., & Rizzo, J. F. (2013). Developments on the Boston 256-channel retinal implant. In *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 1–6. <https://doi.org/10.1109/ICMEW.2013.6618445>
19. Menzel-Severing, J. *et al.* Implantation and explantation of an active epiretinal visual prosthesis: 2-year follow-up data from the EPIRET3 prospective clinical trial. *Eye* **26**(4), 501–509. <https://doi.org/10.1038/eye.2012.35> (2012).
20. Shivdasani, M. N. *et al.* Evaluation of stimulus parameters and electrode geometry for an effective suprachoroidal retinal prosthesis. *J. Neural Eng.* **7**(3), 036008. <https://doi.org/10.1088/1741-2560/7/3/036008> (2010).
21. Lewis, P. M., Ackland, H. M., Lowery, A. J. & Rosenfeld, J. V. Restoration of vision in blind individuals using bionic devices: a review with a focus on cortical visual prostheses. *Brain Res.* **1595**, 51–73. <https://doi.org/10.1016/j.brainres.2014.11.020> (2015).
22. Yue, L., Weiland, J. D., Roska, B. & Humayun, M. S. Retinal stimulation strategies to restore vision: fundamentals and systems. *Prog. Retin. Eye Res.* **53**, 21–47. <https://doi.org/10.1016/j.preteyeres.2016.05.002> (2016).
23. Paraskevoudi, N. & Pezaris, J. S. Eye movement compensation and spatial updating in visual prosthetics: mechanisms, limitations and future directions. *Front. Syst. Neurosci.* **12**, 73. <https://doi.org/10.3389/fnsys.2018.00073> (2019).
24. Burr, D. Eye movements: keeping vision stable. *Curr. Biol.* **14**(5), R195–R197. <https://doi.org/10.1016/j.cub.2004.02.020> (2004).
25. Inaba, N. & Kawano, K. Eye position effects on the remapped memory trace of visual motion in cortical area MST. *Sci. Rep.* **6**(1), 22013. <https://doi.org/10.1038/srep22013> (2016).
26. Klier, E. M. & Angelaki, D. E. Spatial updating and the maintenance of visual constancy. *Neuroscience* **156**(4), 801–818. <https://doi.org/10.1016/j.neuroscience.2008.07.079> (2008).
27. Rao, H. M., Mayo, J. P. & Sommer, M. A. Circuits for presaccadic visual remapping. *J. Neurophysiol.* **116**(6), 2624–2636. <https://doi.org/10.1152/jn.00182.2016> (2016).
28. Brickner, R. M. Oscillopsia: a new symptom commonly occurring in multiple sclerosis. *Arch. Neurol. Psychiatry* **36**(3), 586. <https://doi.org/10.1001/archneurpsyc.1936.0226090139009> (1936).
29. Gresty, M. A., Hess, K. & Leech, J. Disorders of the vestibulo-ocular reflex producing oscillopsia and mechanisms compensating for loss of labyrinthine function. *Brain* **100**(4), 693–716. <https://doi.org/10.1093/brain/100.4.693> (1977).
30. Evans, N. The significance of nystagmus. *Eye* **3**(6), 816–832. <https://doi.org/10.1038/eye.1989.126> (1989).
31. Hafed, Z. M. & Krauzlis, R. J. Microsaccadic suppression of visual bursts in the primate superior colliculus. *J. Neurosci.* **30**(28), 9542–9547. <https://doi.org/10.1523/JNEUROSCI.1137-10.2010> (2010).
32. Hafed, Z. M., Stingl, K., Bartz-Schmidt, K.-U., Gekeler, F. & Zrenner, E. Oculomotor behavior of blind patients seeing with a subretinal visual implant. *Vis. Res.* **118**, 119–131. <https://doi.org/10.1016/j.visres.2015.04.006> (2016).
33. Leopold, D. A. & Logothetis, N. K. Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Exp. Brain Res.* **123**(3), 341–345. <https://doi.org/10.1007/s002210050577> (1998).
34. Kagan, I. & Hafed, Z. M. Active vision: microsaccades direct the eye to where it matters most. *Curr. Biol.* **23**(17), R712–R714. <https://doi.org/10.1016/j.cub.2013.07.038> (2013).
35. Coppola, D. & Purves, D. The extraordinarily rapid disappearance of entopic images. *Proc. Natl. Acad. Sci.* **93**(15), 8001–8004. <https://doi.org/10.1073/pnas.93.15.8001> (1996).
36. Brindley, G. S. & Lewin, W. S. The sensations produced by electrical stimulation of the visual cortex. *J. Physiol.* **196**(2), 479–493. <https://doi.org/10.1113/jphysiol.1968.sp008519> (1968).
37. Bradley, D. C. *et al.* Visuotopic mapping through a multichannel stimulating implant in primate V1. *J. Neurophysiol.* **93**(3), 1659–1670. <https://doi.org/10.1152/jn.01213.2003> (2005).
38. Davis, T. S. *et al.* Spatial and temporal characteristics of V1 microstimulation during chronic implantation of a microelectrode array in a behaving macaque. *J. Neural Eng.* **9**(6), 065003. <https://doi.org/10.1088/1741-2560/9/6/065003> (2012).
39. Dobbela, W. H. & Mladejovsky, M. G. Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. *J. Physiol.* **243**(2), 553–576. <https://doi.org/10.1113/jphysiol.1974.sp010766> (1974).
40. Pezaris, J. S. & Reid, R. C. Demonstration of artificial visual percepts generated through thalamic microstimulation. *Proc. Natl. Acad. Sci.* **104**(18), 7670–7675. <https://doi.org/10.1073/pnas.0608563104> (2007).
41. Schmidt, E. M. *et al.* Feasibility of a visual prosthesis for the blind based on intracortical micro stimulation of the visual cortex. *Brain* **119**(2), 507–522. <https://doi.org/10.1093/brain/119.2.507> (1996).
42. Sinclair, N. C. *et al.* The appearance of phosphenes elicited using a suprachoroidal retinal prosthesis. *Investig. Ophthalmol. Vis. Sci.* **57**(11), 4948. <https://doi.org/10.1167/iovs.15-18991> (2016).
43. Stronks, H. C. & Dagnelie, G. Phosphene mapping techniques for visual prostheses. In *Visual Prosthetics* (ed. Dagnelie, G.) 367–383 (Springer, 2011). https://doi.org/10.1007/978-1-4419-0754-7_19.
44. Tehovnik, E. J. & Slocum, W. M. Phosphene induction by microstimulation of macaque V1. *Brain Res. Rev.* **53**(2), 337–343. <https://doi.org/10.1016/j.brainresrev.2006.11.001> (2007).
45. Veraart, C. *et al.* Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Res.* **813**(1), 181–186. [https://doi.org/10.1016/S0006-8993\(98\)00977-9](https://doi.org/10.1016/S0006-8993(98)00977-9) (1998).

46. Erickson-Davis, C. & Korzybska, H. What do blind people “see” with retinal prostheses? Observations and qualitative reports of epiretinal implant users. *PLoS ONE* **16**(2), e0229189. <https://doi.org/10.1371/journal.pone.0229189> (2021).
47. Brady-Simmons, C., Van Der Biest, R. & Bozeman, L. Miami lighthouse for the blind and visually impaired case study: vision rehabilitation for the first Florida resident to receive the Argus II “bionic eye”. *J. Vis. Impair. Blind.* **110**(3), 177–181. <https://doi.org/10.1177/0145482X1611000304> (2016).
48. Sabbah, N. *et al.* Importance of eye position on spatial localization in blind subjects wearing an Argus II retinal prosthesis. *Invest. Ophthalmol. Vis. Sci.* **55**(12), 8259–8266. <https://doi.org/10.1167/iovs.14-15392> (2014).
49. Prabhu, D. *et al.* Effect of camera position on egocentric localisation with simulated prosthetic vision. *Eng. Res. Express* **3**(1), 015016. <https://doi.org/10.1088/2631-8695/abcd51> (2021).
50. Titchener, S. A. *et al.* Oculomotor responses to dynamic stimuli in a 44-channel suprachoroidal retinal prosthesis. *Transl. Vis. Sci. Technol.* **9**(13), 31. <https://doi.org/10.1167/tvst.9.13.31> (2020).
51. Bourkiza, B., Vurro, M., Jeffries, A. & Pezaris, J. S. Visual acuity of simulated thalamic visual prostheses in normally sighted humans. *PLoS ONE* **8**(9), e73592. <https://doi.org/10.1371/journal.pone.0073592> (2013).
52. McIntosh, B. P. (2015). *Intraocular and extraocular cameras for retinal prostheses: effects of foveation by means of visual prosthesis simulation.* <http://digitalibrary.usc.edu/cdm/ref/collection/p15799coll3/id/530967>. Accessed 22 July 2020.
53. Rasia, K. E. K. & Pezaris, J. S. Improvement in reading performance through training with simulated thalamic visual prostheses. *Sci. Rep.* **8**(1), 16310. <https://doi.org/10.1038/s41598-018-31435-0> (2018).
54. Titchener, S. A., Shivdasani, M. N., Fallon, J. B. & Petoe, M. A. Gaze Compensation as a technique for improving hand-eye coordination in prosthetic vision. *Transl. Vis. Sci. Technol.* **7**(1), 2. <https://doi.org/10.1167/tvst.7.1.2> (2018).
55. Vurro, M., Crowell, A. M. & Pezaris, J. S. Simulation of thalamic prosthetic vision: reading accuracy, speed, and acuity in sighted humans. *Front. Hum. Neurosci.* **8**, 816. <https://doi.org/10.3389/fnhum.2014.00816> (2014).
56. da Cruz, L. *et al.* The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss. *Br. J. Ophthalmol.* **97**(5), 632–636. <https://doi.org/10.1136/bjophthalmol-2012-301525> (2013).
57. Mansfield, J. S., Ahn, S. J., Legge, G. E. & Luebker, A. A new reading-acuity chart for normal and low vision. *Ophthalmic Vis. Opt. Noninvasive Assess. Vis. Syst. Tech. Dig.* **3**(232–235), 232–235 (1993).
58. Crossland, M. D., Legge, G. E. & Dakin, S. C. The development of an automated sentence generator for the assessment of reading speed. *Behav. Brain Funct.* **4**(1), 14. <https://doi.org/10.1186/1744-9081-4-14> (2008).
59. Stronks, H. C. & Dagnelie, G. The functional performance of the Argus II retinal prosthesis. *Expert Rev. Med. Devices* **11**(1), 23–30. <https://doi.org/10.1586/17434440.2014.862494> (2014).
60. Sommerhalder, J. *et al.* Simulation of artificial vision: II. Eccentric reading of full-page text and the learning of this task. *Vis. Res.* **44**(14), 1693–1706. <https://doi.org/10.1016/j.visres.2004.01.017> (2004).
61. Gilchrist, I. D., Brown, V. & Findlay, J. M. Saccades without eye movements. *Nature* **390**(6656), 130–131. <https://doi.org/10.1038/36478> (1997).
62. Barry, M. P. & Dagnelie, G. Hand-camera coordination varies over time in users of the Argus II retinal prosthesis system. *Front. Syst. Neurosci.* **10**, 41. <https://doi.org/10.3389/fnsys.2016.00041> (2016).
63. Killian, N. J., Vurro, M., Keith, S. B., Kyada, M. J. & Pezaris, J. S. Perceptual learning in a non-human primate model of artificial vision. *Sci. Rep.* **6**(1), 36329. <https://doi.org/10.1038/srep36329> (2016).
64. Cehajic Kapetanovic, J. *et al.* Highest reported visual acuity after electronic retinal implantation. *Acta Ophthalmol.* **98**(7), 736–740. <https://doi.org/10.1111/aos.14443> (2020).
65. Garcia, S., Petrini, K., Rubin, G. S., Da Cruz, L. & Nardini, M. Visual and non-visual navigation in blind patients with a retinal prosthesis. *PLoS ONE* **10**(7), e0134369. <https://doi.org/10.1371/journal.pone.0134369> (2015).
66. Sommerhalder, J. & Pérez-Fornos, A. Prospects and limitations of spatial resolution. In *Artificial Vision* (ed. Gabel, V. P.) 29–45 (Springer, 2017). https://doi.org/10.1007/978-3-319-41876-6_4.
67. Pezaris, J. S. & Reid, R. C. Simulations of electrode placement for a thalamic visual prosthesis. *IEEE Trans. Biomed. Eng.* **56**(1), 172–178. <https://doi.org/10.1109/TBME.2008.2005973> (2009).
68. Stingl, K. *et al.* Interim results of a multicenter trial with the new electronic subretinal implant Alpha AMS in 15 patients blind from inherited retinal degenerations. *Front. Neurosci.* **11**, 445. <https://doi.org/10.3389/fnins.2017.00445> (2017).
69. Humayun, M. S. *et al.* Interim results from the international trial of second sight’s visual prosthesis. *Ophthalmology* **119**(4), 779–788. <https://doi.org/10.1016/j.ophtha.2011.09.028> (2012).
70. Ramanaukas, N. Calibration of video-oculographical eye-tracking system. *Elektronika Ir Elektrotechnika* **7**(8), 65–68 (2006).
71. Zhu, Z., Ji, Q., & Bennett, K. P. (2006). Nonlinear eye gaze mapping function estimation via support vector regression. In *18th International Conference on Pattern Recognition (ICPR06)*, 1132–1135. <https://doi.org/10.1109/ICPR.2006.864>
72. Barsingerhorn, A. D., Boonstra, F. N. & Goossens, J. Development and validation of a high-speed stereoscopic eyetracker. *Behav. Res. Methods* **50**(6), 2480–2497. <https://doi.org/10.3758/s13428-018-1026-7> (2018).
73. Dobelle, W. H., Turkel, J., Henderson, D. C. & Evans, J. R. Mapping the representation of the visual field by electrical stimulation of human visual cortex. *Am. J. Ophthalmol.* **88**(4), 727–735. [https://doi.org/10.1016/0002-9394\(79\)90673-1](https://doi.org/10.1016/0002-9394(79)90673-1) (1979).
74. Everitt, B. S. & Rushton, D. N. A method for plotting the optimum positions of an array of cortical electrical phosphenes. *Biometrics* **34**(3), 399–410 (1978).
75. Kar, A. & Corcoran, P. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms. *IEEE Access* **5**, 16495–16519. <https://doi.org/10.1109/ACCESS.2017.2735633> (2017).

Author contributions

Both authors designed the experiment. N.P. collected the data. Both authors analyzed the data, and wrote the manuscript text. J.P. prepared the figures. Both authors reviewed the manuscript.

Funding

This work was supported by: the William M. Wood Foundation, Bank of America, Trustee; the NIMA Foundation; Peter D. Pezaris; Ana and Philippe Laffont, John and Genevive Wolfe; William and Lori Goldenthal; Michael Gersh; and Richard Burkitt. The content is solely the responsibility of the authors and does not necessarily represent views of the sponsors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.S.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021