



OPEN

## Reading text works better than watching videos to improve acuity in a simulation of artificial vision

Katerina Eleonora K. Rassia<sup>1</sup>, Konstantinos Moutoussis<sup>1</sup> & John S. Pezaris<sup>2,3</sup>✉

Simulated artificial vision is used in visual prosthesis design to answer questions about device usability. We previously reported a striking increase in equivalent visual acuity with daily use of a simulation of artificial vision in an active task, reading sentences, that required high levels of subject engagement, but passive activities are more likely to dominate post-implant experience. Here, we investigated the longitudinal effects of a passive task, watching videos. Eight subjects used a simulation of a thalamic visual prosthesis with 1000 phosphenes to watch 23 episodes of classic American television in daily, 25-min sessions, for a period of 1 month with interspersed reading tests that quantified reading accuracy and reading speed. For reading accuracy, we found similar dynamics to the early part of the learning process in our previous report, here leading to an improvement in visual acuity of  $0.15 \pm 0.05$  logMAR. For reading speed, however, no change was apparent by the end of training. We found that single reading sessions drove about twice the improvement in acuity of single video sessions despite being only half as long. We conclude that while passive viewing tasks may prove useful for post-implant rehabilitation, active tasks are likely to be preferable.

Contemporary visual prostheses provide only a crude approximation to normal vision, thus post-implant therapies play an important role in an over-all treatment plan. The ideal rehabilitation strategy for patients receiving visual prostheses remains an open question: what sort of activities would best assist the recipients to adapt to their new visual modality? To help answer such questions, we have previously studied improvements in visual acuity through near-daily use of an active, reading task with normal, sighted subjects viewing a simulation of artificial vision<sup>1</sup>. Here, we extend that work to study the effects of a passive, video viewing task that does not require the same levels of engagement and concentration. We postulate that at-home passive tasks are likely to occupy a larger fraction of daily living than clinic-based active tasks, so it may prove advantageous to incorporate passive tasks into a comprehensive therapeutic strategy.

**Predicting influence of passive tasks versus active tasks.** The idea that active engagement is required to drive Visual Perceptual Learning (VPL) is rooted in the intuitive notion that unbridled visual plasticity needs to be tempered to prevent the constant barrage of visual input leading to an undesirable outcome. The tempering force that would modulate VPL is canonically thought to be the conscious effort of attention (see review by Sasaki and colleagues<sup>2</sup>), but many investigations have now shown that even without attention, VPL remains possible<sup>3–8</sup>. Similar modulatory roles are thought to be played by motor action that reinforces accompanying visual perception<sup>9</sup> and feedback as to the correctness of responses<sup>10</sup>, both factors serving to amplify VPL. The effect is seen for feedback even when that feedback is only internal such as successful recognition of an object<sup>11,12</sup>. Underscoring the importance of modulation of VPL are experiments that bridge the gap between engagement and passivity by recording active experience of one group of subjects in a visually-based virtual reality exploration task and then playing those recordings to a second group of subjects to create passive experiences<sup>13–16</sup>. This comparison using otherwise identical stimuli showed that the effects of active versus passive engagement, such as the ability to identify and remember visual targets or features, are best characterized as graded rather than

<sup>1</sup>Cognitive Science Laboratory, Department of History and Philosophy of Science, National and Kapodistrian University of Athens, Athens, Greece. <sup>2</sup>Department of Neurosurgery, Massachusetts General Hospital, Boston, MA, USA. <sup>3</sup>Department of Neurosurgery, Harvard Medical School, Boston, MA, USA. ✉email: pezaris.john@mgh.harvard.edu

absolute. We are left, therefore, with the expectation that a passive task lacking feedback would drive VPL more slowly than an active task that included it.

**Quantitative assessment of performance.** To predict the potential for improving device utility through VPL, and to suggest rehabilitation training profiles, the field has used simulations of artificial vision with normal-sighted subjects performing psychophysical tasks. These tasks typically employ simple quantifiable behaviours (visual recognition, reading, visuo-motor interaction) in virtual reality setups that simulate the way vision would appear through a prosthesis to measure some aspect of visual perception as a model for blind individuals with implanted devices. Like other researchers, we use visual acuity as the primary performance metric as it is widely understood and recognized in scientific and clinical settings, as well as among the lay population.

Adaptation to artificial vision has been directly studied, or observed as a secondary finding, in many previous studies using active tasks. Experimental paradigms that have been used include visually-guided mobility and navigation<sup>17–20</sup>, object and face recognition<sup>21,22</sup>, letter recognition<sup>23–26</sup>, reading<sup>1,27–31</sup> or combinations of the above<sup>32,33</sup>. Reading tasks<sup>34</sup> in particular have proven to be a robust means to assess visual acuity in our laboratory<sup>1,31,35</sup>, and we continue their use here to measure the effects of training.

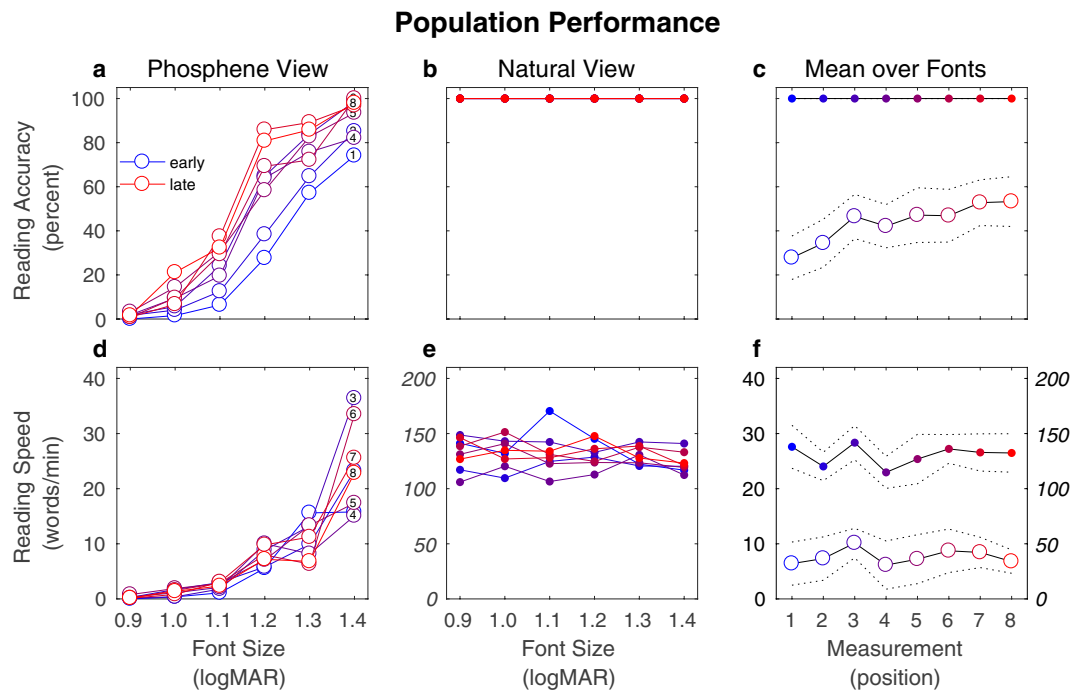
**Previous reports of learning effects in simulations of artificial vision.** Through a line of inquiry using simulated artificial vision, our laboratory has observed instances of VPL in both human<sup>1,25,31</sup> and animal models<sup>26</sup>. Our most extensive study thus far has been a longitudinal experiment<sup>1</sup> that assessed both accuracy and speed of reading performance through a simulation of artificial vision<sup>31,36</sup>. Subjects read 40 novel sentences per day conforming to the MNREAD criteria<sup>34</sup> for 40 sessions over approximately 8 weeks. This training resulted in a general doubling of equivalent visual acuity across the population, with substantial improvements in both reading accuracy and reading speed. Prior to that work, hints of learning effects had been found over a brief period of exposure with humans<sup>25</sup> and more substantial effects over longer periods with non-human primates<sup>26</sup>.

**Learning effects with reading tasks from other laboratories.** Other groups have found that reading under simulations of artificial vision improves with practice for a wide range of conditions. In an early study, Hayes and colleagues found that multiple sessions lead to a twofold increase in reading speed with a hand-manipulated camera<sup>32</sup>. Sommerhalder and colleagues observed impressive increases in reading accuracy and reduced response times in reading 4-letter words, after 1 month of daily training with a retinally-stabilized eccentric field<sup>28</sup>. The same group expanded their study to full-page reading and reported improvements with daily training that asymptoted after 2 months<sup>29</sup>. Dagnelie and colleagues also found that reading short paragraphs of text through a pixelized display benefits from practice for a variety of difficulty levels<sup>30</sup>. Fu and colleagues reported that a similar brief course of daily practice improved reading performance across difficulty levels<sup>24</sup>. Pérez Fornos and colleagues demonstrated that training of more than 1 month allowed subjects to read at 15° eccentricity with the same accuracy as with central reading, but interestingly also with improved performance on other visuo-motor tasks<sup>18</sup>.

**Examples of VPL in non-reading tasks from other groups.** In addition to reports using reading tasks, examples of VPL for a wide range of non-reading tasks have been published in studies that simulated artificial vision, typically demonstrating measurable effects within a small number of sessions. For instance, Xia and colleagues reported that multi-object recognition improves after 5 days of 2-h sessions<sup>22</sup>, and Chen and colleagues reported the ability to identify Landolt C orientation plateaus between 15 to 20 sessions<sup>23</sup>. Dagnelie and colleagues described decreasing error rates and completion times for object counting and placement tasks using checkers on a checkerboard with 17 or fewer daily 1-h sessions<sup>17</sup>. In a related study, Srivastava and colleagues described quite substantial decreases in completion time for a similar object placement task along with maze explorations<sup>19</sup>. In contrast, van Rheede and colleagues found trends toward improvement for object placement and wayfinding with three sessions that did not obtain significance<sup>33</sup>, and without confidence that they measured improved perception rather than an increased familiarity with the practiced tasks<sup>20</sup>. Finally, Thompson and colleagues found no change in accuracy, but a substantial decrease in response time in a single session of nearly 200 trials of a face recognition task<sup>21</sup>.

**Observations of learning effects in clinical trials of artificial vision.** A small handful of reports have studied learning in a clinical setting with recipients of implanted retinal visual prostheses. Patients with the Alpha IMS or AMS devices (Retina Implant AG) showed substantial learning effects on a wide range of behavioral tests<sup>37</sup> that span from generally improved visuomotor abilities or elimination of nystagmus initially preventing successful fixation to visual objects<sup>38</sup> to recognizing small words and shapes (1.39 logMAR, 2.2 logMAR of visual acuity respectively) following an intensive 5-day training regime, 3 years after implantation<sup>39</sup>. Patients with the Argus II device (Second Sight Medical Products, Inc.) showed improvements in performance over the first weeks of use in a battery of tasks<sup>40</sup> that tended to wane over extended time<sup>41</sup>, although Castaldi and colleagues showed a positive correlation in performance in a detection task vs time since implant<sup>42</sup>. The Argus II results were likely confounded by the training regimen required for implant recipients<sup>43</sup>.

**Combining a passive activity and an active assessment.** Previous work across the field has largely focused on tasks that require concerted effort rather than ones that do not, despite the expectation that passive viewing will dominate real post-implant experience. To address this gap, we investigated VPL while using a simulated visual prosthesis through the passive experience of video viewing. Given the substantial improvements we



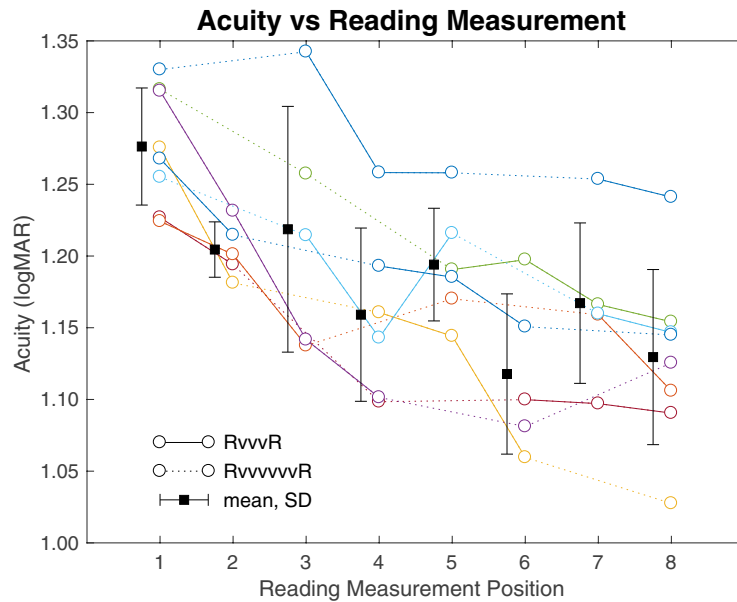
**Figure 1.** Population performance on reading accuracy and speed. The reading accuracy [(a–c), in units of percent correct] and reading speed [(d–f) in units of words per minute] are shown for population means at each of the eight possible measurement positions (colored traces, numbered) in the sequence of sessions. As not all subjects are administered a test at each measurement position in the sequence in order to create the RvvvR/RvvvvvR spans examined elsewhere, the number of subjects varies from curve to curve here (see Fig. 9). Colors shift from earlier (blue) to later (red) measurements, with each trace in the left column (a,d) also identified by measurement position. The left (a,d) and middle (b,e) columns show curves that are the mean over subjects, while the right column (c,f) takes an additional mean over font sizes with dashed lines showing standard deviations across subjects. Phosphene view measurements (open circles) show leftward progression of the reading accuracy curves, but no similar progression in reading speed. Natural view measurements that serve as a control condition (closed circles) show normal values for reading accuracy—all overplotted at 100%—and speed without significant learning effects. Note that the lower right subfigure (f) has two different scales for the two different curves, left for phosphene view (open circles), and right for natural view (closed circles).

previously reported with the active task of reading<sup>1</sup>, and reports from the literature showing the skill of seeing through artificial vision to be highly responsive to training, we hypothesized that the passive task of video viewing could be used to drive a similar effect, but that the learning rate might be slower. We repurposed the reading task for its readout of acuity rather than its training value, and interleaved it sporadically with a video viewing task. Although we minimized the time spent with the reading task by reducing its length and presenting it only infrequently, the tests still represented some experiential time, so we expected effects to remain from them.

To compensate for training originating from both video and reading tasks, we designed a stutter-step schedule in which we systematically varied the placement of a handful of reading tests (R) within the larger set of video-viewing sessions (v), so as to create spans within the overall sequence with either three (R, v, v, R; abbreviated as RvvvR) or six (RvvvvvR) video sessions between reading tests. This schedule, carefully syncopated across subjects, allowed us to measure acuity improvements for runs of consecutive video viewing sessions bounded by single reading tests. We used data from the two lengths of video viewing chains, RvvvR and RvvvvvR, to linearly decode  $\beta_v$  and  $\beta_R$ , the gain factors driving changes in acuity due to video viewing and reading, respectively (see “Methods”).

## Results

Eight subjects (8 total; 2 male, 6 female), recruited from students at the University of Athens, Greece, who had documented and assessed ability in reading English above that required for the tasks, completed the experiment. Subjects had self-reported normal or corrected-to-normal visual acuity, without any major visual defect. A ninth subject was disqualified due to gaze tracking issues, and their data are not reported here. Each subject came to the laboratory for a total of 23 sessions; the first session included a Snellen acuity assessment; subsequent sessions included a reading test through our simulation of artificial vision, and/or watching an episode of a television program shown through the same artificial vision simulation.



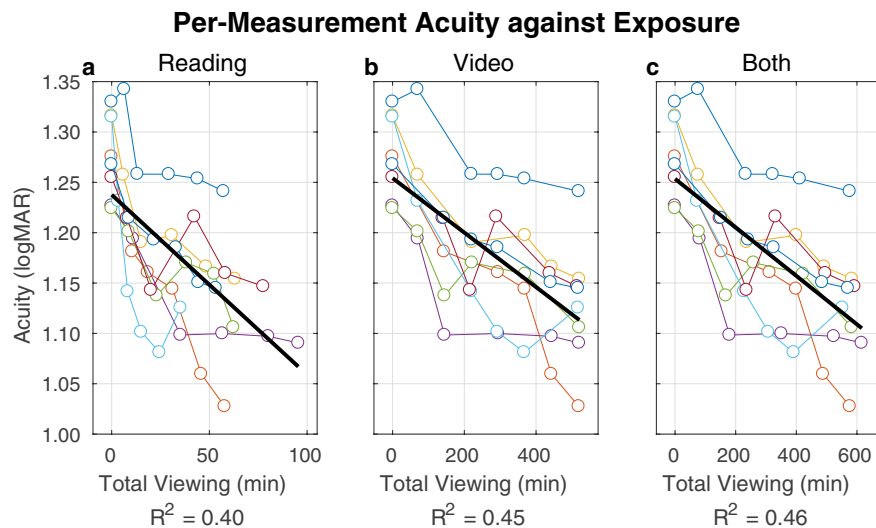
**Figure 2.** Acuity over time. The equivalent acuity is shown for each reading test for each subject. Each colored trace represents one subject. Acuity measurements from reading tests that are separated by three video sessions (RvvvR) are connected with a solid line, while those separated by six video sessions (RvvvvvR) are connected with a dashed line. Each subject displays a progression to improved acuity (lower values on vertical axis) through their participation in the experiment (horizontal axis), with the population as a whole moving strongly downward.

**Reading accuracy.** Before exploring the details of results from the two different spans, we examined overall performance across the full sequence of sessions, here for reading accuracy, and in the next section for reading speed. We found that as a population, subjects improved (Fig. 1) in reading accuracy (percentage of words read correctly) through the sessions but did not improve in reading speed (number of correctly read words per minute). As expected, reading accuracy was near zero at the smallest font sizes and followed a sigmoidal increase toward the larger font sizes. Over the suite of eight reading tests the sigmoidal profile shifted leftward toward smaller fonts as ability increased for the population (Fig. 1a). The mean accuracy pooled across font sizes (Fig. 1c) rose significantly from  $28 \pm 7\%$  to  $53 \pm 9\%$  (Wilcoxon rank sum,  $p = 0.003$ ) echoing the first part of the learning curve from our previous report<sup>1</sup>.

**Reading speed.** Reading speed was also near zero for the smallest font sizes and increased with larger font sizes (Fig. 1d), but did not show evidence of plateauing for the font sizes used here (larger sizes would likely have been necessary), consistent with our previous study. Over the course of the experiment, the curves did not move appreciably for the population. The mean speed pooled across font sizes (Fig. 1f) showed initial hints of improvement over the first three measurements, but did not change significantly following these first measurements (Wilcoxon rank sum,  $p = 0.72$ ,  $6.4 \pm 3.1$  WPM at start,  $6.8 \pm 1.8$  WPM at end).

**Acuity over time.** Equivalent acuity was then extracted from logistic curves fitted to the reading accuracy measurements from each subject and examined for longitudinal effects (Fig. 2). For each subject, there was a striking change over the eight measurements in a manner equivalent to a statistically significant acuity improvement of  $-0.15 \pm 0.05$  logMAR ( $t$  test of the paired differences,  $p = 0.0001$ ) from a starting value of  $1.28 \pm 0.04$  logMAR to a finishing value of  $1.13 \pm 0.06$  logMAR. The level of variability in acuity assessments was elevated as expected due to having limited the number of repeated measurements in each reading task in order to minimize the amount of time subjects spent in that activity. Despite the resulting uncertainty in acuity values, there remained a clear effect of improvement in acuity for each subject through the duration of the experiment.

**Relative contributions of R and v sessions to acuity.** We proceeded to analyze the two factors of reading time and video time in driving acuity improvement. An initial assessment showed both factors were predictive of visual acuity ( $R^2 = 0.39$  for reading,  $R^2 = 0.44$  for video,  $R^2 = 0.45$  for reading and video together;  $p < 10^{-6}$  for all three conditions; see Fig. 3). We then applied a synchronous decoding technique (see “Methods”), sorting the incremental improvements in acuity into short (RvvvR) and long (RvvvvvR) spans between measurements (Fig. 2) to collect the associated factors  $y_1$  ( $-0.025 \pm 0.038$ ;  $t$  test for mean being non-zero,  $p = 0.004$ )



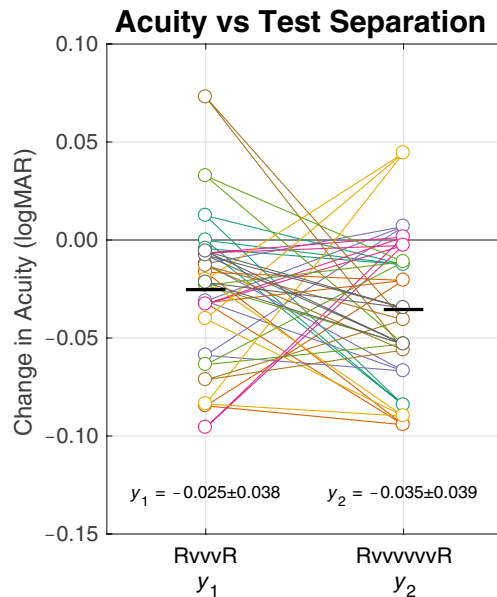
**Figure 3.** Acuity vs phosphene exposure. Acuity (vertical axes) is shown at each of the six measurements made for each subject (colored lines) is plotted against either the total time spent in the reading task [(a), left plot], watching videos [(b), middle], or the combination [(c), right]. For each plot, a linear regression of the data is shown (heavy black line) and the  $R^2$  coefficient given below the axis. The best fit is delivered when acuity is considered against total viewing time (c, reading and video), with nearly as good a fit with video time alone (b), and a somewhat worse fit with reading time alone (a), suggesting that learning was driven by both reading and video experiences.

and  $y_2$  ( $-0.035 \pm 0.039$ ,  $p=0.002$ ) allowed us to decode the per-session learning rate for the active reading task  $\beta_R = -0.015 \pm 0.084$  logMAR/session, and the passive video task  $\beta_V = -0.003 \pm 0.018$  logMAR/session (Fig. 4). These values were both significantly non-zero ( $t$  test for  $\beta_R$ ,  $p=0.001$ ; for  $\beta_V$ ,  $p=0.001$ ) and were statistically distinct (paired  $t$  test,  $p=0.02$ ). To validate this method, we also performed a linear regression of acuity versus cumulative time spent reading and viewing video at each acuity measurement, which yielded learning rates of  $-0.0006$  logMAR/min ( $+0.0005$ ,  $-0.0017$ , 95% CI) for reading and  $-0.0002$  logMAR/min ( $+0.0000$ ,  $-0.0004$ , 95% CI) for video viewing. When multiplied by the mean session lengths ( $t_{\text{reading}} = 13.6$  min,  $t_{\text{video}} = 24.7$  min), this second set of values became  $\beta_V' = -0.009$  logMAR/session and  $\beta_V'' = -0.005$  logMAR/session respectively, agreeing reasonably well with our primary finding for  $\beta_R$  and  $\beta_V$ . We therefore draw two conclusions. First, a single reading session drove about twice the acuity improvement of a single video session despite being about half as long. And second, considering the total number of sessions spent in each task, the overall acuity gain was 56% from reading (mean improvement,  $-0.08$  logMAR) and 44% from video watching (mean improvement,  $-0.06$  logMAR).

To ensure the initial transient that can be seen in Fig. 2 was not unduly influencing our results, we repeated the primary synchronous decoding analysis, excluding the first through fourth measurements. While the uncertainty increased and statistical power was lost as there were fewer data points, the results were highly consistent with  $\beta_R$  being substantially larger than  $\beta_V$ .

**Video viewing behavior.** Subjects were not provided guidance on what to do when watching the videos, and some were truly passive, hardly moving their gaze location from the center of the screen, whereas others appeared to be more intently following the on-screen action. We quantified this range of behavior by measuring the total scan path length of gaze location per session and found a trending increase from  $279 \pm 76$  to  $338 \pm 66$  screen widths per video (Wilcoxon rank sum,  $p=0.1$ ) over the 21 video sessions. A similar, although more significant, increase was found for the reading task with the scan path increasing from  $8.4 \pm 2.5$  to  $14 \pm 5$  screen widths per trial (Wilcoxon rank sum,  $p=0.003$ ) over the reading sessions (see “Discussion”). The correlation coefficient between the scan path length at each reading session and the subsequent video session, once scan paths were corrected for the per-subject mean, was  $r(46) = 0.42$ ,  $p=0.02$ , suggesting the two conditions have a link to an underlying common factor such as session-by-session motivation levels.

**Experimental design validation and control condition.** The experiment was intentionally designed to minimize the fraction of time during the reading task, so as to limit the known learning effects from that task against the intended measurement of learning effects during the video task. We were largely successful with the design, as the subjects spent  $13 \pm 3\%$  of total exposure to the phosphene view simulation in reading



**Figure 4.** Determining  $\beta_R$  and  $\beta_V$  from  $y_1$  and  $y_2$ . Change in reading acuity is shown (open circles), segregated for reading tests separated by three video viewing sessions (RvvvR) and six video viewing sessions (RvvvvvR) in the syncoated schedule. Pooled measurements for improvements  $y_1$  and  $y_2$  refer to the change in reading acuity between three video viewing sessions (RvvvR) and six video viewing sessions (RvvvvvR), respectively. Improvements  $y_1$  and  $y_2$  are assumed to contain the influence from one reading session test session  $R$  ( $\beta_R$ ) and the three or six video sessions  $v$  ( $\beta_V$ ) in each span [see Eqs. (1) and (2)]. Links are drawn (colored lines) between pairs of  $y_1$  and  $y_2$  values from the same subject. The mean change in acuity (black bars) for both  $y_1$  and  $y_2$  is negative, indicating an improvement in reading ability, but as both RvvvR and RvvvvvR spans have the same number of  $R$  sessions, the larger number of encompassed  $v$  sessions represented by the value  $y_2$  creates a larger improvement as compared to  $y_1$ . That difference allows us to deduce the influence of each session type independently, calculating values for  $\beta_V$  and  $\beta_R$  through linear decomposition in Eqs. (3) and (4) (see “Decoding acuity gain factors for reading and video sessions”).

( $4483 \pm 1223$  s), as compared to video viewing ( $30,389 \pm 1050$  s). The ratio of the mean exposures was therefore 1:6.8, or 1 min of phosphene reading for about every 7 min of video viewing.

Reading in the Natural View condition, with text shown normally in unmodified form on the screen, was our primary control (Fig. 1, see also “Methods”). Reading accuracy at all font sizes was 100% for all subjects in this condition, as expected. Mean reading speed pooled across fonts for the population was not significantly different (Wilcoxon rank sum,  $p = 0.51$ ) between the first ( $138 \pm 15$  WPM) and last ( $132 \pm 14$  WPM) measurements, and qualitatively did not appear to be affected through the experiment, matching behavior observed in our previous report with a different cohort of subjects drawn from a similar pool<sup>1</sup>.

## Discussion

We found that passive viewing of videos through our simulation of artificial vision was not nearly as effective for improving visual acuity as reading for the same amount of time. Although significant increases in population reading accuracy were observed, significant improvement for reading speed was not apparent. Our initial hypothesis that passive viewing would be a useful tool in rehabilitation was found to be partially correct: while there was a specific increase found in reading skill, namely reading accuracy, it was not reflected in a universal improvement in proficiency with phosphene vision, as an equivalent increase was not found in reading speed. The most compelling interpretation for this dichotomy is a combination of two related aspects, first that our assumption that passive experience would result in an overall sharpening of ability was incorrect, and second that the sets of skills exercised by the reading task only partially overlapped with those exercised by the video task.

Previous work in the laboratory utilizing multiple phosphene patterns has established that effects of training transfer from one pattern to another in an active, letter recognition task<sup>26</sup>. Thus, we might reasonably expect transferability of skill to be present here as well. Despite this expectation, we did not find evidence of transfer from the video viewing task to all aspects of reading skill.

**Comparison of active task here to previous work.** The cumulative time spent reading here is  $74.7 \pm 20.4$  min ( $n = 8$  subjects), equivalent to the cumulative time in the first four sessions out of the 40 total in our previous experiment<sup>1</sup> at  $77.6 \pm 12.6$  min ( $n = 6$  subjects). Here, there was a significant improvement of  $-0.08 \pm 0.02$  logMAR attributable to the reading task, versus the slightly higher  $-0.10 \pm 0.08$  logMAR for the

equivalent time in the previous study. While the two sets of improvements were not significantly different (Wilcoxon rank sum,  $p=0.49$ ), the possibly higher underlying rate in the previous study may stem from the reinforcing effects of frequent training that were not as strong here due to intentional gaps between reading sessions to reduce that task's influence.

**Dilution of reinforcement.** The larger gaps between reading sessions here, and the weaker influence of video sessions on learning may have diluted reinforcement effects and impeded reading speed development. Reinforcement effects that varied with time between sessions were seen in our earlier work<sup>1</sup>, where reading accuracy was found to be more robust to gaps in training than reading speed. Whereas short gaps of a few days tended to not affect improvements made in reading accuracy, they worked against or eliminated improvements in reading speed such that only a strong gain in a day's session overcame a low-level of persistent forgetting. With that observation in mind, one explanation for the lack of improvement in reading speed here is that the gain during video viewing, as indicated by the per-session  $\beta_v$ , might not have been large enough to sustain improvements across sessions. Another explanation, perhaps not mutually exclusive, is that the act of reading combines two aspects of visual perception, one of raw acuity used to recognize visual patterns and extract probabilistic information on underlying object shape, and a second of word recognition from that probabilistic shape description; if we assume that the first is trained both under reading as well as video experience, but the second aspect is trained only during reading, then it would be reasonable to see a disparity in reading accuracy versus reading speed as observed here.

**Scan path lengths.** Increases in scan path length in the reading task as performance improves are slightly counter-intuitive. We explored this observation through a re-analysis of the more extensive data from our previous report<sup>1</sup>, and found that when subjects are doing very poorly, they generally have short scan lengths in the reading test because they quickly abandon sentences under difficult reading conditions. As their skill improves, they abandon less and spend more time searching and exploring, increasing the scan length, which peaks when the mean performance is 50%. As performance continues to improve, the scan length decreases again as the amount of searching drops and the gaze behavior more closely resembles the left-to-right, line-by-line pattern from normal reading. Thus, we believe the general increase in scan path length seen here reflects a still-continuing improvement in skill like the first part of the more fully completed examples in our previous research. Scan path lengths would therefore be expected to decline again here with additional training. We might imagine that the similar, but less pronounced scan path change observed in the video task would follow a similar up-and-down profile; without prior evidence to support this speculation, however, certainty can come only with additional investigation.

**Transferability of visual skill.** In the design of our experiment, we have assumed that there would be a transferability for skills learned during the passive task (specifically, the ability to recognize objects that are viewed through the phosphene pattern), to the active task where our quantitative measurements are made, and thus any specific improvement driven by video viewing would be reflected in a generalized improvement in all aspects of the reading assessment. This idea of transferability, or generalization from one set of experiences to another, has been the subject of substantial investigation in the visual system, often under the umbrella of VPL.

Transfer of visual skills has been observed in such diverse conditions as with subjects impaired by amblyopia<sup>44–48</sup>, central vision loss<sup>49,50</sup>, presbyopia<sup>51,52</sup>, low vision in children<sup>53</sup>, macular degeneration<sup>54</sup>, cortical blindness from stroke<sup>55,56</sup>, visual function loss<sup>57,58</sup>, and impaired vision<sup>59,60</sup>. And transfer of perceptual learning has been demonstrated in situations such as in shaping a preferred retinal locus in the visual periphery<sup>50</sup>, moving regions of sensitivity in areas of the visual cortex<sup>54</sup>, improving letter recognition<sup>46,61,62</sup>, seeing biological motion in noise<sup>63</sup>, and eye-hand coordination in sports<sup>64,65</sup>. Such changes are seen especially when training is based on a broad stimulus set<sup>65–67</sup>, retinal locations<sup>56</sup> or uses video games<sup>67–73</sup>. Training to respond to these stimuli has been reported to improve visual acuity and contrast sensitivity for both central and peripheral vision of subjects<sup>67</sup>, actual field performance of baseball players<sup>65</sup>, as well as reading<sup>73</sup>. The wide palette of modalities where skill transfer has been found suggests that transfer between the two tasks used in this study should be possible.

Observations have been made, however, about limitations or specificity of transfer. For example, McGovern and colleagues observed improvements that transferred between the related visual tasks of orientation, curvature, and global form discrimination after 10 sessions of 400 trials<sup>74</sup>. In particular, both orientation and global form discrimination transferred to the other two tasks, however, the curvature task transferred only to the orientation task. This asymmetry of action demonstrates that visual skill development can be highly specific.

**Transferability in studies with simulations of artificial vision.** Narrowing the range of studies to those that are closest to the present report, two groups have studied the transferability of trained skills in simulations of artificial vision. The first group, Sommerhalder and colleagues used a reading task to investigate the generalization of monocularly presented stimuli to stimuli presented to the fellow eye<sup>28,29</sup>. They simulated an eccentric (non-foveal) placement of a visual prosthesis and had normally sighted subjects perform reading tasks with that simulation. Subjects had a series of 1-h sessions reading 4-letter words for a period of 1 month and subsequently 1-h sessions of reading full-page text for a period of 2 months. VPL was observed to be successfully transferred from the trained to the untrained eye.

The second group, Wang and colleagues investigated the transfer between two related tasks, object-to-name labeling and name-to-object identification in a cross-validated experiment using two subject cohorts<sup>75</sup>. In their labeling task, subjects were presented a single object through a simulation of artificial vision and had to choose the correct label out of a set of possible words shown in the clear. In the identification task, subjects were

presented instead a single label in the clear and had to choose the correct object out of a set of possible objects, shown in a simulation of artificial vision. In a 1-day session of 64 trials, visual recognition significantly transferred from the labeling task to the identification task with substantial effect in the first cohort but not vice versa in the reverse cohort. In a follow-up study, the group investigated persistence of learning with their labeling task as well as the effect of viewing images compared to videos<sup>76</sup>. Training with the labeling task led to an improved performance, with the improvement with videos being more pronounced than with still images. Next, using their identification task, they found that performance with videos exceeded that with still images in all blocks, but not with as robust a difference. These asymmetries between the nature of both the task and the stimuli confirm the potential for specificity in visual skill development.

**Stutter-step design and recovery of gain factors for R and v.** We selected a stutter-step design using two different length runs of video viewing sessions bounded by reading sessions in order to be able to dissect out the relative influences of the reading and video viewing sessions. An alternative design would have been to, for example, employ separate control and experimental groups both of whom would be administered regularly spaced reading assessments, but only the experimental group would have seen videos in between reading assessments (see our previous report describing an experiment with a similar condition<sup>1</sup>, and the comparison here in section “[Comparison of active task here to previous work](#)”). A third alternative would be to employ a crossover design where videos would be seen by one half of the subjects in the first half of the schedule, and by the other group in the second half. Neither of these approaches, or of the many others that were considered, would have completely compensated for longitudinal effects while also allowing subjects to be their own controls and, importantly, allowing for the rigorous separation of the two factors  $\beta_R$  and  $\beta_v$ . The stutter-step strategy selected supports analysis through the method of synchronous decoding developed for signal detection theory, a powerful means to extract signal in the face of noise, allows each subject to be their own control, and maximizes the statistical power available from a limited number of subjects.

**Video watching as a passive task.** We selected watching videos as a passive task because we wanted to use an activity that recipients of visual prostheses might experience without the need for a special-purpose rehabilitative apparatus. We reasoned that video watching might be effortlessly and pleasantly incorporated into post-implant daily routines with little prodding required from rehabilitation specialists. While the visual aspect of watching videos was expected to be the primary driver of the effects seen here, we included the audio track of each episode in our simulations to more accurately emulate the experience of a prosthesis recipient sitting in front of a television. We did not, however, explore whether the presence of audio itself was a facilitating factor.

## Conclusions

Although the idea of being able to improve the utility of a visual prosthesis through an enjoyable pastime like watching television programs is highly attractive, in our hands, this process does not have nearly the impact as an active, at times challenging task which provides clear-cut automatic feedback for correct answers. While there appear to be benefits from passively viewing videos, the gains are much slower per unit time, and do not appear to transfer universally to other visual skills. During post-implant rehabilitation, therefore, having patients engage in an active process appears to be preferable to ensure successful treatment with a visual prosthesis device. Further research will help identify self-administered active and passive tasks that overcome specificity to allow broad improvements in prosthetic vision, while being as enjoyable as possible to encourage patient engagement in post-implant rehabilitation.

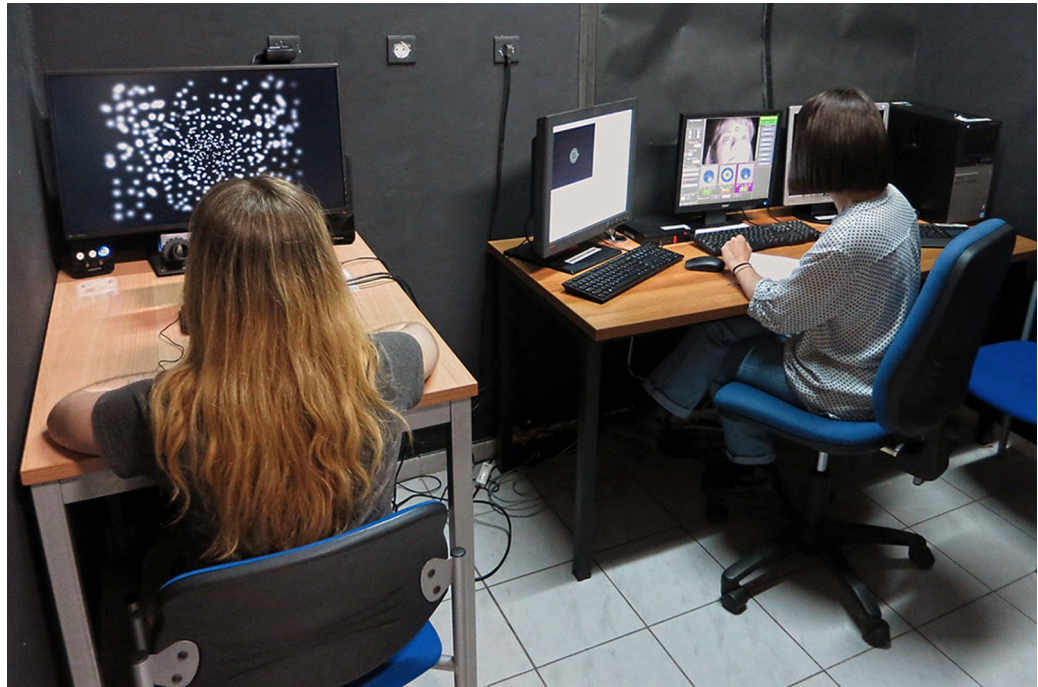
## Methods

The methods employed in this report are very similar to those previously described<sup>1</sup> with differences as highlighted below. Briefly, we used a simulation (Fig. 5) that included 1000 phosphenes in a center-weighted pattern<sup>36</sup> that were activated in a gaze-contingent fashion<sup>77</sup>, tracking instantaneously measured eye position on a frame-by-frame basis so as to approximately stabilize their position in retinal coordinates. The phosphenes were used both as a filter to measure the average image luminance at each phosphene’s location, and then as a display, illuminated as a 2D-Gaussian at the matching luminance. This method is often called *veridical encoding*. The simulation was used to provide two different tasks to the subjects, a reading task that was used as metric of visual performance (Figs. 6, 7), and a video task that was the focus of this work (Fig. 8); the first task requires active participation from the subject, and was designed to be a small fraction of the total experience, while the second task has no specific burden of action from the subject, and was designed to be the dominant portion of the total experience. Each of the two tasks are described below in more detail.

**Differences: extended viewing and visual load.** Reading tests here and in our previous work involved brief periods of phosphene viewing (approximately 30 s on average per trial) separated by inter-trial intervals, interleaved with natural viewing and presented through multiple trials per test, whereas video viewing was for a much more extended, continuous time, typically 24 min, that did not include interruptions. The level of subject engagement required was very different between the reading tests and video viewing, as the act of reading demands active focus to understand each word, whereas watching light entertainment has no goal-directed need to parse visual information.

**Differences: more complicated schedule.** Although subjects shared the same number of video-viewing sessions, each subject’s six pseudo-periodic measurements of their reading performance took place distributed





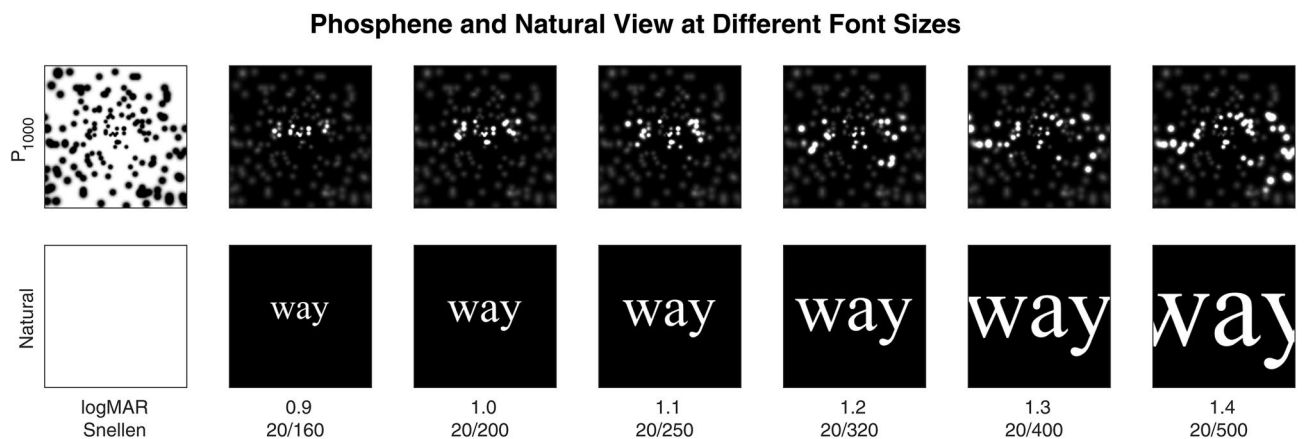
**Figure 5.** Apparatus. The experimental setup is shown in this photograph with a subject seated on the left in front of the stimulus monitor, and the gaze-tracking camera just visible to the left of their head. The subject's position with arms on the table helps reduce head movement. The stimulus monitor displays a phosphene-view image of one frame from a video. On the right, the experimenter is sitting in front of the behavioral control system and gaze tracker. This image has been intentionally over-exposed to help reveal detail, as the room is normally dim.

across eight different possible positions, in a per-subject pattern that was determined beforehand. This pattern resulted in gaps between reading tests of either three (RvvvR) or six (RvvvvvR) videos, with each subject having three of the shorter gaps and two of the longer gaps.

**Differences: lower system latency, higher monitor refresh rate.** In contrast to the previous experiment for which the eye-tracking system was operating at 500 Hz, the monitor's refresh rate was 60 Hz, and the overall system-latency was expected to be 35 ms total, here, we operated the eye-tracking system (SR Research, EyeLink 1000+) at 1000 Hz, used a monitor that could support 144 Hz refresh rate (Asus ROG PG279Q), optimized the custom simulation code, and estimate the latency from eye position measurement to display update to be 11 ms (1 ms gaze measurement, plus 7 ms frame update, plus 3 ms monitor lag).

**Reading task and stimuli.** Reading tests used the MNREAD corpus of simple, three-line sentences presented in a range of font sizes (logMAR 0.9 to 1.4 in 0.1 increments) to assess visual acuity<sup>34,78</sup>, modified so that subjects viewed the sentences shown on a computer monitor either through natural vision as a control condition, or through a simulation of a thalamic visual prosthesis<sup>1,31,35</sup> with 1000 phosphenes that spanned the visual field<sup>36</sup>. The reading test method has been described previously in detail<sup>1,25,26,31,35</sup>, and is summarized here. In a series of trials, subjects are shown novel sentences from the corpus at different font sizes in a pre-determined, pseudo-random sequence. Each font size and viewing condition is presented twice in a given reading test. The number of correctly read words is scored by the experimenter, and the time taken to read each sentence is recorded automatically. From these values, profiles of reading accuracy (percent of correctly read words) and reading speed (number of correctly read words per minute) are developed across the calibrated font sizes. Reading accuracy versus font size can be fitted with a sigmoidal curve where the 50% level is considered the visual acuity of the subject for the given viewing condition. All font sizes presented were substantially larger than native acuity of the subjects, and thus reading accuracy was 100% for the Natural condition that used text presented without filtering. An example of the range of font sizes against the phosphene pattern is given in Fig. 6, while an example sentence seen under phosphene view with different gaze positions is given in Fig. 7.

**Video stimuli.** During video sessions, subjects were presented full episodes of the classic American comedy, "I Love Lucy," viewed through the visual prosthesis simulation. We used these shows because of the ready availability of the material on recorded media, the light and entertaining subject matter, and the expectation that

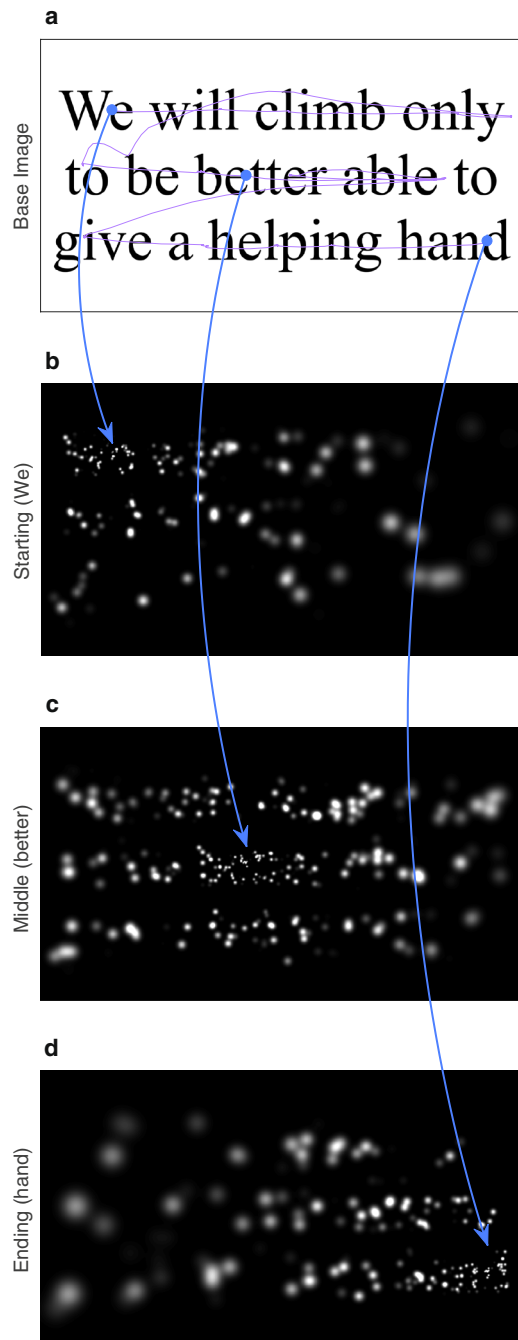


**Figure 6.** Reading test stimuli. This figure shows the twelve possible stimulus conditions for the example word *way*. The left column depicts the central 10° of viewing although it is presented here in negative colours. Stimuli were always presented to the subjects with white on black. In the right columns, gaze location is at the center of the word so that the highest density of phosphene layout overlays the text. Viewing conditions were either presented in Natural view for control (bottom row) or  $P_{1000}$ , a simulation of a thalamic visual prosthesis with 1000 phosphenes (top row). Font size ranged from logMAR 0.9 to 1.4, corresponding to a Snellen acuity from 20/160 to 20/500 (left to right). The full three-line simple sentences occupied the entire video screen at the largest font size (see Fig. 7), although only the central-most patch is shown here. The  $P_{1000}$  phosphene map spans the entire visual field, although only the central portion is shown here, with some 250 phosphenes.

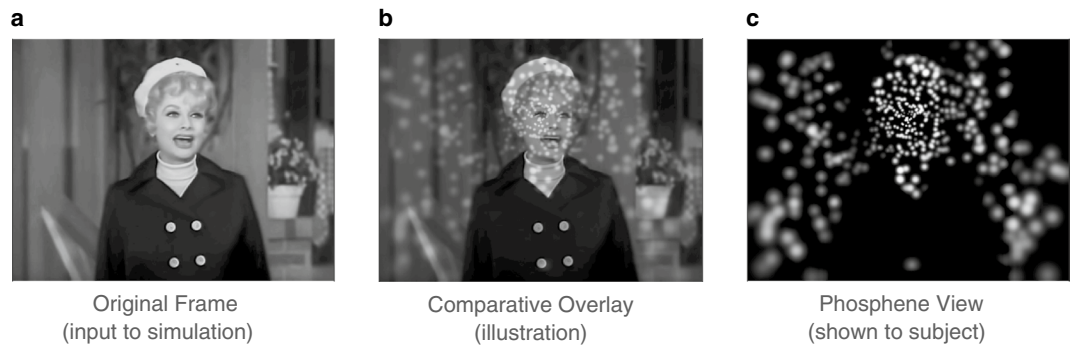
while our subjects would likely be generally familiar with the series, none of them would have recently, if ever, seen the show. We compiled various publicly available best-of lists to identify the most engaging and popular episodes and selected a suite of 21 that were presented in order. Episodes were shown through our simulation of artificial vision using a real-time gaze-contingent architecture (see below) and were presented with normal audio that was synchronized to the simulation. Videos were typically 24–27 min long, although subjects often broke off their viewing at the closing credits. A representative frame in its original state and as viewed through the simulation is given in Fig. 8.

**Gaze-contingent architecture.** Visual stimuli were presented on a contemporary LCD monitor (ROG PG279Q by Asus, Inc.) set to 1600 by 900 pixels resolution and a frame rate of 144 Hz (6.9 ms between frame updates). This monitor was selected because of its high-quality IPS panel, fast refresh rate, and extremely low latency (3.25 ms measured by TFT Central, [https://www.tftcentral.co.uk/reviews/asus\\_rog\\_swift\\_pg279q.htm](https://www.tftcentral.co.uk/reviews/asus_rog_swift_pg279q.htm)). Video signals sent to the monitor were computed in real-time based on subject gaze position measured through a head-free tracker (EyeLink 1000+, by SR Research, Inc.) running at 1000 Hz. For each monitor refresh, the center of the phosphene pattern was computationally translated to the most recently read gaze location and each phosphene used as a local-averaging filter at its position on the current frame from the television episode to construct a phosphene-view frame displayed on the subject monitor. The ratio between monitor refresh (144 Hz) and episode frame rate (23.96 Hz) of 6.01 meant that each video frame was presented 6 times to the subject, with one frame shown 7 times approximately every 4 seconds to resynchronize; when videos were viewed in the clear during development, these periodic resynchronizations were imperceptible. With this monitor and gaze tracker, we estimate the system latency to have been approximately 10 ms. Additional details of the gaze tracking and filtering process can be found in our previous publications<sup>1,25,26,31,35</sup>.

**Syncopated schedule.** A pre-determined testing schedule was used, developed on the basis of eight subjects with 23 visits each to the laboratory. At the first session, each subject was given an informal Snellen chart evaluation to verify that they had normal or corrected-to-normal vision. At subsequent sessions they were either given a reading test, or watched an episode of the television series, or both, as determined by the testing schedule. Positions in the schedule were created for reading tests every three sessions of video viewing, with a final test on the last day, alone, for a total of eight potential testing days. We utilized an interleaved syncopation when determining when subjects would be administered reading tests within those eight possible positions such that each subject skipped two tests in the 23 sessions (see Fig. 9). This schedule resulted in gaps between reading tests of either three (RvvvR) or six (RvvvvvR) video sessions, with each subject having two tests separated by six videos, and three tests separated by three videos. Across the population, the first and last test positions were fully populated and thus produced eight measurements, whereas the second through seventh test positions were partially populated so produced either five or six measurements.



**Figure 7.** Reading screenshots. These panels depict three points during the reading of a typical sentence, “We will climb only to be better able to give a helping hand,” at the largest font size. The subject reads each of the three lines in turn, scanning along the text. (a) Gaze position over time (purple trace) is overlaid on the base image used to compute the phosphene view animation presented to the subject. The base image is never itself visible during phosphene view conditions. Three example points at the start, middle, and end of the scanning are highlighted and connected to their equivalent locations on the lower panels that are snapshots from the trial’s gaze-contingent presentation. (b) The phosphene pattern shown to the subject at the start of this trial when the subject is gazing at the start of the first line, and is about to read the word, *We*. (c) The subject has scanned through the first line and is mid-way through the second line during this second screen capture. As the subject’s gaze position shifts, the phosphene pattern shifts in rigid concert, and we now see the subject looking at the word, *better*. (d) The subject has nearly finished the trial, scanning through all three lines, and is about to read the last word, *hand*. The number of phosphenes landing on the monitor area that are potentially activated are 500, 640, and 430 for subplots, (b), (c) and (d), respectively. Of these phosphene locations, 260, 400, and 250 have been activated, respectively, to represent the image of the text.



**Figure 8.** Video stimuli. **(a)** A representative stimulus frame. Video frames are used as input to the simulation along with the current gaze position that determines the location of the phosphene field as it is overlaid on the image. **(b)** An example overlay of the phosphene pattern for the gaze position at the center of the frame. Each phosphene will represent the mean brightness of the image behind it, thus the brightest points are at the woman's white hat and face, and the darkest are over her coat and portions of the background. The phosphene pattern is updated in position based on gaze location and brightness with each video frame. Phosphenes are smaller and denser toward the point of regard, here near the center of the face, and become larger and more separated in the periphery. **(c)** The resulting constructed phosphene image as shown to the subject. The head and neck of the actress can be made out against larger rectangular fields that correspond to portions of the background. With some effort, the eyes and mouth are discernible—while the example screen shot suggests it might be quite difficult to perceive fine detail, with an animated presentation the experience is surprisingly improved as motion information becomes available to the viewer.

**Decoding acuity gain factors for reading and video sessions.** For the short spans between reading tests/acuity measurements,  $R_{vvvR}$ , we assumed there would be a constant factor of improvement from  $R$  (the second  $R$  measures the influence of the first  $R$ , but is assumed to carry no influence on its own measurement) and a different, constant factor from each  $v$ , thus generating an overall improvement  $y_1$ :

$$y_1 = 3\beta_V + \beta_R \quad (1)$$

For the long spans between measurements,  $R_{vvvvvR}$ , we assumed these same factors would combine with twice as many instances of  $v$ , to yield an overall improvement  $y_2$ :

$$y_2 = 6\beta_V + \beta_R \quad (2)$$

These expressions for the measured improvements  $y_1$  and  $y_2$  create a system of two linear equations with two unknowns, allowing us to decode the influence from reading  $\beta_R$ , and video watching  $\beta_V$ :

$$\begin{aligned} y_1 - y_2 &= 3\beta_V + \beta_R - (6\beta_V + \beta_R) \\ &= 3\beta_V + \beta_R - 6\beta_V - \beta_R \\ &= -3\beta_V \\ \beta_V &= (y_2 - y_1)/3 \end{aligned} \quad (3)$$

$$\begin{aligned} 2y_1 - y_2 &= 2(3\beta_V + \beta_R) - (6\beta_V + \beta_R) \\ &= 6\beta_V + 2\beta_R - 6\beta_V - \beta_R \\ &= \beta_R \\ \beta_R &= 2y_1 - y_2 \end{aligned} \quad (4)$$

Equations (3) and (4) were thus used to convert  $y_1$  and  $y_2$  measurements of acuity improvements from one reading test to the next (Fig. 4) into separated values for  $\beta_V$  and  $\beta_R$ .

We note that accepting this model for extracting  $\beta_V$  and  $\beta_R$  from  $y_1$  and  $y_2$  requires understanding that there is no a priori expectation for statistical independence of  $y_1$  and  $y_2$ . For example, in the hypothetical case where video viewing has no influence whatsoever, the value of  $\beta_V$  would be 0 and the model formed by Eqs. (1) and (2) implies that the observations of  $y_1$  and  $y_2$  would be statistically indistinguishable.

**Ethics statement.** The protocol used in this experiment was approved by the Institute Review Board of the Massachusetts General Hospital and the Ethics Board of the Cognitive Sciences Department in the Department of History and Philosophy of Science at the National and Kapodistrian University of Athens. It conformed to the Declaration of Helsinki. The protocol was classified as a minimal risk study with informed consent obtained from each subject. Subjects were provided modest monetary compensation for their participation.

Session	Subject Number																								Meas Pos
	1			2			3			4			5			6			7			8			
	S	R	V	S	R	V	S	R	V	S	R	V	S	R	V	S	R	V	S	R	V	S	R	V	
1	■			■			■			■			■			■			■			■			
2		■	■		■	■		■	■		■	■		■	■		■	■		■	■		■	■	1
3																									
4																									
5				■									■			■							■		2
6																									
7																									
8		■											■			■					■				3
9																									
10																									
11		■			■								■			■					■			■	4
12																									
13																									
14		■			■								■			■					■			■	5
15																									
16																									
17				■				■				■				■					■			■	6
18																									
19																									
20		■											■			■					■			■	7
21																									
22																									
23		■			■				■				■			■					■			■	8

**Figure 9.** Testing schedule. The testing schedule was used to coordinate the reading tests that were performed every three (RvvvR) or six (RvvvvvR) video sessions in syncoordinated fashion across the eight subjects, as shown in this table. Sessions are shown across rows and subjects in groups of three columns, one for each of Snellen acuity test (yellow squares), reading test (blue squares), and video viewing (red squares). Eight potential positions appear for the reading tests starting with Session 2 and ending with Session 23, as shown in the rightmost column. The majority of days for each subject consisted of just video viewing. Each day that there was both a reading test and a video viewing, the reading test was administered first. The three or six video viewings between reading tests were used to drive the decoupling of learning effects from the two tasks.

Received: 17 March 2021; Accepted: 12 April 2022

Published online: 28 July 2022

## References

- Rassia, K. E. K. & Pezaris, J. S. Improvement in reading performance through training with simulated thalamic visual prostheses. *Sci. Rep.* **8**, 16310. <https://doi.org/10.1038/s41598-018-31435-0> (2018).
- Sasaki, Y., Nanez, J. E. & Watanabe, T. Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* **11**, 53–60. <https://doi.org/10.1038/nrn2737> (2010).
- Skrandies, W. & Fahle, M. Neurophysiological correlates of perceptual learning in the human brain. *Brain Topogr.* **7**, 163–168. <https://doi.org/10.1007/BF01186774> (1994).
- Clapp, W. C. *et al.* Effects of long-term potentiation in the human visual cortex: A functional magnetic resonance imaging study. *NeuroReport* **16**, 1977–1980. <https://doi.org/10.1097/00001756-200512190-00001> (2005).
- Teyler, T. J. *et al.* Long-term potentiation of human visual evoked responses. *Eur. J. Neurosci.* **21**, 2045–2050. <https://doi.org/10.1111/j.1460-9568.2005.04007.x> (2005).
- Gutnisky, D. A., Hansen, B. J., Iliescu, B. F. & Dragoi, V. Attention alters visual plasticity during exposure-based learning. *Curr. Biol.* **19**, 555–560. <https://doi.org/10.1016/j.cub.2009.01.063> (2009).
- Beste, C., Wascher, E., Güntürkün, O. & Dinse, H. R. Improvement and impairment of visually guided behavior through LTP- and LTD-like exposure-based visual learning. *Curr. Biol.* **21**, 876–882. <https://doi.org/10.1016/j.cub.2011.03.065> (2011).
- Clapp, W. C., Hamm, J. P., Kirk, I. J. & Teyler, T. J. Translating long-term potentiation from animals to humans: A novel method for noninvasive assessment of cortical plasticity. *Biol. Psychiatry* **71**, 496–502. <https://doi.org/10.1016/j.biopsych.2011.08.021> (2012).
- James, K. H. & Atwood, T. P. The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. *Cogn. Neuropsychol.* **26**, 91–110. <https://doi.org/10.1080/02643290802425914> (2009).
- Herzog, M. H. & Fahle, M. The role of feedback in learning a Vernier discrimination task. *Vis. Res.* **37**, 2133–2141. [https://doi.org/10.1016/S0042-6989\(97\)00043-6](https://doi.org/10.1016/S0042-6989(97)00043-6) (1997).
- Seitz, A. R. & Watanabe, T. Is subliminal learning really passive?: Psychophysics. *Nature* **422**, 36–36. <https://doi.org/10.1038/422036a> (2003).
- Cortese, A., Lau, H. & Kawato, M. Unconscious reinforcement learning of hidden brain states supported by confidence. *Nat. Commun.* **11**, 4429. <https://doi.org/10.1038/s41467-020-17828-8> (2020).
- Christou, C. G. & Bühlhoff, H. H. View dependence in scene recognition after active learning. *Mem. Cogn.* **27**, 996–1007. <https://doi.org/10.3758/BF03201230> (1999).

14. James, K. H., Humphrey, G. K., Vilis, T. & Corrie, B. “Active” and “passive” learning of three-dimensional object structure within an immersive virtual reality environment. *Behav. Res. Methods Instrum. Comput.* **34**, 383–390. <https://doi.org/10.3758/BF031954668> (2002).
15. Meijer, F. & Van der Lubbe, R. H. J. Active exploration improves perceptual sensitivity for virtual 3D objects in visual recognition tasks. *Vis. Res.* **51**, 2431–2439. <https://doi.org/10.1016/j.visres.2011.09.013> (2011).
16. Chrastil, E. R. & Warren, W. H. Active and passive spatial learning in human navigation: Acquisition of survey knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* **39**, 1520–1537. <https://doi.org/10.1037/a0032382> (2013).
17. Dagnelie, G., Walter, M. & Yang, L. Playing checkers: Detection and eye–hand coordination in simulated prosthetic vision. *J. Mod. Opt.* **53**, 1325–1342. <https://doi.org/10.1080/09500340600619197> (2006).
18. Pérez Fornos, A., Sommerhalder, J., Pittard, A., Safran, A. B. & Pelizzone, M. Simulation of artificial vision: IV. Visual information required to achieve simple pointing and manipulation tasks. *Vis. Res.* **48**, 1705–1718. <https://doi.org/10.1016/j.visres.2008.04.027> (2008).
19. Srivastava, N. R., Troyk, P. R. & Dagnelie, G. Detection, eye–hand coordination and virtual mobility performance in simulated vision for a cortical visual prosthesis device. *J. Neural Eng.* **6**, 035008. <https://doi.org/10.1088/1741-2560/6/3/035008> (2009).
20. Josh, H., Yong, B. & Kleeman, L. Mobile, real-time simulator for a cortical visual prosthesis. In *Proceedings of the International Conference on Biomedical Electronics and Devices* 37–46 (SciTePress—Science and Technology Publications, 2012).
21. Thompson, R. W., Barnett, G. D., Humayun, M. S. & Dagnelie, G. Facial recognition using simulated prosthetic pixelized vision. *Investig. Ophthalmol. Vis. Sci.* **44**, 5035. <https://doi.org/10.1167/iovs.03-0341> (2003).
22. Xia, P., Hu, J. & Peng, Y. Adaptation to phosphene parameters based on multi-object recognition using simulated prosthetic vision: Phosphene parameters adaptation. *Artif. Organs* **39**, 1038–1045. <https://doi.org/10.1111/aor.12504> (2015).
23. Chen, S. C., Hallum, L. E., Lovell, N. H. & Suaning, G. J. Learning prosthetic vision: A virtual-reality study. *IEEE Trans. Neural Syst. Rehabil. Eng.* **13**, 249–255. <https://doi.org/10.1109/TNSRE.2005.851771> (2005).
24. Fu, L., Cai, S., Zhang, H., Hu, G. & Zhang, X. Psychophysics of reading with a limited number of pixels: Towards the rehabilitation of reading ability with visual prosthesis. *Vis. Res.* **46**, 1292–1301. <https://doi.org/10.1016/j.visres.2005.11.011> (2006).
25. Bourkiza, B., Vurro, M., Jeffries, A. & Pezaris, J. S. Visual acuity of simulated thalamic visual prostheses in normally sighted humans. *PLoS One* **8**, e73592. <https://doi.org/10.1371/journal.pone.0073592> (2013).
26. 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**, 36329. <https://doi.org/10.1038/srep36329> (2016).
27. Brindley, G. S. The number of information channels needed for efficient reading. *J. Physiol.* **177**, 44–46 (1964).
28. Sommerhalder, J. *et al.* Simulation of artificial vision: I. Eccentric reading of isolated words, and perceptual learning. *Vis. Res.* **43**, 269–283. [https://doi.org/10.1016/S0042-6989\(02\)00481-9](https://doi.org/10.1016/S0042-6989(02)00481-9) (2003).
29. Sommerhalder, J. *et al.* Simulation of artificial vision: II. Eccentric reading of full-page text and the learning of this task. *Vis. Res.* **44**, 1693–1706. <https://doi.org/10.1016/j.visres.2004.01.017> (2004).
30. Dagnelie, G., Barnett, D., Humayun, M. S. & Thompson, R. W. Paragraph text reading using a pixelized prosthetic vision simulator: Parameter dependence and task learning in free-viewing conditions. *Investig. Ophthalmol. Vis. Sci.* **47**, 1241. <https://doi.org/10.1167/iovs.05-0157> (2006).
31. 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).
32. Hayes, J. S. *et al.* Visually guided performance of simple tasks using simulated prosthetic vision. *Artif. Organs* **27**, 1016–1028. <https://doi.org/10.1046/j.1525-1594.2003.07309.x> (2003).
33. van Rheede, J. J., Kennard, C. & Hicks, S. L. Simulating prosthetic vision: Optimizing the information content of a limited visual display. *J. Vis.* **10**, 32. <https://doi.org/10.1167/10.14.32> (2010).
34. Mansfield, J. S., Legge, G. E., Luebker, A. T. & Cunningham, K. *MNREAD Acuity Charts* (University of Minnesota, 1994).
35. Paraskevoudi, N. & Pezaris, J. S. Full gaze contingency provides better reading performance than head steering alone in a simulation of prosthetic vision. *Sci. Rep.* **11**, 11121. <https://doi.org/10.1038/s41598-021-86996-4> (2021).
36. Pezaris, J. S. & Reid, R. C. Simulations of electrode placement for a thalamic visual prosthesis. *IEEE Trans. Biomed. Eng.* **56**, 172–178. <https://doi.org/10.1109/TBME.2008.2005973> (2009).
37. 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).
38. Zrenner, E. *et al.* Subretinal electronic chips allow blind patients to read letters and combine them to words. *Proc. Biol. Sci.* **278**, 1489–1497 (2011).
39. Kapetanovic, J. *et al.* Highest reported visual acuity after electronic retinal implantation. *Acta Ophthalmol.* **98**, 736–740. <https://doi.org/10.1111/aos.14443> (2020).
40. Humayun, M. S. *et al.* Preliminary 6 month results from the Argus II epiretinal prosthesis feasibility study. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2009**, 4566–4568. <https://doi.org/10.1109/IEMBS.2009.5332695> (2009).
41. da Cruz, L. *et al.* Five-year safety and performance results from the Argus II retinal prosthesis system clinical trial. *Ophthalmology* **123**, 2248–2254. <https://doi.org/10.1016/j.ophtha.2016.06.049> (2016).
42. Castaldi, E. *et al.* Visual BOLD response in late blind subjects with Argus II retinal prosthesis. *PLoS Biol.* **14**, e1002569. <https://doi.org/10.1371/journal.pbio.1002569> (2016).
43. Erickson-Davis, C. & Korzybska, H. What do blind people ‘see’ with retinal prostheses? Observations and qualitative reports of epiretinal implant users. *PLoS One* **16**, e0229189. <https://doi.org/10.1371/journal.pone.0229189> (2021).
44. Levi, D. M. & Polat, U. Neural plasticity in adults with amblyopia. *Proc. Natl. Acad. Sci.* **93**, 6830–6834. <https://doi.org/10.1073/pnas.93.13.6830> (1996).
45. Levi, D. M., Polat, U. & Hu, Y.-S. Improvement in Vernier acuity in adults with amblyopia. *Investig. Ophthalmol.* **38**, 18 (1997).
46. Polat, U. Restoration of underdeveloped cortical functions: Evidence from treatment of adult amblyopia. *Restor. Neurol. Neurosci.* **26**, 413–424 (2008).
47. Levi, D. M. & Li, R. W. Perceptual learning as a potential treatment for amblyopia: A mini-review. *Vis. Res.* **49**, 2535–2549. <https://doi.org/10.1016/j.visres.2009.02.010> (2009).
48. Hussain, Z., Webb, B. S., Astle, A. T. & McGraw, P. V. Perceptual learning reduces crowding in amblyopia and in the normal periphery. *J. Neurosci.* **32**, 474–480. <https://doi.org/10.1523/JNEUROSCI.3845-11.2012> (2012).
49. Chung, S. T. L. Improving reading speed for people with central vision loss through perceptual learning. *Investig. Ophthalmol. Vis. Sci.* **52**, 1164. <https://doi.org/10.1167/iovs.10-6034> (2011).
50. Liu, R. & Kwon, M. Integrating oculomotor and perceptual training to induce a pseudofovea: A model system for studying central vision loss. *J. Vis.* **16**, 10. <https://doi.org/10.1167/16.6.10> (2016).
51. Polat, U. Making perceptual learning practical to improve visual functions. *Vis. Res.* **49**, 2566–2573. <https://doi.org/10.1016/j.visres.2009.06.005> (2009).
52. Polat, U. *et al.* Training the brain to overcome the effect of aging on the human eye. *Sci. Rep.* **2**, 278. <https://doi.org/10.1038/srep00278> (2012).
53. Nyquist, J. B., Lappin, J. S., Zhang, R. & Tadin, D. Perceptual training yields rapid improvements in visually impaired youth. *Sci. Rep.* **6**, 37431. <https://doi.org/10.1038/srep37431> (2016).

54. Baker, C. I. Reorganization of visual processing in macular degeneration. *J. Neurosci.* **25**, 614–618. <https://doi.org/10.1523/JNEUROSCI.3476-04.2005> (2005).
55. Huxlin, K. R. *et al.* Perceptual relearning of complex visual motion after V1 damage in humans. *J. Neurosci.* **29**, 3981–3991. <https://doi.org/10.1523/JNEUROSCI.4882-08.2009> (2009).
56. Das, A., Tadin, D. & Huxlin, K. R. Beyond blindsight: Properties of visual relearning in cortically blind fields. *J. Neurosci.* **34**, 11652–11664. <https://doi.org/10.1523/JNEUROSCI.1076-14.2014> (2014).
57. Fronius, M., Cirina, L., Cordey, A. & Ohrloff, C. Visual improvement during psychophysical training in an adult amblyopic eye following visual loss in the contralateral eye. *Graefes Arch. Clin. Exp. Ophthalmol.* **243**, 278–280. <https://doi.org/10.1007/s00417-004-1014-8> (2005).
58. Ostrovsky, Y., Andalman, A. & Sinha, P. Vision following extended congenital blindness. *Psychol. Sci.* **17**, 1009–1014. <https://doi.org/10.1111/j.1467-9280.2006.01827.x> (2006).
59. Huang, C.-B., Zhou, Y. & Lu, Z.-L. Broad bandwidth of perceptual learning in the visual system of adults with anisometric amblyopia. *Proc. Natl. Acad. Sci.* **105**, 4068–4073. <https://doi.org/10.1073/pnas.0800824105> (2008).
60. Zhou, J. *et al.* The eye limits the brain's learning potential. *Sci. Rep.* **2**, 364. <https://doi.org/10.1038/srep00364> (2012).
61. Chung, S. T. L., Legge, G. E. & Cheung, S. Letter-recognition and reading speed in peripheral vision benefit from perceptual learning. *Vis. Res.* **44**, 695–709. <https://doi.org/10.1016/j.visres.2003.09.028> (2004).
62. Polat, U., Ma-Naim, T., Belkin, M. & Sagi, D. Improving vision in adult amblyopia by perceptual learning. *Proc. Natl. Acad. Sci.* **101**, 6692–6697. <https://doi.org/10.1073/pnas.0401200101> (2004).
63. Grossman, E. D., Blake, R. & Kim, C.-Y. Learning to see biological motion: Brain activity parallels behavior. *J. Cogn. Neurosci.* **16**, 1669–1679. <https://doi.org/10.1162/0898929042568569> (2004).
64. Clark, J. F., Ellis, J. K., Bench, J., Khoury, J. & Graman, P. High-performance vision training improves batting statistics for university of Cincinnati baseball players. *PLoS One* **7**, e29109. <https://doi.org/10.1371/journal.pone.0029109> (2012).
65. Deveau, J., Ozer, D. J. & Seitz, A. R. Improved vision and on-field performance in baseball through perceptual learning. *Curr. Biol.* **24**, R146–R147. <https://doi.org/10.1016/j.cub.2014.01.004> (2014).
66. Xiao, L.-Q. *et al.* Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.* **18**, 1922–1926. <https://doi.org/10.1016/j.cub.2008.10.030> (2008).
67. Deveau, J. Broad-based visual benefits from training with an integrated perceptual-learning video game. *Vis. Res.* **99**, 134–140 (2014).
68. Green, C. S. & Bavelier, D. Action video game modifies visual selective attention. *Nature* **423**, 534–537. <https://doi.org/10.1038/nature01647> (2003).
69. Green, C. S. & Bavelier, D. Action-video-game experience alters the spatial resolution of vision. *Psychol. Sci.* **18**, 88–94. <https://doi.org/10.1111/j.1467-9280.2007.01853.x> (2007).
70. Li, R., Polat, U., Makous, W. & Bavelier, D. Enhancing the contrast sensitivity function through action video game training. *Nat. Neurosci.* **12**, 549–551 (2009).
71. Green, C. S., Li, R. & Bavelier, D. Perceptual learning during action video game playing. *Top. Cogn. Sci.* **2**, 202–216. <https://doi.org/10.1111/j.1756-8765.2009.01054.x> (2010).
72. Li, R. W., Ngo, C., Nguyen, J. & Levi, D. M. Video-game play induces plasticity in the visual system of adults with amblyopia. *PLoS Biol.* **9**, 11 (2011).
73. Deveau, J. & Seitz, A. R. Applying perceptual learning to achieve practical changes in vision. *Front. Psychol.* **5**, 1166. <https://doi.org/10.3389/fpsyg.2014.01166> (2014).
74. McGovern, D. P., Webb, B. S. & Peirce, J. W. Transfer of perceptual learning between different visual tasks. *J. Vis.* **12**, 4–4. <https://doi.org/10.1167/12.11.4> (2012).
75. Wang, L., Sharifian, F., Napp, J., Nath, C. & Pollmann, S. Cross-task perceptual learning of object recognition in simulated retinal implant perception. *J. Vis.* **18**, 22. <https://doi.org/10.1167/18.13.22> (2018).
76. Wang, L., Marek, N., Steffen, J. & Pollmann, S. Perceptual learning of object recognition in simulated retinal implant perception—The effect of video training. *Transl. Vis. Sci. Technol.* **10**, 22. <https://doi.org/10.1167/tvst.10.12.22> (2021).
77. 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).
78. Mansfield, J. S., Atilgan, N., Lewis, A. M. & Legge, G. E. Extending the MNREAD sentence corpus: Computer-generated sentences for measuring visual performance in reading. *Vis. Res.* **158**, 11–18. <https://doi.org/10.1016/j.visres.2019.01.010> (2019).

## Author contributions

K.E.K.R. designed the experiment, collected data, performed analysis, and wrote the manuscript. K.M. provided guidance and reviewed the manuscript. J.S.P. designed the experiment, performed analysis, and wrote the manuscript.

## Funding

Research reported in this publication was supported by: The William M. Wood Foundation, Bank of America, Trustee; The Fulbright Foundation in Greece; The Haseotis Family Foundation; Peter D. Pezaris; Ana and Philippe Laffont; John and Genevieve Wolfe; William and Lori Goldenthal. The content is solely the responsibility of the authors and does not necessarily represent views of their 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](http://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) 2022