

# Learning to see again: Perceptual learning of simulated abnormal on- off-cell population responses in sighted individuals

**Rebecca B. Esquenazi**

Department of Psychology, University of Washington, USA



**Kimberly Meier**

Department of Psychology, University of Washington, USA



**Michael Beyeler**

Department of Computer Science, University of California, Santa Barbara, Santa Barbara, California, USA  
Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, California, USA



**Geoffrey M. Boynton**

Department of Psychology, University of Washington, USA



**Ione Fine**

Department of Psychology, University of Washington, USA



Many forms of artificial sight recovery, such as electronic implants and optogenetic proteins, generally cause simultaneous, rather than complementary firing of on- and off-center retinal cells. Here, using virtual patients—sighted individuals viewing distorted input—we examine whether plasticity might compensate for abnormal neuronal population responses. Five participants were dichoptically presented with a combination of original and contrast-reversed images. Each image ( $I$ ) and its contrast-reverse ( $I'$ ) was filtered using a radial checkerboard ( $F$ ) in Fourier space and its inverse ( $F'$ ).  $[I * F'] + [I' * F]$  was presented to one eye, and  $[I * F] + [I' * F']$  was presented to the other, such that regions of the image that produced on-center responses in one eye produced off-center responses in the other eye, and vice versa. Participants continuously improved in a naturalistic object discrimination task over 20 one-hour sessions. Pre-training and post-training tests suggest that performance improvements were due to two learning processes: learning to recognize objects with reduced visual information and learning to suppress contrast-reversed image information in a non-eye-selective manner. These results suggest that, with training, it may be possible to adapt to the unnatural on- and off-cell population responses produced by electronic and optogenetic sight recovery technologies.

## Introduction

Dramatic progress has been made in sight restoration technologies over the last decade. Four types of retinal electronic devices have been implanted in patients, two of which are commercially approved (da Cruz et al., 2016; Rizzo et al., 2014; Stingl et al., 2015), with several others in active development (Ayton, Blamey, Guymer, & Luu, 2014; Ferlauto et al., 2018; Fujikado et al., 2016; Genovesi-Ebert et al., 2014; Hornig, 2017; Lorach et al., 2015; Palanker, Le Mer, Mohand-Said, Muqit, & Sahel, 2020; Saunders et al., 2014). Other groups are actively implanting (Beauchamp et al., 2020; Bosking, Beauchamp, & Yoshor, 2017; Bosking, Sun, et al., 2017; Morillas et al., 2007; Murphrey, Maunsell, Beauchamp, & Yoshor, 2009) or developing implants for cortical stimulation (Chen, Wang, Fernandez, & Roelfsema, 2020; Troyk, 2017). Gene therapy has been approved for Leber congenital amaurosis, a photoreceptor disorder, and more than a dozen human gene therapy trials for sight restoration are underway (Cho, Bolo, Park, Sengillo, & Tsang, 2019). The first optogenetic clinical trials have begun (Sahel et al., 2021) and others will likely launch within the next 2 years (Liu, Fattah, & Degenaar, 2020). Multiple other technologies such as stem cell transplantation are also in development (Cuevas, Parmar, & Sowden, 2019; Garita-Hernandez

Citation: Esquenazi, R. B., Meier, K., Beyeler, M., Boynton, G. M., & Fine, I. (2021). Learning to see again: Perceptual learning of simulated abnormal on- off-cell population responses in sighted individuals. *Journal of Vision*, 21(13):10, 1–20, <https://doi.org/10.1167/jov.21.13.10>.

<https://doi.org/10.1167/jov.21.13.10>

Received March 17, 2021; published December 22, 2021

ISSN 1534-7362 Copyright 2021 The Authors



et al., 2019; Gasparini, Llonch, Borsch, & Ader, 2019). Within a decade, many blind individuals are likely to have a wide range of options for sight restoration (Fine & Boynton, 2015; Ghezzi, 2015; Roska & Sahel, 2018; Scholl et al., 2016). Here, we focus on the neural signals elicited by retinal electronic and optogenetic/small molecule photoswitch technologies.

Retinal implants such as the Argus II (Second Sight) and Alpha-IMS (Retina Implant AG) and cortical implants such as CortiVis and the Orion (Second Sight) convert visual input from a camera into electrical impulses that trigger an array of microelectrodes that stimulate retinal or visual cortical cells (Mills, Jalil, & Stanga, 2017; Schmidt et al., 1996; Weiland, Walston, & Humayun, 2016). Current retinal and cortical electronic implants seem to be capable of providing useful (Erickson-davis & Korzybska, 2020) though relatively poor vision (Stingl et al., 2017). Limiting factors of these devices include a relatively small number of electrodes, degenerative and/or surgical damage to the retina, retinal axonal stimulation, and difficulties maintaining close proximity between the electrodes and the retinal or cortical surface (Ahuja et al., 2013; de Balthasar et al., 2008; Rizzo et al., 2014; Stingl et al., 2015, 2017).

Optogenetic proteins create light-sensitive ion channels that make cells responsive to light (Bamann, Nagel, & Bamberg, 2010). In the context of sight recovery, these optogenetic proteins are delivered to remaining retinal cells (such as ganglion, amacrine, or bipolar cells) to create artificial photoreceptors (Busskamp, Picaud, Sahel, & Roska, 2012). Optopharmacological tools such as photoswitch compounds (Kramer, Mourot, & Adesnik, 2013) elicit light sensitivity by dynamically activating and deactivating ion channels within remaining retinal cells via exposure to particular wavelengths of light (Polosukhina et al., 2012; Tochitsky et al., 2014).

Critically, the vision provided by all of these technologies differs substantially from normal sight, even if implants are extremely high resolution. In biologically natural vision, stimuli that excite on-center retinal cells always inhibit off-center cells in the same retinal location, and vice versa. It will be extremely challenging for electronic, optogenetic, and optopharmacological approaches to selectively stimulate on- and off-cells in a naturalistic complementary manner, in which activity in on-cells is accompanied by the suppression of off-cells. For current electronic prostheses, this simultaneous stimulation of on- and off-cells likely plays a negligible role in limiting resolution (Zrenner, 2002) and a minor role in limiting sensitivity (Ni & Maunsell, 2010). However, as these technologies improve, the unselective stimulation of populations of neurons whose natural firing patterns are anti-correlated is likely to become more of a concern.

Engineering solutions to address the problem of simultaneous on- and off-cell stimulation are being attempted within the context of both electrical implants and optogenetic technologies (for a recent review see Tong, Meffin, Garrett, & Ibbotson, 2020).

Visual cortical prostheses implanted in later regions of the visual pathway would entirely bypass the problem of unselective on- and off-cell stimulation. However, the selectivity of single neurons in areas beyond V2 is highly complex, making the consequences of unselective stimulation unpredictable. Thus, current electronic prostheses, such as the Orion (NCT02983370; Beauchamp et al., 2020) and CortiVis (NCT02747589; Chen et al., 2020) are located as close to the V1 foveal pole as possible.

Another engineering approach is to develop devices that mimic naturalistic stimulation patterns. This goal is technically challenging because it requires not only selectively stimulating on- and off-cells, but also requires identifying these cells *in vivo*. Recently, Shah et al. (2019) electrically recorded the response patterns of monkey on- and off- retinal ganglion cells (RGCs) to a white noise stimulus and used the data to construct a dictionary of RGC activity patterns and their corresponding visual percepts. Dictionary patterns were then combined linearly to selectively stimulate on- and off- RGCs, to generate a firing pattern whose predicted percept closely resembled the original white noise image. Although promising, this approach requires generating cell-specific RGC dictionaries, maintaining stable cell-specific stimulation over time, and possibly developing a more sophisticated nonlinear model (Demb, Haarsma, Freed, & Sterling, 1999; Hochstein & Shapley, 1976) to adequately replicate population responses for complex naturalistic images with widely varying visual properties.

In optogenetic therapy, compounds are being designed to make on- and off- RGCs differentially responsive to specific wavelengths of light and/or selectively produce excitatory (on) and inhibitory (off) responses with the goal of creating distinct firing patterns in each cell type (Barrett, Panesar, Scally, & Pacey, 2012; Berry et al., 2017). Creating naturalistic patterns will further require selective transfection of on- and off-cells, optogenetic proteins with narrow spectral sensitivity, and fast temporal dynamics that are within safe light levels.

Thus, for the foreseeable future, optogenetic, small molecule photoswitches, and electronic prostheses will not be able to selectively stimulate on- versus off- retinal or cortical cells. Will individuals be able to adapt to these abnormal population responses (Beyeler, Rokem, Boynton, & Fine, 2017; Fine, Cepko, & Landy, 2015)?

When on-cell pathways are compromised at birth, the impact seems to be relatively minor. Individuals with complete Schubert–Bornschein congenital stationary night blindness type 1 genetic deficits have severely

compromised on-bipolar pathways (Bijveld et al., 2013; Cibis & Fitzgerald, 2001). Yet these patients show surprisingly good visual performance under photopic conditions, with an average visual acuity of 0.3 logarithm of the minimum angle of resolution (logMAR) (Snellen 20/40) (Zeitz, Robson, & Audo, 2015) and report no perceptual difficulties beyond their acuity loss. Thus, off-pathways alone carry enough visual information to produce near-normal vision. However, individuals with congenital stationary night blindness type 1 have experienced stimulation of just one bipolar cell pathway input since birth. The ability to learn to decode simultaneous on- and off-pathway stimulation in adulthood is less clear.

To date, only one study has directly compared the perceptual decoding of light with electrical stimulation. Ni and Maunsell (2010) trained macaques to detect a phosphene induced by microstimulation within V1 over the course of several months. Training resulted in a significant decrease of detection thresholds, a roughly 10-fold decrease in the threshold current. This improvement in the ability to detect microstimulation at specific V1 sites resulted in a decrease in the ability to detect visual stimuli presented at the same retinotopic location (threshold currents increased by a factor of 1.7–7.0). Retraining with a visual stimulus then interfered with detecting electrical stimulation, suggesting that the optimal detection of electrical stimulation required a long-term reconfiguration of the decoding of neuronal responses within V1. Adaptation to electrical stimulation was slow (>10,000 trials), and both training and learning were retinotopically specific; it, therefore, remains an open question whether patients can learn to decode abnormal on- and off-cell population responses under more naturalistic learning conditions.

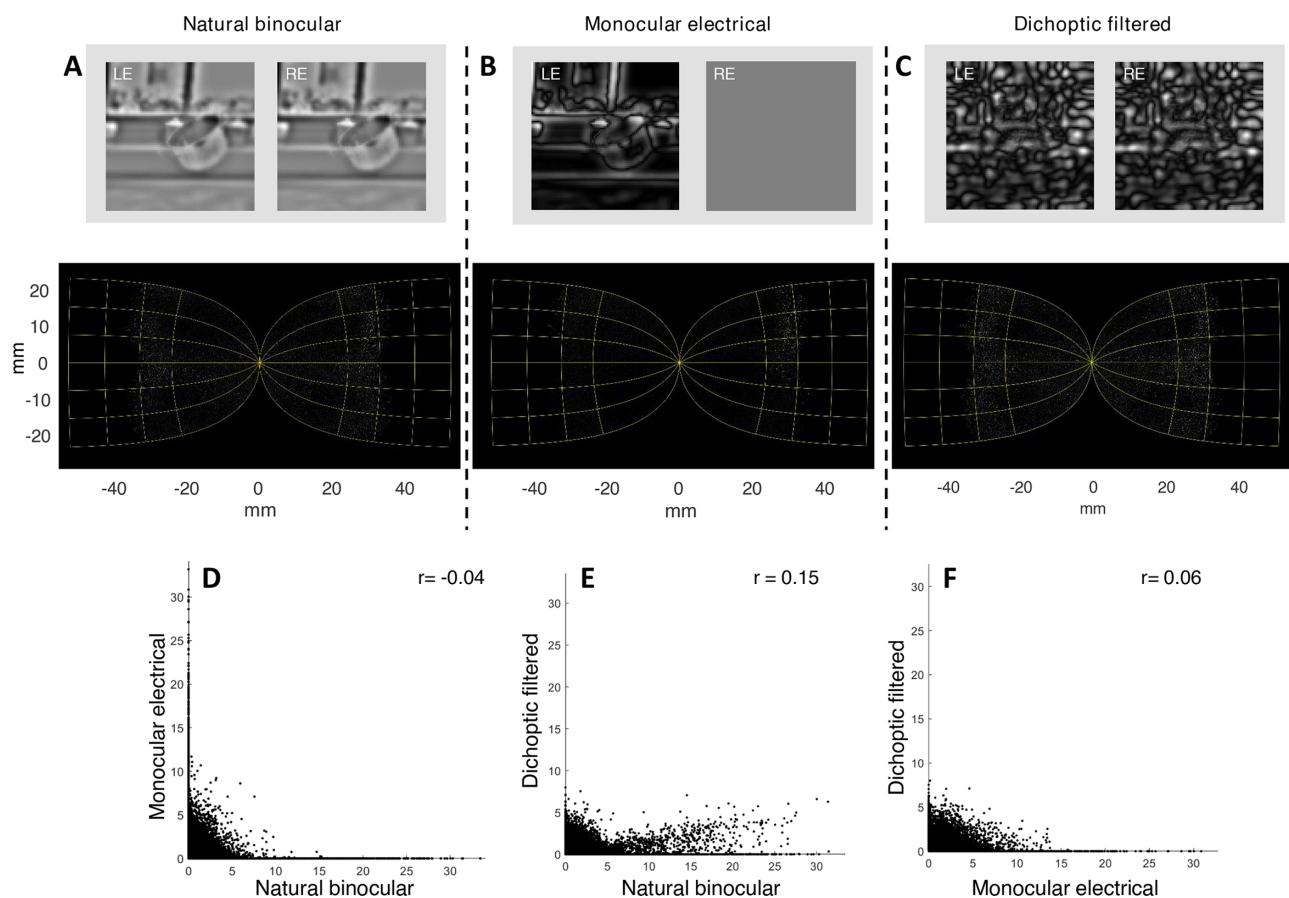
Patients with cochlear implants learn to make use of their distorted input with remarkable speed, even when implanted in adulthood (Fallon, Irvine, & Shepherd, 2008). This plasticity is rapid and persistent and requires mere hours of training (Fritz, Shamma, Elhilali, & Klein, 2003). However, plasticity may be very different for visual implants. Primary areas of the visual hierarchy in cortex (V1) show far less plasticity in adulthood than primary auditory (A1) or somatosensory (S1) cortical areas (Beyeler, Boynton, Fine, & Rokem, 2017; Ghose, Yang, & Maunsell, 2002). The reason for this difference remains unclear, but a possible explanation is that, compared with visual pathways, there is significantly more subcortical processing within somatosensory and auditory pathways. Thus, primary cortical areas A1 and S1 can be considered as higher in their respective processing pathways than V1 within the visual processing hierarchy and may, as a result, be more plastic (Haak & Beckmann, 2019).

The aim of the current study was to produce abnormal population responses within V1 that serve as

a rough proxy for the abnormal population responses elicited by electronic sight restoration technologies. Five participants were trained in an object discrimination task using a dichoptic (a different image to each eye) presentation. Images were convolved with filters via multiplication in the Fourier domain. Our filter, F (**Figure 2A**), was defined as a radial checkerboard in Fourier space such that, when convolved with an image, I, only one-half of the total combination of spatial frequencies and orientations in I were passed through. Convoluting with the filter's inverse ( $F'$ ) passed the other half of the spatial frequencies and orientations in I. Images and filters were combined such that  $[I * F'] + [I' * F]$  was presented to one eye, and  $[I * F] + [I' * F']$  to the other (where  $*$  denotes two-dimensional convolution). Thus, regions of the resulting image that produced on-cell responses in one eye produced off-cell responses in the other eye at the corresponding visual location, and vice versa.

Although it is impossible to recreate the effects of electronic retinal or cortical stimulation in sighted individuals, the dichoptic stimuli described here create a visual stimulus that similarly scrambles the cortical input. **Figure 1** shows simulated cortical responses (see Methods) for three example stimuli: a natural binocular image (**Figure 1A**), monocular electronic retinal stimulation (**Figure 1B**), and the filtered dichoptic images used in our experiment (**Figure 1C**). The upper images of **Figures 1A–C** show the predicted retinal input into cortex (in visual space). For natural cortical stimulation, the input from retina to cortex is a slightly blurred version of the original image. Most of this represents receptive field sizes in the retina (however, to limit computational time, the retinal filter bank was slightly low pass, which may have also contributed to blurring; see Methods). The lower images of **Figures 1A–C** show cortical responses for approximately 32,500 typical V1 cells, distributed evenly over the cortex. For all types of stimulation, responses are relatively sparse over the cortical surface, due to the orientation and spatial frequency selectivity of individual cells. **Figures 1D–F** show selected pair-wise correlations between cortical responses (in arbitrary response units) across these three stimulation protocols. The population responses produced by natural binocular stimulation are uncorrelated with the population responses elicited by monocular electrical stimulation (**Figures 1A, B, and D**). Similarly, the population responses produced by our filtered dichoptic stimuli are also only weakly correlated with naturalistic population responses (**Figures 1A, C, and E**).

These low correlation values should not be interpreted as implying that there was no relationship between the population responses elicited by these unnatural stimulation methods and natural stimulation. Rather, some cells were correlated, whereas others were anticorrelated. Nonetheless, these simulations show that



**Figure 1.** (A–C) Examples of simulated cortical responses for three stimuli: the original binocular image, monocular electrical retinal stimulation, and the dichoptic filtered images used in our experiment. The cortical input image for each panel is shown as an inset. For all types of stimulation, responses are relatively sparse over the cortical surface, due to the selectivity of individual cells. (D–F) Cross-correlations between cortical responses (in arbitrary response units) across these three stimulation protocols. LE, left eye; RE, right eye.

learning to interpret either electrical stimulation or the stimuli used in the current study, requires remarkable flexibility in decoding cortical population responses.

The population responses for our dichoptic stimuli are also dissimilar from the population responses for monocular electrical stimulation, which is typical of current sight restoration methods (Figure 1B, C, and E). Thus, our paradigm should only be considered a proxy for electrical sight restoration methods, in that our manipulation effectively disrupts early population responses—not because the population responses elicited by dichoptic stimulation directly resemble the population responses elicited by electrical stimulation.

For comparison, cortical responses to monocular and binocular natural input are strongly correlated ( $r = \sim 0.8$ , data not shown). Interestingly, even heavy blurring of the input to cortex also produces simulated neuronal population responses that are highly correlated ( $r > \sim 0.9$ ) with the responses to unblurred stimuli (this general result holds for all neurally realistic parameters, because most V1 cells are tuned for

relatively low spatial frequencies). This finding may explain why recognition is robust to substantial blurring and suggests that pixelation and/or blurring do not fully capture the difficulty of interpreting the neural signals generated by electrical or optogenetic stimulation.

Participants in our study showed significant improvement in object discrimination over 14, one-hour training sessions. The transfer of learning in pre-training and post-training tests were used to examine the mechanisms underlying performance improvements. Collectively, our results suggest that it may be possible to adapt to the unnatural on- and off-cell population responses likely to be produced by electronic and optogenetic sight recovery technologies.

## Methods

This study was approved by the University of Washington's Institutional Review Board (Study

#3868), and carried out in accordance with the Code of Ethics of the Declaration of Helsinki. Informed consent was obtained before the start of the first experimental session.

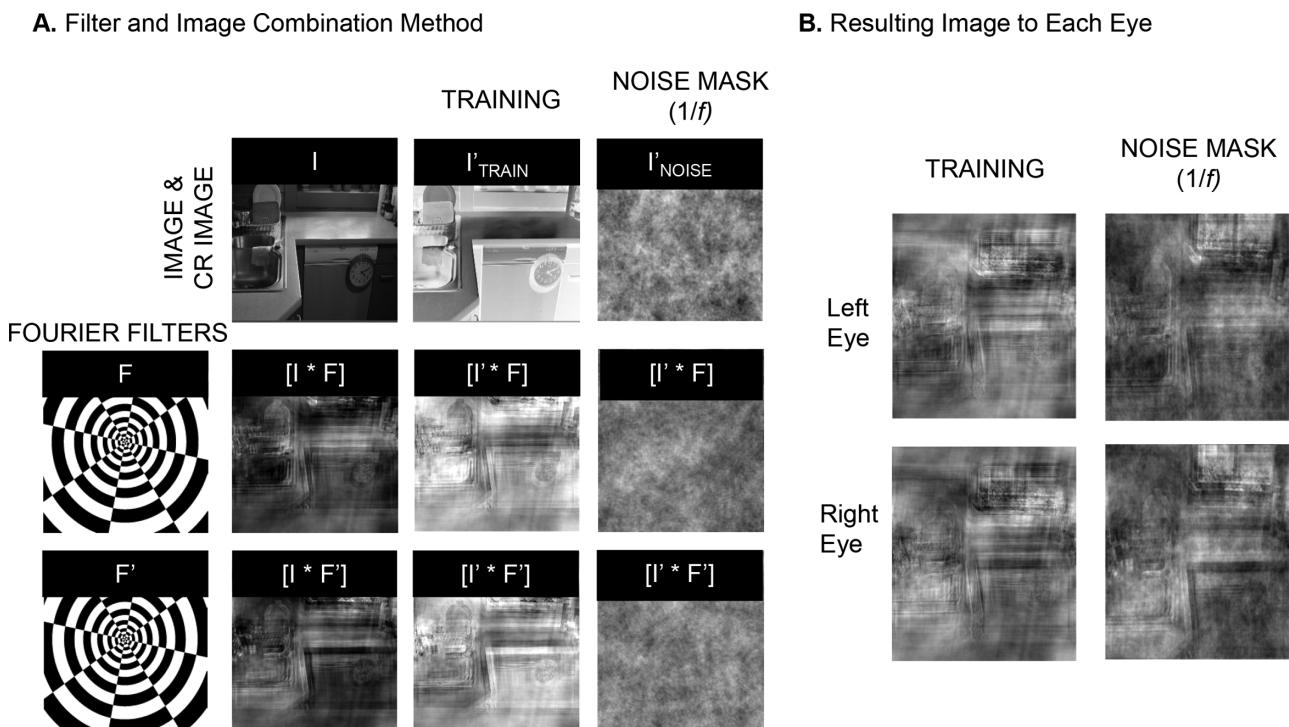
## Participants

Five naïve observers (4 males) aged 25 to 32 years ( $M = 27$ ) were recruited through word of mouth at the University of Washington. Binocular and monocular visual acuity was assessed using FrACT (Bach, 1996, 2007), and stereoacuity was assessed using the Randot Stereotest (Stereo Optical Co. Inc.). All observers had normal or corrected to normal visual acuity (defined as  $\geq 0.2$  logMAR or 20/30 Snellen), no interocular acuity differences greater than 0.1 logMAR, and normal stereoacuity. The suppression check of the Randot Stereotest was also used to confirm that no participant experienced abnormal suppression.

## Stimulus and procedure

Stimuli were presented dichoptically using a custom-built stereoscope that consisted of two cold mirrors mounted on posts, rotated at a 45° angle to capture input from a monitor and reflect it separately into each eye. A 32" LED monitor at a viewing distance of 136 cm with 2560 × 1440 pixel resolution projected to each cold mirror. Each monitor spanned 28.9°, and these monitors were the only light source in the room. Each stimulus spanned 768 × 768 pixels (8.84°). Stimuli had a mean luminance of 132 cd/m<sup>2</sup> and were presented on a mid-grey background with luminance 80 cd/m<sup>2</sup>.

At the beginning of each session, participants used a nonius task to align the screens to account for any fixation disparities. Because our goal was to isolate the ability to learn to decode abnormal on- and off-population responses (rather than to simulate a specific sight restoration technology), all stimuli were presented at high resolution.



**Figure 2.** Example of filtering for dichoptic presentation. (A) The two upper left panels show an example scene ( $I$ ), and the contrast-reversed version of that scene ( $I'$ ). The upper right panel shows the noise mask 1/f noise. The leftmost panels show two filters:  $F$  and  $F'$ . Filters images represent amplitudes in the Fourier domain, with spatial frequency increasing with distance from the center of the image and orientation changing with polar angle. The filters are paired complements, so the full spatial frequency and orientation content of the scenes is divided equally across the two filters. The lower middle panels show the convolution of the original ( $I$ ), contrast reversed ( $I'$ ), and 1/f NOISE images with Fourier filters  $F$  and  $F'$ . (B) Examples of filtered images presented to left and right eyes for the training condition and 1/f noise condition. Although these images do not resemble the perceptual experience of simultaneous on- and off-cell stimulation, interpretation of these images requires an analogous process of interpreting a garbled population response.

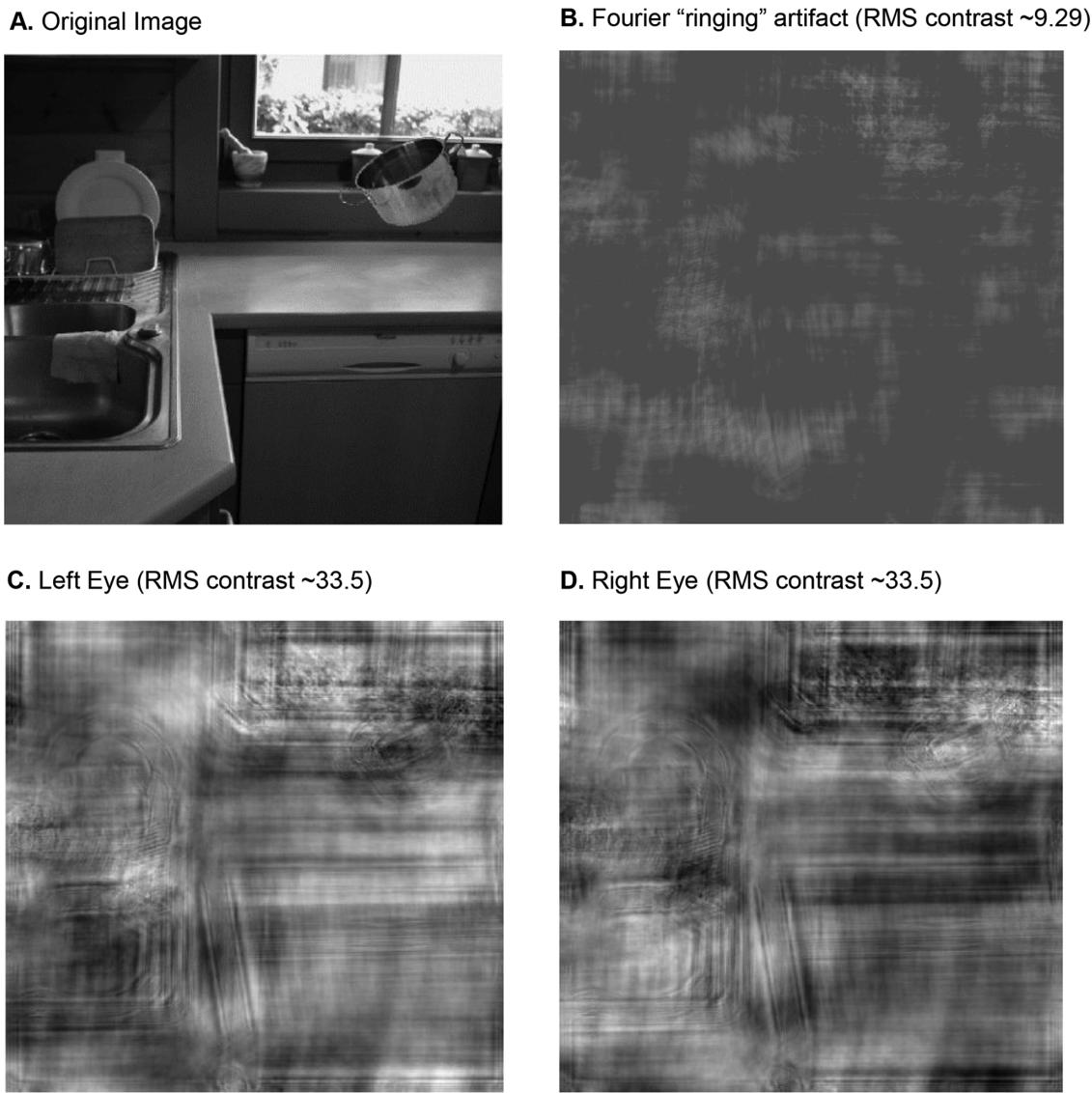


Figure 3. (A) Unfiltered image with a pot (upper right corner) overlaid as part of the object discrimination task. (B) The Fourier artifact. (C, D) Example of the resulting dichoptic stimuli after the filter was applied to the top image (see Filtering section of the main text).

Stimuli were created by manipulating the spatial frequency and orientation information of naturalistic scenes in the Fourier domain, see [Figures 2](#) and [3](#).

#### **Stimulus set**

Seventeen scenes of different household settings (e.g., kitchen, living room, bathroom, bedroom) from the SCEGRAM Database ([Öhlschläger & Võ, 2017](#)) were used as backgrounds. A separate set of 45 various household objects (e.g., clock, rolling pin, dish soap) were overlaid onto each scene ([Figure 4B](#)). To minimize image-specific learning, objects could take on one of six possible logarithmically spaced sizes ranging from 22.2% (1.96°) to 66.7% (5.90°) of the original object (768 × 768 pixels or 8.84°). Each object was located

randomly within the scene and was rotated by up to 30° in either direction. Thus, there were more than 13,000 unique images in the training set.

#### **Filtering**

Binarized radial checkerboard filters in Fourier space were used to present separate spatial frequency and orientation information to each eye ([Figure 2](#)). In the Fourier domain, increasing spatial frequency is represented by distance from the center of the image, and orientation is represented along the polar angle dimension.

The spatial frequency content of the filters,  $F_f$ , was defined as:  $F_f = \sin(2\pi n f_0 f^{\frac{1}{n}})$ , where  $f$  is the spatial frequency of the Fourier image,  $f_0 = 13$  and controls

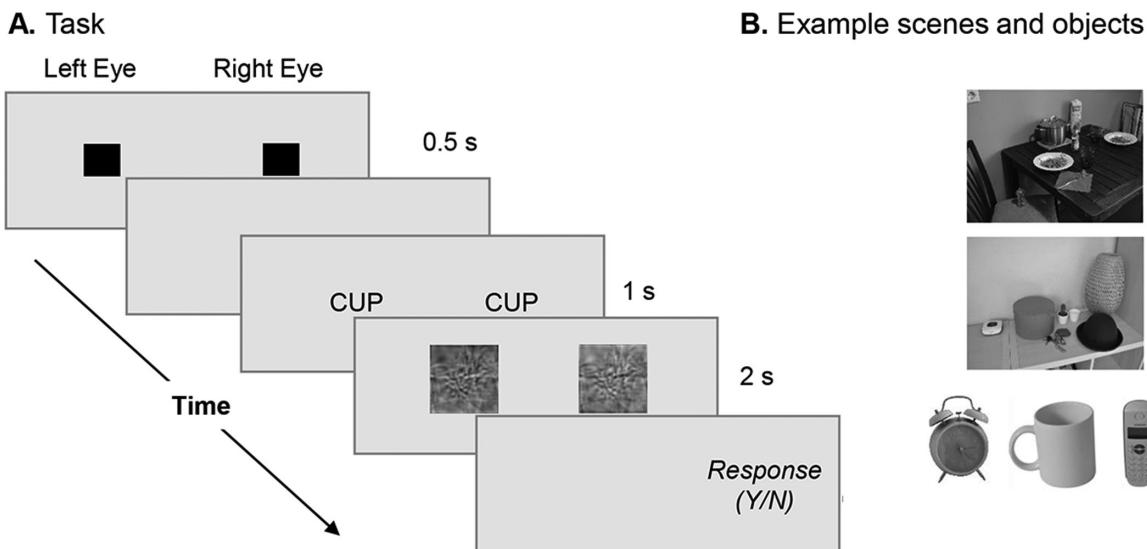


Figure 4. (A) The object discrimination task. During each trial, the participant reported whether or not the cued object was present within the scene. (B) Examples of unfiltered scenes and objects from the SCEGRAM database.

the overall frequency of the radial rings, and  $n$  describes the increase in ring width as a function of spatial frequency. The orientation content of the filter,  $F_9$ , was defined as:  $F_9 = \alpha_0 \alpha$ , where  $\alpha$  is the orientation of the Fourier image and  $\alpha_0$  defines the number of radial spokes.

The radial checkerboard filter was built with sharp edges in the Fourier domain, which leads to image artifacts—often seen as ringing in the spatial domain (see Figure 3). Our choice of a binary filter (rather than a filter with smooth edges in the Fourier domain) was motivated by the desire to minimize shared spatial frequency and orientation information within the images presented to each eye. These Fourier artifacts were relatively subtle, with a root mean square contrast of approximately one-third that of the original images (Figure 3B) and was unlikely to be the primary cause of masking. Pilot data (not shown) from three participants using a Fourier filter with smooth edges in the Fourier domain that almost entirely eliminated these artifacts, resulted in a very similar level of performance, and a similar or faster rate of learning.

Strong contours in our images, depending on their alignment with the orientation ( $F_9$ ) and spatial frequency ( $n$ ), and radial spokes ( $\alpha_0$ ) of our filters, resulted in strong striping (horizontal striping can be seen in  $[I * F]$  and  $[I' * F']$  in Figure 2). These stripes are due to missing alternating frequency bands, and it can be seen that the striping occurs at complementary frequencies in  $[I * F]$  and  $[I' * F']$ . The orientation, frequency, and strength of the striping depends on the orientation of the strong contours of the image in relationship to the filter bands.

It is unlikely that participants learned to make use of these Fourier artifacts to perform the task

(e.g., by recognizing objects based on a characteristic ringing or striping structure) because the object images in the task were always presented at random locations, orientations, and sizes, and the overall scaling of the background also varied over each trial.

The final filters were the product of  $F_f$  and  $F_9$ , with one being the negative of the other:  $F = F_f \times F_9$ , and  $F' = -F_f \times F_9$ . Finally, these filters were scaled  $F = (F + 1)/2$ , and binarized to values 0 and 1. Each filter was the complement of the other, so the full spatial frequency and orientation content of both the original and the contrast-reversed scene were divided equally across the two filters and thus the two eyes.

The two upper left panels of Figure 2 show an example scene (I), and the contrast-reversed version of that scene (I'). The leftmost panels show the two radial checkerboard Fourier filters  $F$  and  $F'$ . The original (I) and the contrast-reversed scene I' ( $I' = 1 - I$ ) were each converted into the Fourier domain, multiplied with one of the two Fourier filters, and then converted back to image space using the inverse Fourier transform. The top panels in Figure 2A show original (left), contrast-reversed (middle) and 1/f images (right). The bottom two panels (Figure 2A) show the four examples of possible filtering:  $I * F$ ,  $I * F'$ ,  $I' * F$ , and  $I' * F'$  (where  $*$  denotes two-dimensional convolution) for both the training and 1/f noise stimuli.

In the training paradigm, we presented the left eye (Figure 2B, Figure 3) the sum of two filtered images,  $[I * F'] + [I' * F]$ , such that one-half of the spatial frequency and orientation content was based on the original image and the other half was based on the contrast reversed image. In the right eye, we presented

the sum of the complementary filtered images,  $[I * F] + [I' * F']$ .

Thus, the monocular input contains half the information from the original image, and one-half of the information from a contrast-reversed image, so in theory, all the information from the original image is retained; however, the normal pattern of population responses (wherein on- and off-cells with similar spatial frequencies and orientations tend to be highly correlated in their firing) is disrupted. In the binocular input, all the information from both the original image and the contrast-reversed image is retained. In the absence of suppression, on- and off-cells with identical tuning profiles would be simultaneously stimulated by the combination of input from the left and the right eye, in a close analogue of the effects of electrical stimulation.

Note that the sum,  $[I * F] + [I' * F']$ , equals the original image  $I$ , and  $[I' * F] + [I' * F']$ , equals the original contrast reversed image  $I'$ ; thus, all the spatial frequency and orientation information of both the original and contrast reversed image is preserved. Thus, with optimal decoding, stimuli are lossless. The sum of the distorted images in each eye results in a blank image.

### Task

A brief fixation cue (0.5 second) began each trial (Figure 4A). After a 0.5-second pause, a word cue told the participants what the target object was (e.g., cup, clock). After the word cue, a scene with an overlaid object was displayed for up to 2 seconds, or until the participant responded with a key press. To create a dynamic scene that more closely resembled naturalistic retinal input, and to encourage generalizable learning by creating more variation in the retinal image, there was a simulated panning action within each 2-second trial. The field of view drifted to the right or left, at a rate that was uniformly distributed between 0.21 and 0.52°/s. The image also expanded or contracted at a maximum rate of 0.35°/s.

Participants performed a two-alternative forced-choice object discrimination task, judging whether or not the scene contained the cued object. In each trial, there was a 50% chance that the scene contained the prompted object, or a different distractor object. Auditory feedback was provided after each trial to indicate whether the answer was correct or incorrect. Participants were not given specific instructions on where to look within the scene.

### Learning protocol

Each experimental session consisted of 400 trials. Participants were offered a break every 40 trials to

mitigate fatigue, and only performed one session per day. Participants carried out 20 approximately 1-hour sessions in total. The first and last three sessions contained 100 trials of the training stimulus set and 100 trials of each of the three (monocular, filter-switched, 1/f noise) pre-test and post-test conditions, for a total of 400 trials. The remaining 14 sessions consisted of 400 trials of the training set.

## Pre-test and post-test conditions

### Monocular presentation

Participants were shown the filtered image to the left or right eye only (randomly interleaved across trials). A blank gray screen matched in mean luminance was presented to the other eye.

### Filter switched

Left and right eye filters were switched across the two eyes, such that the eye trained to view  $[I * F'] + [I' * F]$  received  $[I * F] + [I' * F']$ , and vice versa.

### 1/f noise

The contrast-reversed image  $I'$  was replaced by a 1/f noise pattern, such that the eye trained to view  $[I * F'] + [I' * F]$  received  $[I * F'] + [1/f * F]$ , and the eye trained to view  $[I * F] + [I' * F']$  received  $[I * F] + [1/f * F']$ .

## Statistical analyses

Hits, misses, correct rejections, and false alarms from conditions within each session of two-alternative forced-choice trials were converted into d-prime ( $d'$ ) units (Green & Swets, 1966). Linear mixed models were fit to the data using the lme4 package in R (R Core Team, 2018).

### Learning over time for the trained stimulus set

Two separate linear mixed effects regression analyses were carried out to examine performance (in  $d'$ ) on the trained stimulus set as a function of experimental session. Participants were treated as a random factor and the session number was treated as a fixed factor in both models. The first model included all 20 sessions, including the first three and last three sessions, which only included 100 trials each of the trained stimuli. The second model was restricted to the middle 14 sessions that all contained 400 trials each of the trained stimuli. We also calculated a learning index for each participant, which represents the proportion increase

in  $d'$  normalized to the average  $d'$  of the three pre-test sessions (Fine & Jacobs, 2002).

### Performance in pre-test and post-test conditions

A full model assessing differences in condition, time, and the interaction of condition by time was carried out. Fixed effects were treated as two different factors: (1) the ‘condition’ factor contained four levels consisting of all types of pre-tests and post-tests, and (2) time, which assessed the progression of learning between the pre-test and the post-test. Each analysis was conducted as a linear mixed effects model with participant as a random effect factor. Three planned comparisons were used to assess performance differences in each of the three pre-test and post-test conditions, compared with the trained stimulus.

## Modeling

For illustrative purposes, we provide the results of some simple simulations of expected population responses in V1 for a variety of stimulation paradigms (Figure 1).

We modeled retinal receptive fields as a bank of circular center-surround difference of Gaussian filters (with 4 sizes), with centers fixed at twice the size of the surround, with the overall size scaling as a function of eccentricity (Watson, 2014). Both on-center and off-center difference of Gaussian filters were modeled.

For natural vision (including our dichoptic stimuli), we assumed the response strength of each retinal cell could be described as the rectified sum of the dot product of each retinal receptive field with the image projected onto the retina.

In the case of electrical stimulation, we assumed tiny electrodes flush to the retinal surface, so current spread was not modeled. We further assumed unselective stimulation of on- and off-cells, without axonal stimulation (Beyeler et al., 2019), with the response strength of each retinal cell being linearly related to the electrical stimulation current, which was in turn linearly related to the luminance of the stimulus.

Transformation from the retinal to the cortical surface was carried out using a template derived from a conformal map developed by Schwartz et al. (Polimeni, Balasubramanian, & Schwartz, 2006; Schwartz, 1980, 1994), in which two-dimensional visual space is projected onto the two-dimensional flattened cortex as follows:  $w = k \times \log(z + a)$ , where  $z$  is a complex number representing a point in visual space,  $w$  represents the corresponding point on the flattened cortex,  $a = 0.5$  reflects the proportion of V1 devoted to the foveal representation, and  $k = 15$  is an overall scaling factor (Hinds et al., 2008).

Within the cortex, ocular dominance columns and orientation pinwheels were simulated based on work by Rojer and Schwartz (1990). Orientation columns were modeled by bandpass filtering white noise in the complex domain, with the angle representing orientation preference. We then extended the model to include ocular dominance columns as the gradient of the same filtered white noise along a single direction, thereby generating orthogonal ocular dominance and orientation columns.

Individual V1 receptive fields were modeled based on Mata and Ringach (2005), in which ON and OFF maps are simulated as the linear combination of two subregions of opposite sign, distance  $d$  apart (sampled from a distribution that declined logarithmically as a function of cortical distance), with each subregion organized with an antagonistic (push–pull manner). The size of the subregions linearly increased with eccentricity (Freeman & Simoncelli, 2011; Keliris, Li, Papanikolaou, Logothetis, & Smirnakis, 2019).

For both natural and electrical stimulation we approximated the retinal contribution to cortex as a cortical input image created as the linear sum of each retinal cells’ receptive field, weighted by its response strength. In the case of natural stimulation, this cortical input image was very similar to that produced by projecting the image directly onto the cortical surface.

For each cortical cell, the response was calculated as the rectified sum of the dot product of each cortical receptive field with the cortical input image.

When modeling cortical neuronal responses (Figure 1), we projected the entire scene onto the retina, assumed participants were fixating centrally, and assumed no neural noise.

When modeling monocular versus binocular performance on the task (Figure 7) we first calculated noise-free cortical response to the filtered target (POT) as a  $1 \times n$  vector (where  $n$  is the number of cortical cells). This could be considered a *target perceptual template*.

We then calculated noisy cortical responses to both the filtered target object and a distractor object, where Gaussian noise was added to each cell with a standard deviation proportional to the square root of that cells’ response strength. The standard deviation of the Gaussian noise was titrated to produce a  $d'$  value of approximately 1.5 in the binocular condition.

We defined the Euclidian norm of the difference between the perceptual template and the noisy cortical response to the filtered target as the signal. We defined the Euclidian norm of the difference between the perceptual template and the noisy cortical response to the filtered distractor object as the noise. Note that, unlike most signal and noise representations, a small Euclidian norm represents good performance, where the cortical response was similar to the perceptual template. We created signal and noise distributions

(Figure 7) by simulating 1000 independent trials. This was done separately for both binocular and monocular presentations, and these distributions were used to calculate  $d'$ . Although this value of  $d'$  is not directly comparable to the  $d'$  measured in our study (where participants discriminated the target from a wide variety of distractors), it nonetheless provides an estimate of the relative signal to noise available to observers in binocular versus monocular presentations.

We limited these calculations to a subregion of the scene ( $4.42^\circ$ ) containing the object, and assumed that because the task was performed using free viewing, that the participant was foveating that region when performing the task.

## Results

### Performance improvements over time for the trained stimulus set

Figure 5 shows the  $d'$  values for each session, for each individual participant. Figure 5A shows linear mixed

effects regression fits when both pre-test, post-test, and training sessions were included in the data (all 20 sessions). Participant intercepts, which reflect  $d'$  at session 1, ranged between 1.014 and 1.965 with a mean value of  $M = 1.275$ ,  $SD = 0.400$ . The slope estimate of the model, an indicator of learning rate, was  $m = 0.063$ , 95% confidence interval (CI) 0.049–0.075,  $t(94) = 9.524$ ,  $P < .0001$ ,  $SE = 0.007$ , Cohen's  $d = 0.570$ , showing that  $d'$  improved significantly over time. Over the course of 20 sessions, the predicted  $d'$  increased from 1.07 in session 1 to 2.26 in session 20, a percentage increase of 210%. When data were restricted to the 14 training sessions only, the slope estimate of the model was smaller but still significant,  $m = 0.043$  per session, 95% CI = 0.028–0.058,  $t(64) = 5.827$ ,  $P < .0001$ ,  $SE = 0.007$ , Cohen's  $d = 0.37$ .

The decrease in slope when the pre-test and post-test sessions were excluded is a function of rapid learning during the pre-test phase. The largest average session-to-session increase in  $d'$  was found between pre-test sessions 1 and 2 ( $M = 0.768$ ,  $SD = 0.104$ ).

Figure 5B shows individual differences in the rate of learning. For each participant the  $x$ -axis represents the testing session and the  $y$ -axis (Learning Index)

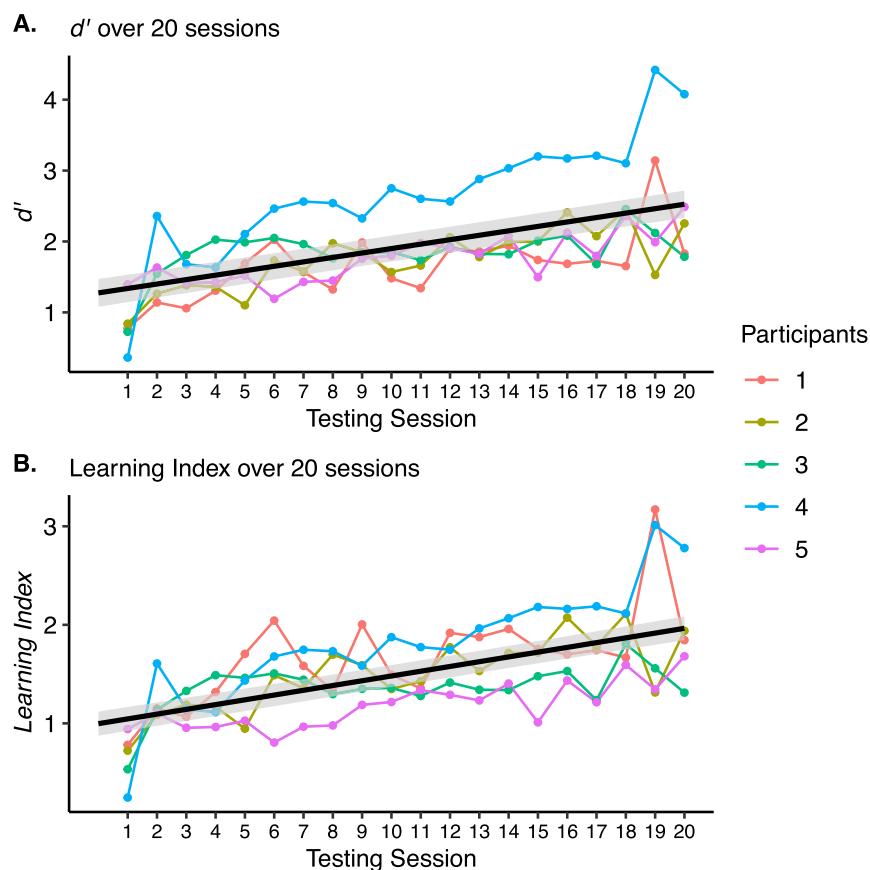
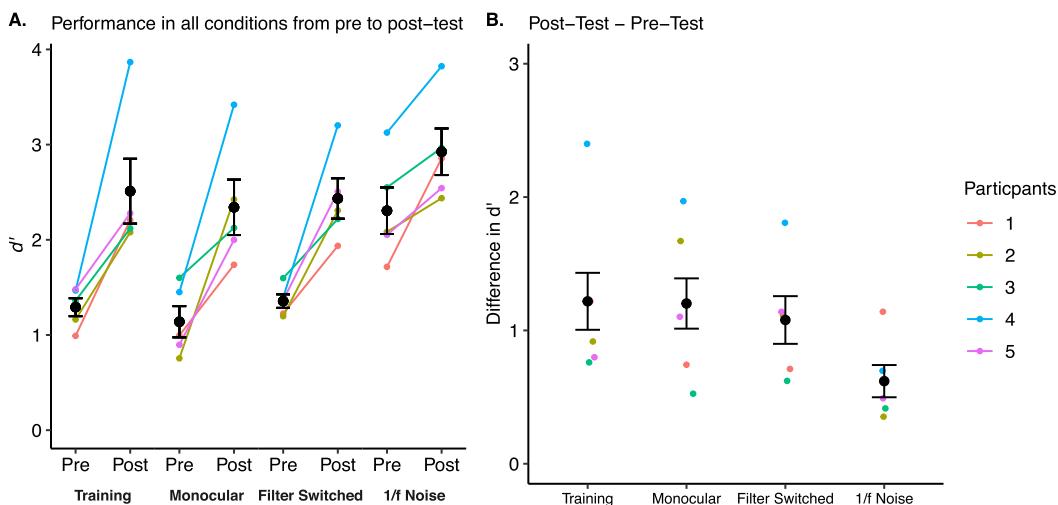


Figure 5. (A)  $d'$  scores for participants in the trained stimulus set. The regression line for data pooled across participants (black) is overlaid on individual participant scores. (B) The rate of learning for all subjects, normalized to the average  $d'$  of the three pre-test sessions. A learning index of greater than 1 shows better performance than the average  $d'$  of the participant's three pre-test sessions, a learning index of less than 1 shows worse performance than the average  $d'$  of the three pre-test sessions.



**Figure 6.** (A)  $d'$  scores for each pre-test/post-test condition. Each pre-test and post-test  $d'$  is calculated as the average of each participant's three runs in each test. Large black data points represent the average  $d'$  (across subjects) for that condition. Error bars represent standard error of the mean for each condition. (B) The change in  $d'$  scores with training, calculated as the average of each participant's three post-tests subtracted by the average of each participant's three pre-tests. Large black data points represent the average difference in  $d'$  (across subjects) for that condition. A larger difference indicates improved performance in the post-test compared with the pre-test. Error bars represent standard error of the mean for each condition. The asterisk\* represents the finding that there was a significant difference in the amount of learning in the training condition compared with the 1/f noise condition.

represents  $d'$  on that session normalized to the average  $d'$  on the first three sessions. The slope of the average learning rate was significant,  $m = 0.048$  per session 95% CI = 0.038–0.058,  $t(94) = 9.520$ ,  $P < .001$ ,  $SE = 0.005$ , Cohen's  $d = 0.610$ . Individual learning rates (slopes) varied across participants, ranging from 0.020 to 0.088, but all five participants had learning rates that were statistically significant,  $m_{\min} = 0.020$ , 95% CI = 0.001–0.037,  $P < 0.033$ ,  $SE = 0.008$ , Cohen's  $d = 0.480$ ;  $m_{\max} = 0.088$ , 95% CI = 0.064–0.111,  $P < 0.001$ ,  $SE = 0.011$ , Cohen's  $d = 0.880$ .

### Pre-tests and post-tests

The purpose of the pre-test and post-tests was to examine the underlying learning mechanisms used by participants over the course of training.

The results of a full model assessing the four stimulus conditions (trained, monocular, filter-switched, and 1/f noise) in the pre-testing and post-testing phases revealed significant main effects. An analysis of variance with Satterthwaite's method on each fixed effect revealed a main effect of condition,  $F(3,108) = 18.544$ ,  $P < .001$ ,  $\eta_p^2 = 0.340$ , and time (pre-test vs post-test)  $F(1,1) = 127.884$ ,  $P < .001$ ,  $\eta_p^2 = 0.540$ . There was a marginally significant interaction effect of condition by time,  $F(3,108) = 2.371$ ,  $P = .075$ ,  $\eta_p^2 = 0.060$ .

To understand how performance on the three pre-test and post-test conditions compared with performance

on the trained stimulus set, three additional planned tests were conducted as separate models, one for each condition. Each transfer of learning condition was compared with performance in the training condition during the pre-test ( $M_{\text{Training}} = 1.292$ ,  $SD_{\text{Training}} = 0.500$ ) and post-test ( $M_{\text{Training}} = 2.510$ ,  $SD_{\text{Training}} = 0.848$ ) (Figure 6).

### Monocular presentation

Comparisons of performance between the trained and monocular stimulus sets showed a significant main effect of time, participants improved in performance from pre-test to post-test,  $F(1,52) = 82.134$ ,  $P < .001$ ,  $\eta_p^2 = 0.61$ , but no main effect of condition,  $F(1,52) = 1.455$ ,  $P = .232$ ,  $\eta_p^2 = 0.030$  or interaction between condition and time,  $F(1,52) = 0.004$ ,  $P = .951$ ,  $\eta_p^2 < 0.001$ . Thus, performance was very similar for the monocular and training condition, in both pre-tests ( $M_{\text{monoc}} = 1.139$ ,  $SD_{\text{monoc}} = 0.447$ ) and post-tests ( $M_{\text{monoc}} = 2.341$ ,  $SD_{\text{monoc}} = 0.664$ ).

### Filter switched

As with the monocular condition, performance was similar for filter-switched and trained conditions, in both the pre-test ( $M_{\text{Filter-switched}} = 1.356$ ,  $SD_{\text{Filter-switched}} = 0.426$ ) and post-test ( $M_{\text{Filter-switched}} = 2.434$ ,  $SD_{\text{Filter-switched}} = 0.512$ ). Comparisons of performance

between the trained stimulus and the filter-switched conditions showed a main effect of time, with performance significantly improving from pre-test to post-test,  $F(1,52) = 73.821, P < .001, \eta_p^2 = 0.590$ . There was no significant main effect of condition,  $F(1,52) = 0.002, P = .963, \eta_p^2 < 0.001$ , or interaction between condition and time,  $F(1,52) = 0.275, P = .602, \eta_p^2 < 0.001$ .

### 1/f noise

The 1/f noise was a less effective mask than contrast-reversed filtered information; participants had relatively high  $d'$  for both pre-tests ( $M_{1/f\text{ Noise Pre}} = 2.305, SD_{1/f\text{ Noise Pre}} = 0.740$ ), and post-tests ( $M_{1/f\text{ Noise Post}} = 2.925, SD_{1/f\text{ Noise Post}} = 0.656$ ) in the 1/f noise condition. We did observe some transfer of learning (Figure 6), with a significant improvement in performance between pre-tests and post-tests,  $F(1,25) = 10.118, P = 0.004, \eta_p^2 = 0.290$ .

Comparisons of performance between the trained stimulus and the 1/f noise condition revealed a significant difference between the two conditions,  $F(1,55) = 25.386, P < .001, \eta_p^2 = 0.32$ , and a significant difference in performance from pre-test to post-test,  $F(1,55) = 42.051, P < .001, \eta_p^2 = 0.43$ . There was also a significant interaction between condition and time: the 1/f and trained stimulus conditions showed different amounts of learning from pre-test to post-test,  $F(1,55) = 4.463, P = .039, \eta_p^2 = 0.08$ .

The increase in  $d'$  scores, from pre-test to post-test, was significantly larger for the trained ( $M_{\text{Training}} = 1.218, SE_{\text{Training}} = 0.214$ ) than for the 1/f noise condition ( $M_{1/f\text{ Noise}} = 0.620, SE_{1/f\text{ Noise}} = 0.121$ ). The post-test performance in the training condition was only slightly worse than performance in the 1/f noise condition by the end of training ( $M_{\text{Training}} = 2.510, M_{1/f\text{ Noise Post}} = 2.925$ ).

## Discussion

Our goal was to examine whether and how sighted participants might learn to use visual input that roughly mimics the distortions caused by simultaneous on- and off-cell stimulation elicited by electronic and optogenetic sight restoration technologies.

It is impossible to replicate simultaneous stimulation of both on- and off-cell pathways; there is no real-world visual stimulus that elicits such a response, because in natural vision responses to visual stimuli in on-cells are always accompanied by the suppression of off-cells. However, the methods described here represent an analogous disruption of population responses. By combining conflicting spatial frequency and orientation information across the two eyes, we likely produced unnatural on- and off-cell input to cortex.

Our goal was to examine the ability of the visual system to learn to compensate for abnormal neuronal population responses, rather than simulate prosthetic vision per se. We therefore used high-resolution stimuli, as opposed to using the pixelated images in many studies of prosthetic vision (Chen, Suaning, Morley, & Lovell, 2009; van Rheede, Kennard, & Hicks, 2010; Wang, Marek, Steffen, & Pollmann, 2021; Wang, Sharifian, Napp, Nath, & Pollmann, 2018). Previous studies of prosthetic performance have focused on distorting the input through blurring or pixelation, in which a great deal of image information is lost. One limitation/difference in this study is that, in contrast with these other methods, our filtering procedure, despite initially making the stimuli perceptually incomprehensible to a naïve observer, was mathematically lossless.

Our filtering method was surprisingly effective at disrupting how perceptually recognizable our stimuli were before training. In a variety of other studies, visual performance has been shown to be robust to adding noise or removing information through low- or high-pass filtering or pixelation (Dagnelie et al., 2007; Kwon & Legge, 2011; Norman, Beers, Holmin, & Boswell, 2009). Indeed, 1/f noise was much less effective as a mask as compared with the contrast reversed image, despite containing similar contrast as a function of spatial frequency, and no image content. It seems likely that the surprising effectiveness of our filtering is due to two related factors. First, our filters produce population responses that are very unlike the natural population code. Second, these population responses no longer have the statistical properties of responses to natural scenes, which typically vary relatively smoothly as a function of spatial, orientation and spatial frequency tuning, except at the borders of objects (Field, Hayes, & Hess, 1993; Geisler, Perry, Super, & Gallogly, 2001).

### Sighted participants versus patients with a prosthesis

One limitation of this study is that there are, of course, major differences between our training protocol in sighted participants and the experience of patients with a prosthesis. One major difference is that patients with a prosthesis have access to distorted information for much more than 1 h/day. However, it is worth noting that current Argus II retinal implant patients, by choice, report using their implant for only a couple of hours per day (Erickson-davis & Korzybska, 2020). The reason for this is unclear, but it seems plausible that the cognitive effort of decoding distorted and pixelated input is a factor.

A second difference is that for patients with a prosthesis the alternative to distorted input is no input,

whereas our virtual patients spend most of their day with normal vision. This factor is likely to have limited plasticity in our study in two ways. First, deprivation (in the case of a prosthesis user) has been shown to have dramatic effects on neurotransmitters associated with both in both responsiveness and plasticity (Park & Fine, 2020). Second, daily alternation with normal visual input may impair adaptation to distorted input (in the case of our sighted participants). In macaques, training to detect electrical stimulation of the cortex causes a large, reversible, retinotopically localized impairment of thresholds for detecting visual stimuli. Retraining on visual detection restores normal light thresholds, but at the cost of increased thresholds for detecting microstimulation. These results naturally raise the concern that optimized decoding for electrical and light stimulation cannot simultaneously coexist within a local cortical region (Ni & Maunsell, 2010). However, the macaques were not trained under conditions that would be designed to promote generalization across the two types of input, and macaques are frustratingly notorious for failing to show generalization of learning under conditions where humans generalize effortlessly. Work done with prisms (Panico, Rossetti, & Trojano, 2020), colored lenses (Engel, Wilkins, Mand, Helwig, & Allen, 2016), and selective attenuation of certain orientations (Bao, Fast, Mesik, & Engel, 2013; Haak, Fast, Bao, Lee, & Engel, 2014) suggests that humans are very capable of switching between perceptual modes, and of course any wearer of corrective lenses is similarly used to rapidly switching between modes of perceptual distortion.

## Fast versus slow learning

Over the 20 sessions that included the trained stimulus,  $d'$  increased by more than 200%. The most rapid learning occurred early, with slower learning after the first session. Many other studies of visual (Fahle, Edelman, & Poggio, 1995) and auditory (Hawkey, Amitay, & Moore, 2004; Wright & Fitzgerald, 2001) perceptual learning similarly show an initial rapid phase of learning, followed by slower improvement (Karni & Sagi, 1993). The performance in the first session likely represents learning specific task demands.

The performance of our participants during the slower phase of learning was very comparable with the learning rates for low-level properties, such as perceptual judgments of spatial frequency or direction of motion (for a review see Fine & Jacobs, 2002) and is thought to be characteristic of learning that occurs relatively early in the visual pathway (Karni & Sagi, 1993). It is possible that a more engaging task (gamification) would result in faster learning during this slower phase of improvement (Achtman, Green, & Bavelier, 2008; Green & Bavelier, 2010, 2012).

## Monocular versus binocular performance

One limitation of our study was that we generated abnormal population responses using conflicting binocular input, whereas prosthetic vision is (currently) monocular. Thus, for our stimuli, binocular rivalry and/or suppression may have affected learning.

Participants performed very similarly in the monocular and binocular conditions, both before and after training, suggesting that most of the improvement in performance with training was not due to participants learning to suppress information from one eye. The fact that there was no discernable decrease in performance when the filters were switched across eyes shows that the majority of the learning that we observed was not eye specific.

There are at least two explanations for these results. One possibility was that the signal to noise available to the observer was identical across binocular and monocular conditions. We used our simple model to estimate the relative signal to noise available in the cortical response to binocular versus monocular presentations. We defined the Euclidian norm of the difference between the perceptual template and the noisy cortical response to the filtered target as the signal. We defined the Euclidian norm of the difference between the perceptual template and the noisy cortical response to the filtered distractor object as the noise. Figure 7 shows simulated histograms representing these distributions for both binocular and monocular simulations. The  $d'$  values were consistently, but only very slightly (<1%) larger for the binocular condition. Thus, for all practical purposes, the signal-to-noise ratio can be considered identical across binocular and monocular conditions.

One possible explanation for similar monocular and binocular performance is that participants were equally efficient at performing the task under both conditions, and all binocular training transferred to the monocular task. A second possibility is that participants suppressed information from one of the two eyes. This suppression may have been global and consistent (e.g., always the right eye) or suppression may have alternated across eyes (either across trials or even within a single trial), and/or have been piecemeal across space. Participants generally perceived a single, coherent image, with no reports of shimmer or luster as would be expected if suppression was partial and/or alternating temporally across eyes. However, these characteristic qualia of incomplete or alternating suppression may not have been particularly noticeable to our participants given our very peculiar stimuli.

## 1/f noise

Performance was initially much better in the 1/f noise condition. As described in Figure 2, in the training

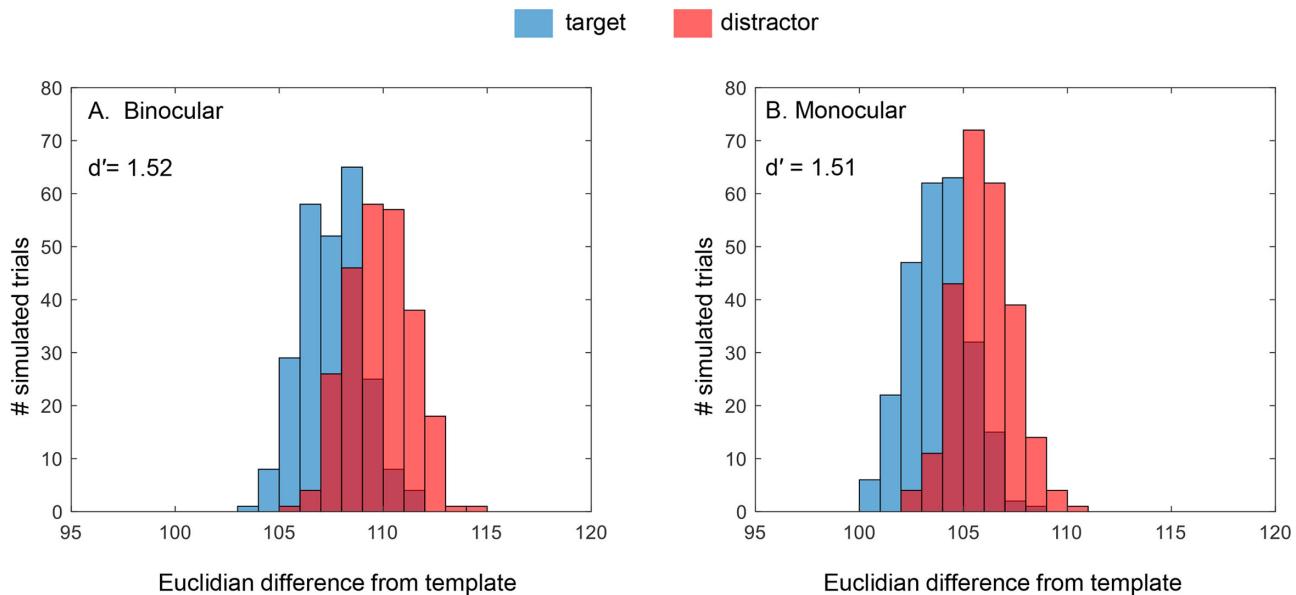


Figure 7. Example histograms showing the Euclidian norm of the difference between the noise-free cortical response to a target object (POT) and noisy cortical responses to the target or a distractor (CLOCK). There were 1,000 trials simulated for each condition. Note that, unlike most signal and noise representations, a small Euclidian norm represents good performance, where the cortical response is similar to the perceptual template.

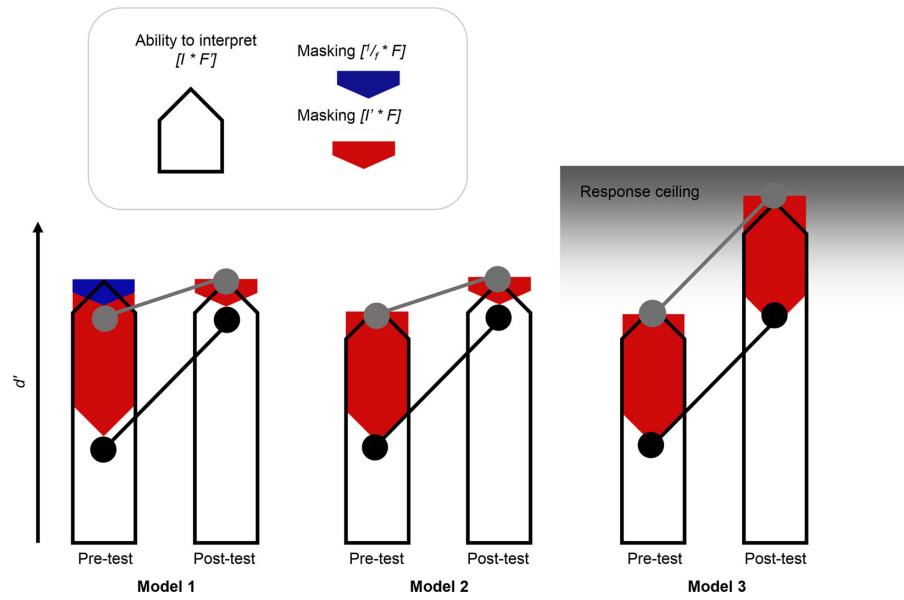


Figure 8. Schematic of possible sources of learning.

condition, in a single eye, participants received  $[I * F'] + [I' * F]$ , whereas in the 1/f noise condition  $I'$  was replaced with 1/f noise, such that the monocular input was  $[I * F'] + [1/f * F]$ . The simplest explanation for the better performance in the 1/f noise condition is that  $[I' * F]$  served as a mask rather than as additional information, and that  $[1/f * F]$  was less effective as a mask, resulting in better performance.

We saw significant (but not complete) transfer of learning to the 1/f noise condition. As schematized,

there are two possible sources of this transfer of learning. One possibility is that the learned ability to suppress  $[I' * F]$  masking information in the training condition, transferred to suppressing the  $[1/f * F]$  in the 1/f noise condition (model 1, Figure 8). A second possibility is that this transfer of learning represented an improved ability to recognize objects that had been passed through our Fourier filters -  $[I * F']$  (models 2 and 3) (Figure 8).

We did not see a full transfer of learning to the 1/f noise condition. One possibility is that this difference represents a lack of (or partial) transfer between learning to suppress  $[I^* F]$  and  $[1/f^* F]$  masking. However, it is also very possible that performance in the 1/f noise condition during the post-test (and therefore the amount of learning transfer) was limited by a ceiling effect ( $M_{1/f \text{ Noise Post}} = 2.925$ ) (Figure 6). We might well have seen greater transfer of learning with a more effective 1/f mask that produced similar initial performance as our training condition.

## Models of learning

Three possible learning models are schematized in Figure 8. These models assume three possible sources of learning: (1) improvement in the ability to interpret  $[I^* F]$  (*empty bars*), (2) learning to discount masking by  $[1/f^* F]$  (*blue bars*), and (3) learning to discount masking by  $[I' * F]$  (*red bars*). All models can predict pre-training and post-training performance across all conditions. For simplicity we only show predicted performance in the training condition (black symbols and lines) and the 1/f noise condition (gray symbols and lines).

In model 1, learning is entirely due to the participants learning to discount within-eye masking information. The  $[1/f^* F]$  mask results in a small amount of masking, which disappears after training. The  $[I' * F]$  mask has a larger impact on performance before training, but effects are considerably decreased with training. In model 2, 1/f noise is an entirely ineffective mask. Some improvement in performance comes from participants getting better at interpreting the  $[I^* F]$  stimulus, and the rest comes from learning to discount  $[I' * F]$  masking. Models 1 and 2 assume that performance improvements in the training condition were primarily the result of learning to discount within-eye masking information. Model 3 assumes that the limited amount of learning in the 1/f noise condition was due to a response ceiling (participants were performing at approximately 95% accuracy by the end of training, which is close to a response ceiling, but was unlikely to be the actual ceiling, given that, by the end of the study, these were highly trained observers who would be expected to have a response ceiling of 97%–98%). According to this model, performance improvements could be entirely driven by improvement in the ability to interpret  $[I^* F]$ . Of course, an intermediate model that falls between these three learning scenarios is equally plausible.

Although distinguishing between these models is obviously an important next step, it is important to note that learning to decode the input provided by prosthetic vision will require both learning to rely more heavily

on neurons that provide interpretable information, and learning to discount neurons that do not.

## Conclusions

These results suggest that it may be possible for patients to adapt to the unnatural on- and off-cell population responses produced by electronic and optogenetic sight recovery technologies. Participants were able to gradually improve in their ability to interpret the cortical input produced by unnatural early on- and off-cell population responses.

The previous literature on perceptual learning and plasticity has mainly focused on two frameworks. The first examines how individuals learn to refine existing perceptual templates by identifying or discriminating a particular set of stimuli or tasks (e.g., the direction of a field of moving dots, or identifying an object in noise; Dosher & Lu, 1998; Fine & Jacobs, 2002). The second examines experiential (i.e., naturalistic viewing conditions) adaptation to sensory loss, for example, within a region of the visual field (Augath et al., 2005; Baseler et al., 2002, 2011; Darian-Smith & Gilbert, 1994; Hiroshi et al., 2015; Masuda, Dumoulin, Nakadomari, & Wandell, 2008), within one eye (Lunghi, Berchicci, Morrone, & Di Russo, 2015; Lunghi, Burr, & Morrone, 2011), or by removing orientation or spatial frequency information (Georges & Sullivan, 1975; Haak et al., 2014; Webster, Georges, & Webster, 2002; Zhang, Bao, Kwon, He, & Engel, 2009). This study frames the role of plasticity in a novel way: is it possible to reconfigure the fundamental building blocks of visual perception in adults? This is a central question both because of its translational importance, and because it examines the adult-analogue of processes that are fundamental to early visual development.

*Keywords:* *vision restoration, perceptual learning*

## Acknowledgments

Research reported in this publication was supported by the National Eye Institute of the National Institutes of Health under Award Number F31EY031164. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Commercial relationships: none.

Corresponding author: Rebecca B. Esquenazi.

Email: resq@uw.edu.

Address: Department of Psychology, University of Washington, USA.

## References

- Achtman, R. L., Green, C. S., & Bavelier, D. (2008). Video games as a tool to train visual skills. *Restorative Neurology and Neuroscience*, 26(4–5), 435–446.
- Ahuja, A. K., Yeoh, J., Dorn, J. D., Caspi, A., Wuyyuru, V., & McMahon, M. J., & Argus II Study Group. (2013). Factors affecting perceptual threshold in Argus II retinal prosthesis subjects. *Translational Vision Science & Technology*, 2(4), 1, <https://doi.org/10.1167/tvst.2.4.1>.
- Augath, M., Smirnakis, S. M., Logothetis, N. K., Schüz, A., Brewer, A. A., Schmid, M. C., ... Wandell, B. A. (2005). Lack of long-term cortical reorganization after macaque retinal lesions. *Nature*, 435(7040), 300–307, <https://doi.org/10.1038/nature03495>.
- Ayton, L. N., Blamey, P. J., Guymer, R. H., & Luu, C. D. (2014). First-in-human trial of a novel suprachoroidal retinal prosthesis. *PLoS One*, 9(12), 1–26, <https://doi.org/10.1371/journal.pone.0115239>.
- Bach, M. (1996). FrACT-Landolt-vision. *Optometry and Vision Science*, 73(1), 49–53.
- Bach, M. (2007). The Freiburg Visual Acuity Test-Variability unchanged by post-hoc re-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 245, 965–971, <https://doi.org/10.1007/s00417-006-0474-4>.
- Bamann, C., Nagel, G., & Bamberg, E. (2010). Microbial rhodopsins in the spotlight. *Current Opinion in Neurobiology*, 20(5), 610–616, <https://doi.org/10.1016/j.conb.2010.07.003>.
- Bao, M., Fast, E., Mesik, J., & Engel, S. (2013). Distinct mechanisms control contrast adaptation over different timescales. *Journal of Vision*, 13(10), 1–11, <https://doi.org/10.1167/13.10.14>.
- Barrett, B. T., Panesar, G. K., Scally, A. J., & Pacey, I. E. (2012). A limited role for suppression in the central field of individuals with strabismic amblyopia. *PLoS One*, 7(5), 1–12, <https://doi.org/10.1371/journal.pone.0036611>.
- Baseler, H. A., Brewer, A. A., Sharpe, L. T., Morland, A. B., Jaägle, H., & Wandell, B. A. (2002). Reorganization of human cortical maps caused by inherited photoreceptor abnormalities. *Nature Neuroscience*, 5(4), 364–370, <https://doi.org/10.1038/nn817>.
- Baseler, H. A., Gouws, A., Haak, K. V., Racey, C., Crossland, M. D., Tufail, A., ... Morland, A. B. (2011). Large-scale remapping of visual cortex is absent in adult humans with macular degeneration. *Nature Neuroscience*, 14(5), 649–657, <https://doi.org/10.1038/nn.2793>.
- Beauchamp, M. S., Oswalt, D., Sun, P., Foster, B. L., Magnotti, J. F., Niketeghad, S., ... Yoshor, D. (2020). Dynamic stimulation of visual cortex produces form vision in sighted and blind humans. *Cell*, 181(4), 774–783.e5, <https://doi.org/10.1016/j.cell.2020.04.033>.
- Berry, M. H., Holt, A., Levitz, J., Broichhagen, J., Gaub, B. M., Visel, M., ... Isacoff, E. Y. (2017). Restoration of patterned vision with an engineered photoactivatable G protein-coupled receptor. *Nature Communications*, 8(1), 1–12, <https://doi.org/10.1038/s41467-017-01990-7>.
- Beyeler, M., Boynton, G. M., Fine, I., & Rokem, A. (2017). pulse2percept: A Python-based simulation framework for bionic vision, (Scipy). *BioRxiv*, 81–88.
- Beyeler, M., Nanduri, D., Weiland, J. D., Rokem, A., Boynton, G. M., & Fine, I. (2019). A model of ganglion axon pathways accounts for percepts elicited by retinal implants. *Scientific Reports*, 9(1), 1–16, <https://doi.org/10.1038/s41598-019-45416-4>.
- Beyeler, M., Rokem, A., Boynton, G. M., & Fine, I. (2017). Learning to see again: Biological constraints on cortical plasticity and the implications for sight restoration technologies. *Journal of Neural Engineering*, 14(5), <https://doi.org/10.1088/1741-2552/aa795e>.
- Bijveld, M. M. C., Florijn, R. J., Bergen, A. A. B., Van Den Born, L. I., Kamermans, M., Prick, L., ... Van Genderen, M. M. (2013). Genotype and phenotype of 101 Dutch patients with congenital stationary night blindness. *Ophthalmology*, 120(10), 2072–2081, <https://doi.org/10.1016/j.ophtha.2013.03.002>.
- Bosking, W. H., Beauchamp, M. S., & Yoshor, D. (2017). Electrical stimulation of visual cortex: Relevance for the development of visual cortical prosthetics. *Annual Review of Vision Science*, 3, 141–166, <https://doi.org/10.1146/annurev-vision-111815-114525>.
- Bosking, W. H., Sun, P., Ozker, M., Pei, X., Foster, B. L., Beauchamp, M. S., ... Yoshor, D. (2017). Saturation in phosphene size with increasing current levels delivered to human visual cortex. *Journal of Neuroscience*, 37(30), 7188–7197, <https://doi.org/10.1523/JNEUROSCI.2896-16.2017>.
- Busskamp, V., Picaud, S., Sahel, J. A., & Roska, B. (2012). Optogenetic therapy for retinitis pigmentosa. *Gene Therapy*, 19(2), 1–7, <https://doi.org/10.1038/gt.2011.155>.
- Chen, S. C., Suaning, G. J., Morley, J. W., & Lovell, N. H. (2009). Simulating prosthetic vision: II. Measuring functional capacity. *Vision Research*, 49(19), 2329–2343, <https://doi.org/10.1016/j.visres.2009.07.003>.
- Chen, X., Wang, F., Fernandez, E., & Roelfsema, P. R. (2020). Shape perception via a high-

- channel-count neuroprosthesis in monkey visual cortex. *Science*, 370(6521), 1191–1196, <https://doi.org/10.1126/science.abd7435>.
- Cho, G. Y., Bolo, K., Park, K. S., Sengillo, J. D., & Tsang, S. H. (2019). Attenuation of inherited and acquired retinal degeneration progression with gene-based techniques. *Molecular Diagnosis and Therapy*, 23(1), 113–120, <https://doi.org/10.1007/s40291-018-0377-1>.
- Cibis, G. W., & Fitzgerald, K. M. (2001). The negative ERG is not synonymous with nightblindness. *Transactions of the American Ophthalmological Society*, 99, 171–176.
- Cuevas, E., Parmar, P., & Sowden, J. C. (2019). Restoring vision using stem cells and transplantation. *Advances in Experimental Medicine and Biology*, 1185, 563–567.
- da Cruz, L., Dorn, J. D., Humayun, M. S., Dagnelie, G., Handa, J., Barale, P. O., ... Greenberg, R. J. (2016). Five-year safety and performance results from the Argus II retinal prosthesis system clinical trial. *Ophthalmology*, 123(10), 2248–2254, <https://doi.org/10.1016/j.ophtha.2016.06.049>.
- Dagnelie, G., Keane, P., Narla, V., Yang, L., Weiland, J., & Humayun, M. (2007). Real and virtual mobility performance in simulated prosthetic vision. *Journal of Neural Engineering*, 4(1), S92–S101, <https://doi.org/10.1088/1741-2560/4/1/S11>.
- Darian-Smith, C., & Gilbert, C. D. (1994). Axonal sprouting accompanies functional reorganization in adult cat striate cortex. *Nature*, 368(6473), 737–740, <https://doi.org/10.1038/368737a0>.
- de Balthasar, C., Patel, S., Roy, A., Freda, R., Greenwald, S., Horsager, A., ... Fine, I. (2008). Factors affecting perceptual thresholds in epiretinal prostheses. *Investigative Ophthalmology and Visual Science*, 49(6), 2303–2314, <https://doi.org/10.1167/iovs.07-0696>.
- Demb, J. B., Haarsma, L., Freed, M. A., & Sterling, P. (1999). Functional circuitry of the retinal ganglion cell's nonlinear receptive field. *Journal of Neuroscience*, 19(22), 9756–9767, <https://doi.org/10.1523/jneurosci.19-22-09756.1999>.
- Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23), 13988–13993, <https://doi.org/10.1073/pnas.95.23.13988>.
- Engel, S. A., Wilkins, A. J., Mand, S., Helwig, N. E., & Allen, P. M. (2016). Habitual wearers of colored lenses adapt more rapidly to the color changes the lenses produce. *Vision Research*, 125, 41–48, <https://doi.org/10.1016/j.visres.2016.05.003>.
- Erickson-davis, C., & Korzybska, H. (2020). What do blind people “see” with retinal prostheses? Observations and qualitative reports of epiretinal implant users. *PloS One*, 16(2), e0229189, <https://doi.org/10.1101/2020.02.03.932905>.
- Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in visual hyperacuity. *Vision Research*, 35(21), 3003–3013, <https://doi.org/10.1126/science.1589770>.
- Fallon, J. B., Irvine, D. R. F., & Shepherd, R. K. (2008). Cochlear implants and brain plasticity. *Hearing Research*, 238(1–2), 110–117.
- Ferlauto, L., Airaghi Leccardi, M. J. I., Chenais, N. A. L., Gilliéron, S. C. A., Vagni, P., Bevilacqua, M., ... Ghezzi, D. (2018). Design and validation of a foldable and photovoltaic wide-field epiretinal prosthesis. *Nature Communications*, 9(1), 1–15, <https://doi.org/10.1038/s41467-018-03386-7>.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field.” *Vision Research*, 33(2), 173–193, [https://doi.org/10.1016/0042-6989\(93\)90156-Q](https://doi.org/10.1016/0042-6989(93)90156-Q).
- Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision*, 2, 190–203, <https://doi.org/10.1167/2.2.5>.
- Fine, I., Cepko, C., & Landy, M. (2015). Vision research special issue: Sight restoration: Prosthetics, optogenetics and gene therapy. *Vision Research*, 111, 115–123, <https://doi.org/10.1016/j.visres.2015.04.012>.
- Fine, I., & Boynton, G. M. (2015). Pulse trains to percepts: The challenge of creating a perceptually intelligible world with sight recovery technologies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1677), 20140208.
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, 14(9), 1195–1204, <https://doi.org/10.1038/nn.2889>.
- Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, 6(11), 1216–1223, <https://doi.org/10.1038/nn1141>.
- Fujikado, T., Kamei, M., Sakaguchi, H., Kanda, H., Endo, T., Hirota, M., ... Nishida, K. (2016). One-year outcome of 49-channel suprachoroidal-transretinal stimulation prosthesis in patients with advanced retinitis pigmentosa. *Investigative Ophthalmology and Visual Science*, 57(14), 6147–6157, <https://doi.org/10.1167/iovs.16-20367>.
- Garita-Hernandez, M., Lampič, M., Chaffiol, A., Guibal, L., Routet, F., Santos-Ferreira, T., ...

- Duebel, J. (2019). Restoration of visual function by transplantation of optogenetically engineered photoreceptors. *Nature Communications*, 10(1), 1–13, <https://doi.org/10.1038/s41467-019-12330-2>.
- Gasparini, S. J., Llonch, S., Borsch, O., & Ader, M. (2019). Transplantation of photoreceptors into the degenerative retina: Current state and future perspectives. *Progress in Retinal and Eye Research*, 69, 1–37, <https://doi.org/10.1016/j.preteyeres.2018.11.001>.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41(6), 711–724, [https://doi.org/10.1016/S0042-6989\(00\)00277-7](https://doi.org/10.1016/S0042-6989(00)00277-7).
- Genovesi-Ebert, F., Allegrini, L., di Bartolo, E., Cinelli, L., Belting, C., Barca, F., ... Rizzo, S. (2014). The Argus II retinal prosthesis: 12-Month outcomes from a single-study center. *American Journal of Ophthalmology*, 157(6), 1282–1290, <https://doi.org/10.1016/j.ajo.2014.02.039>.
- Georgeson, B. Y. M. A., & Sullivan, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *Journal of Physiology*, 252, 627–656.
- Ghezzi, D. (2015). Retinal prostheses: Progress towards the next generation implants. *Frontiers in Neuroscience*, 9, 1–6, <https://doi.org/10.3389/fnins.2015.00290>.
- Ghose, G. M., Yang, T., & Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology*, 87(4), 1867–1888, <https://doi.org/10.1152/jn.00690.2001>.
- Green, C. S., & Bavelier, D. (2010). Training induced learning. *Brain*, 133(4), 692–701, <https://doi.org/10.1037/a0014345.Exercising>.
- Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology*, 22(6), R197–R206, <https://doi.org/10.1016/j.cub.2012.02.012>.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Haak, K. V., & Beckmann, C. F. (2019). Plasticity versus stability across the human cortical visual connectome. *Nature Communications*, 10(1), 1–8, <https://doi.org/10.1038/s41467-019-11113-z>.
- Haak, K. V., Fast, E., Bao, M., Lee, M., & Engel, S. A. (2014). Four days of visual contrast deprivation reveals limits of neuronal adaptation. *Current Biology*, 24(21), 2575–2579, <https://doi.org/10.1016/j.cub.2014.09.027>.
- Hawkey, D. J. C., Amitay, S., & Moore, D. R. (2004). Early and rapid perceptual learning. *Nature Neuroscience*, 7(10), 1055–1056, <https://doi.org/10.1038/nn1315>.
- Hinds, O. P., Rajendran, N., Polimeni, J. R., Augustinack, J. C., Wiggins, G., Wald, L. L., ... Fischl, B. (2008). Accurate prediction of V1 location from cortical folds in a surface coordinate system. *NeuroImage*, 39(4), 1585–1599, <https://doi.org/10.1016/j.neuroimage.2007.10.033>.
- Hiroshi, A., McManus, J. N. J., Ramalingam, N., Li, W., Marik, S. A., Meyer zum Alten Borgloh, S., ... Gilbert, C. D. (2015). Adult cortical plasticity studied with chronically implanted electrode arrays. *Journal of Neuroscience*, 35(6), 2778–2790, <https://doi.org/10.1523/JNEUROSCI.3579-14.2015>.
- Hochstein, S., & Shapley, R. M. (1976). Quantitative analysis of retinal ganglion cell classifications. *Journal of Physiology*, 262, 237–264.
- Hornig, R. (2017). *Artificial vision: A clinical guide*. New York: Springer International Publishing.
- Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, 365(6443), 250–252, <https://doi.org/10.1038/365250a0>.
- Keliris, G. A., Li, Q., Papanikolaou, A., Logothetis, N. K., & Smirnakis, S. M. (2019). Estimating average single-neuron visual receptive field sizes by fMRI. *Proceedings of the National Academy of Sciences of the United States of America*, 116(13), 6425–6434, <https://doi.org/10.1073/pnas.1809612116>.
- Kramer, R. H., Mourot, A., & Adesnik, H. (2013). Optogenetic pharmacology for control of native neuronal signaling proteins. *Nature Neuroscience*, 16(7), 816–823, <https://doi.org/10.1038/nn.3424>.
- Kwon, M. Y., & Legge, G. E. (2011). Spatial-frequency cutoff requirements for pattern recognition in central and peripheral vision. *Vision Research*, 51(18), 1995–2007, <https://doi.org/10.1016/j.visres.2011.06.020>.
- Liu, Y., Fattah, N., & Degenaar, P. (2020). Newcastle visual prosthesis implantable control unit. In *27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. Glasgow, Scotland, November 23–25 (pp. 1–4), <https://doi.org/10.1109/icecs49266.2020.9294853>.
- Lorach, H., Goetz, G., Smith, R., Lei, X., Mandel, Y., Kamins, T., ... Palanker, D. (2015). Photovoltaic restoration of sight with high visual acuity. *Nature Medicine*, 21(5), 476–482, <https://doi.org/10.1038/nm.3851>.
- Lunghi, C., Berchicci, M., Morrone, M. C., & Di Russo, F. (2015). Short-term monocular deprivation alters early components of visual evoked potentials. *Journal of Physiology*, 593(19), 4361–4372, <https://doi.org/10.1113/JP270950>.

- Lunghi, C., Burr, D. C., & Morrone, C. (2011). Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex. *Current Biology*, 21(14), R538–R539, <https://doi.org/10.1016/j.cub.2011.06.004>.
- Masuda, Y., Dumoulin, S. O., Nakadomari, S., & Wandell, B. A. (2008). V1 projection zone signals in human macular degeneration depend on task, not stimulus. *Cerebral Cortex*, 18(11), 2483–2493, <https://doi.org/10.1093/cercor/bhm256>.
- Mata, M. L., & Ringach, D. L. (2005). Spatial overlap of ON and OFF subregions and its relation to response modulation ratio in macaque primary visual cortex. *Journal of Neurophysiology*, 93(2), 919–928, <https://doi.org/10.1152/jn.00668.2004>.
- Mills, J. O., Jalil, A., & Stanga, P. E. (2017). Electronic retinal implants and artificial vision: Journey and present. *Eye*, 31(10), 1383–1398, <https://doi.org/10.1038/eye.2017.65>.
- Morillas, C. A., Romero, S. F., Martínez, A., Pelayo, F. J., Ros, E., & Fernández, E. (2007). A design framework to model retinas. *BioSystems*, 87(2–3), 156–163, <https://doi.org/10.1016/j.biosystems.2006.09.009>.
- Murphrey, D. K., Maunsell, J. H. R., Beauchamp, M. S., & Yoshor, D. (2009). Perceiving electrical stimulation of identified human visual areas. *Proceedings of the National Academy of Sciences of the United States of America*, 106(13), 5389–5393, <https://doi.org/10.1073/pnas.0804998106>.
- Ni, A. M., & Maunsell, J. H. R. (2010). Microstimulation reveals limits in detecting different signals from a local cortical region. *Current Biology*, 20(9), 824–828, <https://doi.org/10.1016/j.cub.2010.02.065>.
- Norman, J. F., Beers, A. M., Holmin, J. S., & Boswell, A. M. (2009). Effective 3-D shape discrimination survives retinal blur. *Attention, Perception, & Psychophysics*, 72(6), 1569–1575, <https://doi.org/10.3758/APP>.
- Öhlschläger, S., & Võ, M. L. H. (2017). SCEGRAM: An image database for semantic and syntactic inconsistencies in scenes. *Behavior Research Methods*, 49(5), 1780–1791, <https://doi.org/10.3758/s13428-016-0820-3>.
- Palanker, D., Le Mer, Y., Mohand-Said, S., Muqit, M., & Sahel, J. A. (2020). Photovoltaic restoration of central vision in atrophic age-related macular degeneration. *Ophthalmology*, 127(8), 1087–1104, <https://doi.org/10.1016/j.ophtha.2020.02.024>. Photovoltaic.
- Panico, F., Rossetti, Y., & Trojano, L. (2020). On the mechanisms underlying Prism Adaptation: A review of neuro-imaging and neuro-stimulation studies. *Cortex*, 123, 57–71, <https://doi.org/10.1016/j.cortex.2019.10.003>.
- Park, W. J., & Fine, I. (2020). New insights into cortical development and plasticity: from molecules to behavior. *Current Opinion in Physiology*, 16, 50–60, <https://doi.org/10.1016/j.cophys.2020.06.004>.
- Polimeni, J. R., Balasubramanian, M., & Schwartz, E. L. (2006). Multi-area visuotopic map complexes in macaque striate and extra-striate cortex. *Vision Research*, 46(20), 3336–3359, <https://doi.org/10.1016/j.visres.2006.03.006>.
- Polosukhina, A., Litt, J., Tochitsky, I., Nemargut, J., Sychev, Y., De Kouchkovsky, I., ... Kramer, R. H. (2012). Photochemical restoration of visual responses in blind mice. *Neuron*, 75(2), 271–282, <https://doi.org/10.1016/j.neuron.2012.05.022>.
- Rizzo, S., Belting, C., Cinelli, L., Allegrini, L., Genovesi-Ebert, F., Barca, F., ... Di Bartolo, E. (2014). The Argus II retinal prosthesis: 12-month outcomes from a single-study center. *American Journal of Ophthalmology*, 157(6), 1282–1290, <https://doi.org/10.1016/j.ajo.2014.02.039>.
- Royer, A. S., & Schwartz, E. L. (1990). Cat and monkey cortical columnar patterns modeled by bandpass-filtered 2D white noise. *Biological Cybernetics*, 62(5), 381–391.
- Roska, B., & Sahel, J.-A. (2018). Restoring vision. *Nature*, 557(7705), 359–367, <https://doi.org/10.1038/s41586-018-0076-4>.
- Sahel, J. A., Boulanger-Scemama, E., Pagot, C., Arleo, A., Galluppi, F., Martel, J. N., ... Roska, B. (2021). Partial recovery of visual function in a blind patient after optogenetic therapy. *Nature Medicine*, 27(7), 1223–1229, <https://doi.org/10.1038/s41591-021-01351-4>.
- Saunders, A. L., Williams, C. E., Heriot, W., Briggs, R., Yeoh, J., Nayagam, D. A., ... Allen, P. J. (2014). Development of a surgical procedure for implantation of a prototype suprachoroidal retinal prosthesis. *Clinical and Experimental Ophthalmology*, 42(7), 665–674, <https://doi.org/10.1111/ceo.12287>.
- Schmidt, E. M., Bak, M. J., Hambrecht, F. T., Kufta, C. V., Rourke, D. K. O., & Vallabhanath, P. (1996). Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. *Brain*, 119, 507–522.
- Scholl, H. P. N., Strauss, R. W., Singh, M. S., Dalkara, D., Roska, B., Picaud, S., ... Sahel, J. A. (2016). Emerging therapies for inherited retinal degeneration. *Science Translational Medicine*, 8(368), 1–11, <https://doi.org/10.1126/scitranslmed.aaf2838>.

- Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Research*, 20(8), 645–669, [https://doi.org/10.1016/0042-6989\(80\)90090-5](https://doi.org/10.1016/0042-6989(80)90090-5).
- Schwartz, E. L. (1994). Topographic mapping in primate visual cortex: History, anatomy, and computation. In *Visual science and engineering: Models and applications* (pp. 293–360). Boca Raton, FL: CRC Press, <https://doi.org/10.1201/9781466593534-27>.
- Shah, N. P., Madugula, S., Grosberg, L., Mena, G., Tandon, P., Hottowy, P., ... Chichilnisky, E. J. (2019). Optimization of electrical stimulation for a high-fidelity artificial retina. *International IEEE/EMBS Conference on Neural Engineering (NER)*. San Francisco, March 20–23 (pp. 714–718), <https://doi.org/10.1109/NER.2019.8716987>.
- Stingl, K., Bartz-Schmidt, K. U., Besch, D., Chee, C. K., Cottrall, C. L., Gekeler, F., ... Zrenner, E. (2015). Subretinal visual implant Alpha IMS - Clinical trial interim report. *Vision Research*, 111, 149–160, <https://doi.org/10.1016/j.visres.2015.03.001>.
- Stingl, K., Schippert, R., Bartz-Schmidt, K. U., Besch, D., Cottrall, C. L., Edwards, T. L., ... Zrenner, E. (2017). Interim results of a multicenter trial with the new electronic subretinal implant alpha AMS in 15 patients blind from inherited retinal degenerations. *Frontiers in Neuroscience*, 11, 445, <https://doi.org/10.3389/fnins.2017.00445>.
- R Computational Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available: [www.r-project.org/](http://www.r-project.org/). Accessed November 24, 2021.
- Tochitsky, I., Polosukhina, A., Degtyar, V. E., Gallerani, N., Smith, C. M., Friedman, A., ... Kramer, R. H. (2014). Restoring visual function to blind mice with a photoswitch that exploits electrophysiological remodeling of retinal ganglion cells. *Neuron*, 81(4), 800–813, <https://doi.org/10.1016/j.neuron.2014.01.003>.
- Tong, W., Meffin, H., Garrett, D. J., & Ibbotson, M. R. (2020). Stimulation strategies for improving the resolution of retinal prostheses. *Frontiers in Neuroscience*, 14, 262, <https://doi.org/10.3389/fnins.2020.00262>.
- Troyk, P. R. (2017). *Artificial vision*. Cham: Springer.
- van Rheede, J. J., Kennard, C., & Hicks, S. L. (2010). Simulating prosthetic vision: Optimizing the information content of a limited visual display. *Journal of Vision*, 10(14), 1–14, <https://doi.org/10.1167/10.14.1>.
- Wang, L., Marek, N., Steffen, J., & Pollmann, S. (2021). Perceptual learning of object recognition in simulated retinal implant perception—The effect of video training. *Translational Vision Science & Technology*, 10(12), 22, <https://doi.org/10.1167/tvst.10.12.22>.
- Wang, L., Sharifian, F., Napp, J., Nath, C., & Pollmann, S. (2018). Cross-task perceptual learning of object recognition in simulated retinal implant perception. *Journal of Vision*, 18(13), 1–14, <https://doi.org/10.1167/18.13.22>.
- Watson, A. B. (2014). A formula for human retinal ganglion cell receptive field density as a function of visual field location. *Journal of Vision*, 14(7), 1–17, <https://doi.org/10.1167/14.7.15>.
- Webster, M. A., Georgeson, M. A., & Webster, S. M. (2002). Neural adjustments to image blur. *Nature Neuroscience*, 5(9), 839–840, <https://doi.org/10.1038/nn906>.
- Weiland, J. D., Walston, S. T., & Humayun, M. S. (2016). Electrical stimulation of the retina to produce artificial vision. *Annual Review of Vision Science*, 2(1), 273–294, <https://doi.org/10.1146/annurev-vision-111815-114425>.
- Wright, B. A., & Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences of the United States of America*, 98(21), 12307–12312, <https://doi.org/10.1073/pnas.211220498>.
- Zeitz, C., Robson, A. G., & Audo, I. (2015). Congenital stationary night blindness: An analysis and update of genotype-phenotype correlations and pathogenic mechanisms. *Progress in Retinal and Eye Research*, 45(October), 58–110, <https://doi.org/10.1016/j.preteyeres.2014.09.001>.
- Zhang, P., Bao, M., Kwon, M., He, S., & Engel, S. A. (2009). Effects of orientation-specific visual deprivation induced with altered reality. *Current Biology*, 19(22), 1956–1960, <https://doi.org/10.1016/j.cub.2009.10.018>.
- Zrenner, E. (2002). Will retinal implants restore vision? *Science*, 295(5557), 1022–1025, <https://doi.org/10.1126/science.1067996>.