Visual Cortical Plasticity and the Implications for Sight Restoration Technologies

Rebecca B. Esquenazi

A dissertation
submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

University of Washington

2022

Reading Committee:
Ione Fine, Chair
Scott Murray
Andrea Stocco

Program Authorized to Offer Degree:

Psychology

University of Washington

Abstract

Visual Cortical Plasticity and the Implications for Sight Restoration Technologies

Rebecca Esquenazi

Chair of the Supervisory Committee:
Ione Fine, PhD Professor
Psychology

Degenerative eye diseases are a leading cause of blindness in adults. While only limited treatment options for diseases such as age related macular degeneration and retinitis pigmentosa exist, there have been a variety of technologies developed that attempt to restore some visual functioning. These include electronic prostheses (either cortical or retinal), optogenetics, and gene therapy. However, to date, only retinal prosthetic devices have been implanted in patients, with cortical implant and optogenetic clinical trials still under way. These technologies are not likely to completely recreate normal vision, but rather provide essential visual input that can improve everyday functioning in patients.

In the current state of sight restoration technology, these devices cause early on- and off-center retinal cells to fire simultaneously, rather than in a biologically complementary fashion (i.e. when on-cells fire, off-cells in the same location are suppressed). This creates deeply unnatural population responses that propagate from the retina to cortex. The central question of this thesis is whether patients have the potential to access cortical plasticity in adulthood that allows them to decode unnatural cell population responses in early visual processing caused by electronic prosthesis technologies.

In Chapter 1, sighted participants were dichoptically presented with a combination of original and contrast-reversed images that resulted in a similar decoding challenge caused by simultaneous stimulation of on- and off-cells. 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. We show that participants continuously improved in a naturalistic object discrimination task over 20 one-hour sessions using the distorted input. Pre- and post-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

In Chapter 2, we ask whether video game training provides any additional learning benefit than that witnessed previously, and examine whether learning generalizes outside of the trained conditions. While we did not find a large degree of transfer to novel stimuli as a result of video game training, we provide additional insights into mechanisms that explain learning on the task.

Collectively, these studies examine the potential of human participants to learn to decode unnatural neural population responses in adulthood. Learning to decode these distorted stimuli will provide developers of retinal prosthesis technologies with a realistic, empirical expectation of the extent to which patients can learn to make use of their implant.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I am extremely grateful to my supervisor, Dr. Ione Fine and co-advisor Dr. Geoffrey Boynton for their advice, continuous support, and patience during my PhD study. Their immense knowledge and plentiful experience encouraged me to complete my study, and they were always willing to lend a helping hand. I would like to thank all the members in the Vis Cog lab, especially Kimberly Meier, a post-doc who was always willing to explain difficult concepts and lend a helping hand with technical work such as coding. It is their kind help and support that have made my study and life in the UK a wonderful time. Finally, I would like to express my gratitude to my incredible husband who never wavered in his consistent support of my studies. Further, my friends who I consider family provided a support system for me in Seattle, for which I am so grateful. Without their support, I don't believe I could have finished my studies.

VITA

EDUCATION

Psychology (Cognition & Perception) – Ph.D. (December 2022)
University of Washington

General Experimental Psychology – M.A. (May 2018) – *Summa Cum Laude*
Psychology – B.A. (May 2015) – *Cum Laude*
California State University, Northridge (CSUN)

AWARDS & SCHOLARSHIP

- **NIH National Research Service Award** (2020 – 2023) – Grant awarded from the National Eye Institute at NIH
- **Vision Training Grant** (2019-2020) – Funding for one year through the UW Department of Biological Structure
- **Tabachnick and Fidell Multivariate Statistics Award** (2018) – CSUN Psychology Department award for outstanding quantitative achievement
- **Top Scholar Award** (2018) – University of Washington Psychology Department award for summer funding
- **Creative Endeavors Scholarship** (2017) – Funding awarded for graduate thesis project
- **CSUN Data Jam Competition** (2017) – First Place Winner – Best Statistical Analysis in CSUN data science competition
- **Sally Casanova Pre-Doctoral Scholarship** (2017-18) – Fellowship awarded to support doctoral program aspirations
- **First Place Winner** – CSUNPosium (2018) – Prize awarded for best oral research presentation

RESEARCH EXPERIENCE

Publications

Esquenazi, R., Meier, K., Beyeler, M., Boynton, GM., Fine I. (2020). *Learning to see again: perceptual learning of simulated abnormal on- off- cell population responses in sighted individuals.* Journal of Vision.

Invited Research Talks

Esquenazi, R., (2020, January). *Learning to see again: Perceptual learning for sight restoration technologies.* Talk at the 2020 Neural Engineering Connection Conference. Seattle, WA.

National Conference Presentations

Esquenazi, R., Beyeler, M., Boynton, GM., Fine I. *Perceptual learning of simulated abnormal on- off- cell population responses in sighted individuals.* Virtual poster presentation at the Eye and Chip Research Congress.

Esquenazi, R., (2019, September). *Learning to see again: Perceptual learning for sight restoration technologies.* Research talk at the 2019 annual meeting of the Optical Society Association. Washington, DC.

Esquenazi, R., Larranaga, D., Awad, M., Drew, S. A. (2018, November). *Grapheme-color synesthetes and inhibitory differences during an antisaccade task*. Poster presentation at the 2018 annual meeting at the Society for Neuroscience, San Diego, CA.

Awad, J.F., **Esquenazi, R**., Hackney, B.C., Delgado, C.Y., Kangavary, A., Gandhi, V., Drew, S.A. (2018, May). *Observing Effects of Virtual Reality on Visual Accommodation*. Poster presentation at the Association for Psychological Sciences Annual Convention, San Francisco, CA.

Doty, T. A., **Esquenazi, R**., Lundqvist, S., Morales, R., McGinnis, C., Gandhi, V., Drew, S. A. (2018, May). *Analysis of Acute and Reoccurring Virtual Reality Usage on Visual Discomfort*. Poster presentation at the Association for Psychological Sciences Annual Convention, San Francisco, CA.

Esquenazi, R., Mcmanus, R., Choi, P., Duarte, A., Katz G., Ainsworth, A., Rutchick, A. (2018, March). *Technologically Facilitated Remoteness Decreases Negative Emotion in a Killing Task*. Poster session at the 2018 annual meeting at the Society for Personality and Social Psychology, Atlanta, GA.

Esquenazi, R., Mosher, R., Buenrostro, J., Lundqvist, S., Drew, S.A. (2017, November). *The implications of immersive virtual reality usage on the visual system*. Poster session at the 2017 annual meeting at the Society for Neuroscience, Washington, D.C.

Doty, T., **Esquenazi, R**., Larranaga, D., Rutchick, A.M., Drew, S.A. (2017, November). *Synesthesia and creativity: a look on writing and word tasks*. Poster session at the 2017 meeting of the Society for Neuroscience, Washington, D.C.

Esquenazi, R., Larragna, D., Rutchick, A., Drew, S. A. (2017, May). *Synesthetes produce different language in creative writing tasks.* Poster session at the 2017 annual meeting of the Association for Psychological Science, Boston, MA.

Hochman, A. S., Awad J., **Esquenazi, R.**, Inicki, A., Drew, S. A., (2017, May). *Red 8 Green 8: Color Shifts in Synesthesia.* Poster session at the 2017 annual meeting of Vision Sciences Society, St. Pete Beach, FL.

Choi, P., McManus, R., Duarte A., **Esquenazi, R.,** Frausto E., Katz, G., Ainsworth A., Rutchick A., (2017, May). *Recall of an Insect-Killing Experience: Biased Recall and Lingering Negative Affect*. Poster session at the 2017 annual meeting at the Association Psychological Science, Boston, MA.

Hochman, A. S., Awad J., **Esquenazi, R.**, Inicki, A., Drew, S. A., (2017, May). *Red 8 Green 8: Color Shifts in Synesthesia.* Poster session at the 2017 annual meeting of Vision Sciences Society, St. Pete Beach, FL.

Regional & State Conference Presentations

Esquenazi, R. (2018, April). *Yielding Agency to Machines: A measure of comfort with technology.* Oral presentation at the 2018 annual Western Psychological Association, Portland, OR.

Esquenazi, R. (2018, April). *An Investigation into the Disinhibited Feedback Theory of Synesthesia.* Oral presentation at the 22nd Annual CSUN Student Research & Creative Works CSUNposium, Northridge, CA.

Awad. M., Larranaga, D., **Esquenazi, R**., Drew, S. A. (2017, April). *Keeping an eye on it: observing accuracy of self-reported re-reading rates.* Poster at the 2017 annual CSUN Symposium, Northridge, CA

Esquenazi, R., Larragna, D., Rutchick, A., Drew, S. A. (2017, April). *Synesthetes produce different language in creative writing tasks*. Poster session at the 2017 annual CSUN Creative Works Symposium, Northridge, CA

Researcher (2018 – Present)

Vision-Cognition Group, *University of Washington*

- Conducted research focused on the perceptual learning and plasticity in simulated retinal prosthesis patients
- Assisted in coding perceptual learning project in MATLAB using custom image filters
- Conducted literature review for perceptual learning project and wrote NRSA F31 Grant

Visiting Student Researcher (Summer 2018)

Banks Lab, Vision Science Department, *UC Berkeley*

- Assisted in research focused on eye movements in natural scenes, and the effects of digital display resolution on visual perception
- Analyzed eye movement data using MATLAB to plot gaze distribution at multiple fixation distances
- Assisted in stimulus generation for experiment involving perception of display resolution

Researcher (2017 - 2018)

Research Assistant (2016 - 2017)

Visual Information Sciences at Northridge Lab, Psychology Department, *Cal State University, Northridge*

- Assisted with research focused on grapheme color synesthesia, and the effects of virtual reality on the visual system
- Collected data using methods of remote eye tracking and autorefraction
- Created psychophysical experiments using Experiment Builder
- Entered and verified quantitative data in Excel and SPSS
- Presented research in the form of posters at regional and national conferences

Lead Researcher (2016 – 2018)

Applied Social Psychology Research Lab, Psychology Department, *California State University, Northridge*

- Assisted with research focused on the effects of technology on social cognition, and trust in machines
- Trained in electrode placement, data extraction and analysis using Biopac-MP150 System & AcqKnowledge 4.2 Software
- Created surveys using Qualtrics software

- Analyzed quantitative experiments and survey data in SPSS using univariate and multivariate approaches
- Formally pitched new projects and experiments to lab primary investigator

Research Assistant (2014 – 2015)
Neuropsychology Research Lab, Psychology Department, *California State University, Northridge*
- Performed cognitive test battery on patients with early onset dementia and mild cognitive impairment
- Scored and analyzed test battery results to record in Excel
- Assisted in writing abstracts an manuscript sections

TEACHING EXPERIENCE

Psychological Statistics Teaching Associate (2017 – 2018)
Psychology Department, CSUN, Northridge, CA

- Instructor of record for Statistical Methods in Psychological Research Lab
- Educated students in common statistical tests such as t-test, chi square, ANOVA, correlation and regression
- Educated students in data processing programs such as Excel and SPSS to analyze and visualize large datasets
- Graded homework assignments and published class grades in Canvas (online course management system)

Psychology Research Methods Teaching Associate (2017 – 2018)
Psychology Department, CSUN, Northridge, CA
- Educated students in the process of psychological research from the question of interest to the manuscript
- Developed discussion platforms and uploaded files in Canvas
- Held office hours to answer students' homework questions and assist with understanding of lecture concepts
- Graded manuscript sections and provided detailed feedback to improve writing skills
- Assisted in scheduling of other tutors
- Maintained and updated website with resources for statistics and research methods

SKILLS
Statistics/Research: Remote Eyelink Eyetracker, Autorefractor, SPSS, R, MATLAB, Python, ACQ Knowledge, Biopac, Qualtrics, Stata, EQS, Tableau
Productivity: Excel, Word, PowerPoint

ACTIVITIES
Professional
Society for Neuroscience (2016 to present)
Association for Psychological Science (2017)
Vision Sciences Society (2017)

Chapter 1. INTRODUCTION

1.1 THE RETINA

The retina is a delicate layer of tissue (0.5mm thick) made up of hundreds of millions of cells located at the back of the eye, and is responsible for transducing light from the outside world into a visual signal. Transduction begins at the photoreceptors, and ends with transmission of a signal to the optic nerve, which carries visual information to the brain for further processing.

Cells that convert light signal from a visual scene are organized into five different layers (**Figure 1.1**). The Outer Nuclear Layer (ONL), Inner Nuclear Layer (INL), and Ganglion Cell Layer (GCL) contain cell bodies of the photoreceptors, horizontal/bipolar/amacrine, and ganglion cells (the GCL contains displaced amacrine cells), The Inner and Outer Plexiform Layers (IPL, OPL) contain the synapses that transmit signals to cells between each layer.

To initiate a signal, photons of light must travel through all layers of the retina to reach the photoreceptors in the outer-most nuclear layer. Photoreceptors begin to transduce light, sending a signal to bipolar cells, which send a signal to the output cells of the retina: the ganglion cells, located in the most inner nuclear layer. Finally, ganglion cells, whose axons form the optic nerve located in the center of the retina, transmit this information to the Lateral Geniculate Nucleus of the thalamus, and ultimately to the primary visual cortex.

Signal transduction in the retina happens on short timescales and in multiple parallel processing streams. Visual information can flow vertically throughout the retina: an electrical signal is initiated by a photoreceptor, whose signal is then converted by bipolar cells and sent to the last information processing stage in the ganglion cell layer. Conversely, information may flow laterally through inhibitory neurons such as the horizontal and amacrine cells. Each processing stream

(lateral or vertical) transmits information about the visual scene to the visual cortex for further processing.

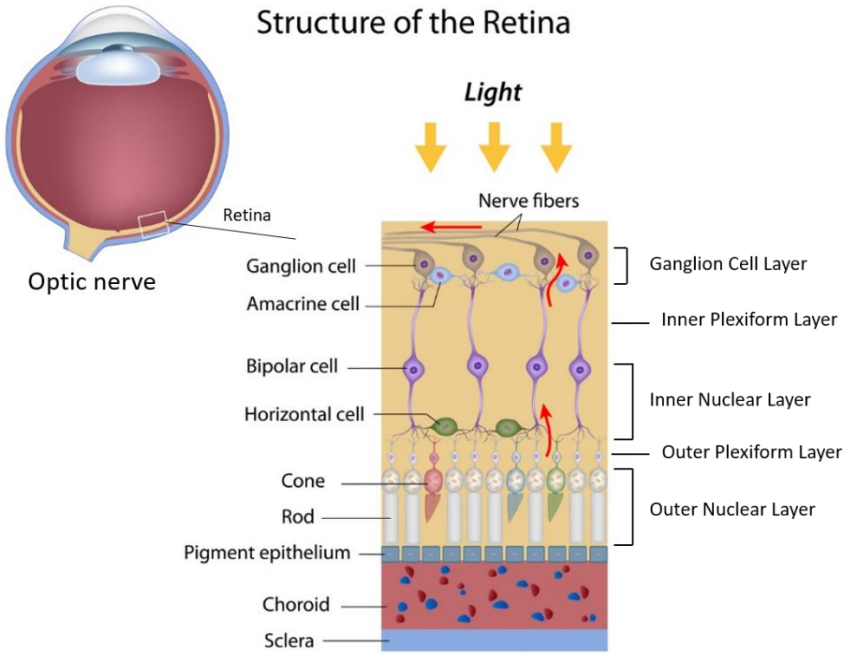


Figure 1.1. Diagram of the eye (top left), and retinal structure. Light travels through the structures of the eye and the initial layers of the retina to the ONL where photoreceptors initiate the transduction of light. The optic nerve transmits retinal signals to the brain. Figure adapted from Discovery Eye Foundation (www.discoveryeye.org/layers-of-the-retina/).

1.2 Retinal Signal Cascade

1.2.1 *Photoreceptors*

The rod and cone photoreceptors of the retina initiate the first stage of visual processing by converting light from a scene into a visual signal. Light-sensitive proteins called opsins within the

outer membranous discs of the photoreceptors convert photons of light into electrical signals that propagate through the various layers of the retina to the brain (Dowling, 1987).

In the dark, the ion channels of photoreceptors remain open and are in a ‘depolarized’ state, actively releasing the excitatory neurotransmitter, glutamate, at the synaptic terminal.. Hyperpolarization of the cell causes photoreceptor ion channels to close, resulting in a decreased release of glutamate.

Photoreceptors are differentially affected by light: each type has its own specialized function within the retina. Rods are extremely sensitive to light, contributing to visual processing in dimly lit conditions. Cones are less sensitive to light and aid in processes such as acuity and color perception (Dowling, 1987).

Rods and cones form a distributed mosaic across the retina in order to sample as much of the visual scene as possible. Different aspects of the visual scene are sampled by each of the two photoreceptors. Cone photoreceptors are in highest density at the center of the retina known as the fovea, and in this region, converge to downstream cells in a 1:1 ratio. Thus, cones are best suited to provide excellent spatial detail about the visual scene. Further, the three subtypes of cones are each responsible for absorbing different wavelengths of light, giving rise to color perception. Rods increase in density with eccentricity away from the fovea and are largely responsible for processing information about motion and light sensitivity.

1.2.2 *Bipolar Cells*

The pathway of signal transduction in the retina continues when bipolar cells receive a signal from the photoreceptors (**Figure 1.1**). Bipolar cells (BCs) can have either metabotropic or ionotropic receptors, each of which cause different responses to the release of glutamate from a photoreceptor

(Dowling, 1987). Metabotropic receptors cause a bipolar cell to depolarize, or 'spike', in response to a decrease in glutamate release (i.e. in scotopic conditions), and hyperpolarize in response to an increase in glutamate (i.e. in photopic conditions). The response pattern is exactly opposite for bipolar cells with ionotropic receptors.

The INL of the retina contains least 13 different types of bipolar cell types, each relaying different information about the retinal image such as luminance, chromaticity (when receiving signals from cones, such as the midget bipolar cell type), and even information about motion (Euler, Haverkamp, Schubert, & Baden, 2014; Kim et al., 2014) to different strata of the IPL. Although morphologically distinct, BCs can be stratified into three different classes: rod bipolar cells (RBCs) which connect exclusively to rods, and two types of cone bipolar cells (CBCs), which connect exclusively to cones, but have axon terminations in the different strata of the IPL.

In the fovea, a single cone sends a signal to a single CBC, giving rise to precise spatial processing of the visual scene. In the rod-dominant periphery, multiple rods converge on a single RBC, and feed their signal into the cone bipolar pathway via direct connections with CBCs.

ON and OFF receptive fields. A major function of retinal signal processing is the decoding of discreet, spatially distinct contrast information about the visual scene which is transmitted to the brain. This function is achieved in part by the receptive fields of cells, which are referred to as the area in visual space that a particular retinal cell is responsible for processing. BCs are tiled across the entire retina, each containing a different receptive field. Although, the receptive fields of nearby cells tend to have some overlap.

The receptive fields of BCs are roughly circular and contain an excitatory center and an inhibitory surround (**Figure 1.2**). On-center BCs have a receptive field that causes the cell to respond

maximally when there is a small spot of light present in the center of its receptive field, and a decrement of light in the surround. Off-center BCs maximally respond to a decrement of light in the center of its receptive field, and illumination in the surround. This concept is known as center-surround antagonism, where a response increase in the center results in a response decrease in the surround. Thus, BCs are the first cells in the retina to initiate the processing of contrast information within the visual scene. On- and -off information is therefore separated into parallel processing streams within the retina.

ON-Bipolar Cells. As reviewed earlier in this section, when a photoreceptor absorbs light, it is immediately hyperpolarized, causing a decrease in the release of glutamate at the synaptic cleft. On-BCs have metabotropic receptors respond to a decrease in glutamate, causing depolarization (**Figure 1.2**). Following the depolarization, on-BCs release glutamate to their downstream synaptic connections and continue the retinal signal cascade.

When a photoreceptor is depolarized in the dark, the increase in glutamate release from a photoreceptor causes the metabotropic ion channels to close, leading to hyperpolarization of the on-BC.

OFF Bipolar Cells. Off-BCs have an equal and opposite response to glutamate release from photoreceptors, due to the presence of ionotropic receptors on their dendrites. When a photoreceptor is depolarized in the dark, ionotropic receptors of off-BCs open and glutamate flows through the BC leading to a depolarization. Consequently, the decrease in glutamate release from photoreceptors in the light causes the hyperpolarization of off-BCs.

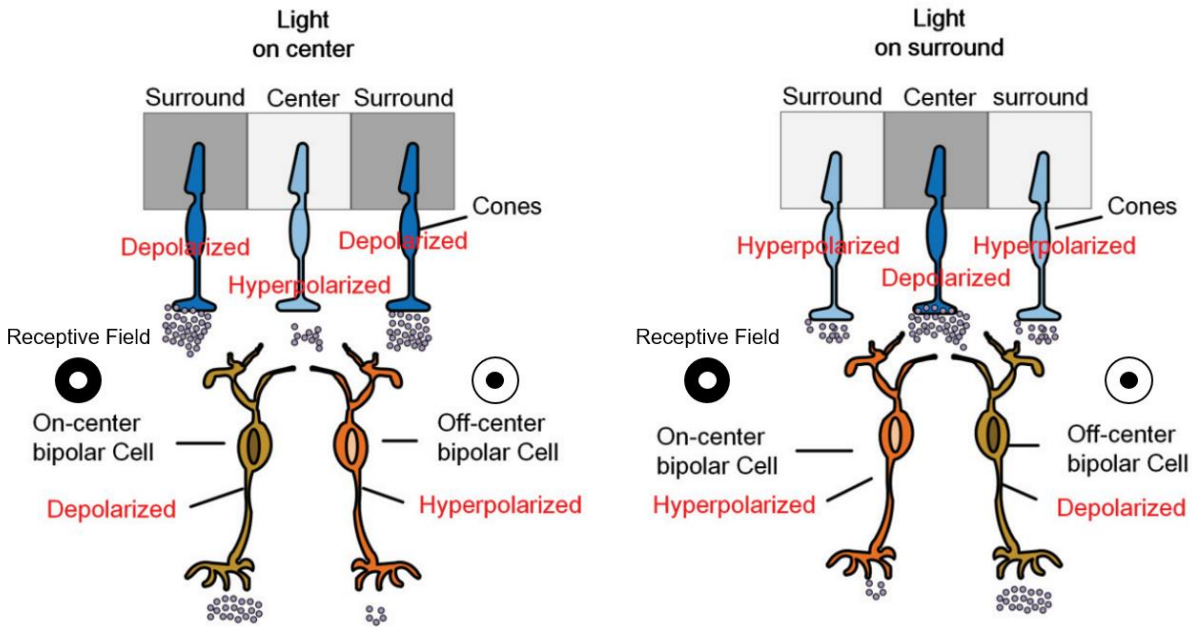


Figure 1.2. Center surround antagonism for bipolar cells. Adapted from Eshraghian, Cho, Baek, Kim, & Eshraghian (2017).

1.2.3 *Inhibitory interneurons of the IPL*

The IPL of the retina is made up of cells that transmit information to the optic nerve either vertically, or horizontally. When a photoreceptor transmits information vertically, signal propagation begins at the photoreceptors, is sent to a BC, and subsequently a retinal ganglion cell (RGC). The retina also processes signals from the photoreceptors that are transmitted horizontally. These photoreceptor signals can be sent to horizontal or amacrine cells, which typically release the inhibitory neurotransmitter GABA. These inhibitory signals travel across the retina to help shape the receptive fields of BCs and RGCs.

Amacrine Cells. Amacrine cells are the most diverse cell type in the retina and have large dendritic spreads with no axons. Two major classes of amacrine cells can be identified based

how many layers of the IPL they synapse to. Diffuse amacrine cells extend through multiple layers through the IPL, whereas stratified amacrine cells synapse into only one or a few layers of the IPL (Dowling, 1987; Masland, 1988; Cajal, 1972). Roughly 42 types have been identified, however the function of all these types is not well understood (Masland, 1988). Amacrine cells synapse directly onto bipolar cells, providing inhibitory input to BCs, which ultimately helps shape receptive fields and enhance contrast detection. In addition to providing cells with inhibitory signals, amacrine cells also gather input from rod BCs in the periphery and relay their signals to cone BCs by synapsing directly onto the dendrites of cone BCs (Dowling, 1987). Finally, amacrine cells can act as modulators for shaping the output of postsynaptic ganglion cells, and may be responsible for shaping the orientation selectivity of RGCs (Dowling, 1987).

Horizontal cells. Horizontal cells contain large dendritic arbors, and are occasionally axonless. They typically have very large receptive fields (sometimes far larger than their dendritic field size) and hyperpolarize in response to light stimulation (Dowling, 1987). By providing BCs with inhibitory input, they seem to be responsible for forming the surround of BC receptive fields. It is reasoned that because the surround region of BC receptive fields exhibit a latency response that is roughly similar to the longer latency response of horizontal cells, that these cells might contribute to the surround portion of BC receptive fields (Dowling, 1987).

1.2.4 *Retinal Ganglion Cells*

BCs and amacrine cells send signals to the retinal ganglion cells in the innermost nuclear layer of the retina. The axons of ganglion cells form the optic nerve which transmits much of the retinal signal to the lateral geniculate nuclei of the thalamus, and ultimately the primary visual cortex (V1) in the occipital lobes of the brain.

The number of morphologically distinct ganglion cell types is still up for debate, however there is a consensus range of anywhere between 15 to 20 (Masland, 2012). Ganglion cells can also be distinguished based on their function within the retina. Some are spatially selective (ON-center and OFF-center), and some are directionally selective. Some ganglion cell types also exhibit color selectivity (e.g. blue ON and blue OFF cells), and some exhibit intrinsic photosensitivity due to the expression of melanopsin, which has been shown to regulate pupillary responses and circadian rhythm (Sanes & Masland, 2015).

RGCs exhibit the same center-surround antagonism as bipolar cells, and inherit the receptive field properties of their upstream bipolar and amacrine cell connections, ultimately exhibiting more complex retinal properties than BCs. For example, ganglion receptive fields have been shown to be orientation selective, a property that is inherited from orientation selective amacrine cells (Bloomfield, 1994). Thus, rather than having a concentric center-surround antagonism (as previously described), the receptive fields of RGCs are elongated along the preferred orientation of the cell, and flanked by two inhibitory zones on either side (in the case of on- RGCs. The opposite is true for off- RGCs).

1.3 Retinal Degeneration that Leads to Progressive Vision Loss

Inherited retinal dystrophies (IRD) are a diverse set of genetic disorders that cause the progressive degeneration of a number of different cell types of the retina. IRDs can range in severity and age of onset, and can result in near or total blindness in late stages. As the IRDs progress (some taking several years to several decades), there is often a decline in patients' quality of life, resulting from emotional, physical, and socioeconomic challenges (Prem Senthil, Khadka, & Pesudovs, 2017).

To date, there have been over 300 identified genes (see web.sph.uth/edu/RetNet/disease.htm for a comprehensive and updated list) and thousands of possible mutations of those genes associated with IRDs. Many of these genes are connected to the regulation of phototransduction proteins, photoreceptor function, and genetic expression (Bessant, Ali, & Bhattacharya, 2001). IRDs can be categorized based on the affected retinal area: degraded functionality in the periphery, such as retinitis pigmentosa and choroideremia, degradation of central vision such as Stargardt disease, or both, such as cone-rod or rod-cone dystrophies (Broadgate, Yu, Downes, & Halford, 2017).

While IRDs are typically characterized as monogenetic diseases (Broadgate et al., 2017), diseases caused by mutations on a single gene, mutations along multiple genes can cause many other retinal diseases. Age-related macular degeneration (AMD) is a common polygenic disease that results in the progressive loss of central vision. AMD is thought to have a contain a genetic component and has been recently associated with mutations on a number of genes, which may be the result of environmental factors such as smoking and being overweight (Fritsche et al., 2016).

Of particular importance here are retinal dystrophies that make up the world's majority cases of blindness. The most common IRD, Retinitis pigmentosa (RP), affects 1 in 3000 individuals (Francis, 2006), initially causing some rod photoreceptor loss ultimately complete blindness in late stages. Age-related macular degeneration causes progressive degeneration of the retinal pigment epithelium and eventually all cones and many rod photoreceptors, and is the world's leading cause of blindness (Francis, 2006).

1.4 Age-Related Macular Degeneration

The prevalence of AMD has been assessed in numerous epidemiologic studies. A recent meta-analysis (Wong et al., 2014) estimates the global prevalence of AMD in persons 45-85 years of

age is an astounding 8.7% (95% CrI 3.98%–15.49%). Although the average age of onset varies between races, AMD is typically more prevalent in Caucasian and Asian populations. Due to an increasingly aging world population, the incidence is expected to rise from 196 million people affected in 2020 to 280 million in 2040.

AMD (for reviews, se de Jong, 2006; Jager, Mieler, & Miller, 2008; Lim, Mitchell, Seddon, Holz, & Wong, 2012) is a disease of the outer retina – specifically, Ruysch's complex which is comprised of the photoreceptors, the retinal pigment epithelium, and Brusch's membrane. These latter two layers of the retina lie posterior to the photoreceptor layer, and aid in the function of the photoreceptor matrix. When functioning properly, cells in the Retinal Pigment Epithelium (RPE) aid in the protection of the retina from disease through multiple growth and repair factors, relieve ocular pressure through the transport of water to the subretinal space, supply pigmentation (cis-chromophore) to bleached photoreceptors, replace discarded photoreceptor outer segments, and aid in the maintenance of the photoreceptor matrix which functions to support photoreceptors and bond the RPE to the retina (Marmor & Wolfensberger, 1998). However, as one ages, debris called drusen begin to form between the RPE and Bruch's membrane, which decreases the permeability of fluid and nutrients from the choroid to the RPE. While drusen can be present in healthy individuals beyond age 50 in the absence of AMD, excess drusen that form in the macula, the central part of the visual field (contains the fovea, foveola), are a hallmark of an AMD diagnosis (de Jong, 2006).

There are two forms of AMD: neovascular (wet AMD), and geographic atrophy (referred to as one form of dry AMD). Neovascular AMD often results from the hemorrhaging of vessels that break through the RPE into the neural retina, and leak fluids that lead to retinal scarring and distorted vision (Lim et al., 2012). While wet AMD is less common, only comprising between 10-15% of

cases, it is the most severe, often resulting in profound visual loss within days or weeks of symptom onset. Dry AMD results in the progressive accumulation of drusen over months to years, and is often initially asymptomatic. As drusen disrupt the supply of nutrients from the choroid to the retina, visual acuity becomes impaired, and central scotomas (blind spots) may begin to form. Advanced dry AMD results in profound vision loss (Snellen visual acuity 20/200; <0.1 logMAR), decreased quality of life due to economic and social impacts, and often results in decreased mental health outcomes for patients.

1.5 Retinitis Pigmentosa

Retinitis pigmentosa (RP) refers to a group of IRDs whose main characteristic is the loss of photoreceptor function (for reviews see Hamel, 2006; Hartong, Berson, & Dryja, 2006; Pagon, 1988). The most common form is rod-cone dystrophy in which the rod photoreceptors in the peripheral retina are the first to deteriorate, followed by the death of cone photoreceptors in late stages. RP can be diagnosed in early stages of the disease through examination of a retinal fundus. Typically, fundus images will show pigmentary deposits in the peripheral retina and loss of vasculature (Hamel, 2006). There have been several hundred genes and mutations linked to the development of RP, however, the exact pathway to photoreceptor death is still unclear.

Symptoms typically begin in childhood but are often ignored by the patient because acuity is largely preserved, usually leaving RP undiagnosed until a patient's teenage years. The key diagnostic factor is an electroretinogram (ERG), a test that measures the electrical activity of the retina recorded by placing a thin electrode on the patient's cornea to record responses of retinal neurons non-invasively (Hamel, 2006). The ERG of a patient with early RP is characterized by a decrease in the amplitude of photoreceptor electrical activity (Hartong et al., 2006). As RP

progresses over a period of years, symptoms (increased night blindness, peripheral blind spots, decreased sensitivity to pale blue and yellow hues, and photophobia) become more obvious to the patient and the fundus image presents with dark deposits in the periphery, while the macular and foveal regions are largely spared. In the end stages, patients develop complete tunnel vision and typically require disability assistance as they only have a few degrees of a functioning visual field. The fundus presents with pigmentation across the entire retina and ERGs are unrecordable.

1.6 VISION RESTORATION

There is currently no known cure for degenerative photoreceptor diseases such as RP and AMD, though there are reports of some effective medications in combination with nutritional regimens that may result in slower disease progression (Hartong et al., 2006). In recent years, sight restoration technologies have emerged as a potential avenue to restore some visual perception in those who suffer from retinal blindness.

Methods of restoring vision include optogenetics, small molecule photoswitches, and electrical stimulation of areas along the visual pathway (retina, optic nerve, LGN, and occipital cortex). Currently, none of these approaches recreate normal vision. Descriptions of these methods, in addition to ways electrical and optogenetic stimulation deviates from natural retinal processing are discussed below.

1.6.1 *Optogenetic stimulation of the retina*

The strategy of optogenetics is to transform surviving retinal into photosensitive units that fire in response to certain wavelengths of light. This method of vision restoration was made possible by the discovery of light-gated cation channels called channelrhodopsins (ChR) (Nagel et al., 2003). The finding that microbial rhodopsins can result in either the depolarization or hyperpolarization

of healthy retinal cells has inspired several different groups to develop optogenetic devices for sight restoration.

This method works by packaging ChR (or ChR2) inside an adeno-associated virus (AAV) vector to deliver to healthy retinal cells. Several proof-of-concept studies in mice targeting retinal ganglion cells and bipolar cells to express ChR2 have resulted light sensitivity for the affected cells (for a review, see Pan, Lu, Bi, Dizhoor, & Abrams, 2015). Once cells are converted into light sensors, patients wear specialized goggles with an attached camera to process the incoming retinal image. To date, there are currently four groups (GenSight Biologics, Allergan, Bionic Sight LLC, and Nanoscope Therapeutics) that are in active clinical trials to evaluate the safety, tolerability, and efficacy of several different opsins to treat advanced RP (for a review, see Prosseda, Tran, Kowal, Wang, & Sun, 2022; see. Sahel et al.,(2021) for first in-human data)

The use of ChRs and optogenetic approaches to vision restoration has exploded in the past decade, but success of this method remains elusive until several challenges can be addressed. The photostimulation system will need to incorporate a device that is small and comfortable enough to be worn by a patient, and be equipped with sophisticated technology that stimulates the newly photosensitive neurons with single-cell precision, which would require an eye-tracking system that tracks the light-sensitive locations of the retina (Prosseda et al., 2022). Further, the ChR2 opsin exhibits very low sensitivity to light and therefore could require potentially damaging levels of light to stimulate retinal cells. Recently, mutant ChRs have been explored, each possessing significantly higher light sensitivity than the original ChR2. However, increasing the light sensitivity with various variants of ChR may come at the cost of decreased temporal kinetics of activation due to prolonged channel opening (Pan et al., 2015). Research in this field is continuing

to grow through the exploration of alternate light-sensitive proteins such as cone opsins, which display similar sensitivity to rhodopsin, and much faster kinetics (Berry et al., 2019).

Small-molecule photoswitches, provide light sensitivity to cells by targeting the opening and closing of existing ion channels for depolarization,as opposed to optogenetics, which involves the creation of a new ion channel (Van Gelder, 2015). However, significant drawbacks in this field are small molecule photoswitches' transient effects (Pan et al., 2015). Thus, there is a continued need to redeliver these molecules to retinal cells in order to have the desired effect. That being said, because of its impermanence, success using this method may serve as a steppingstone to the optogenetic method, which results in permanent, irreversible alterations.

1.6.2 *Electrical stimulation of neurons*

Neurons communicate through electrical signals known as action potentials. Typically, a neuron has a resting membrane potential of -70 mV, which indicates that there are more negatively charged ions inside the cell, relative to the extracellular space around it. The dendrites of each neuron contain receptors that allow chemicals known as neurotransmitters to flow inside of the neuron. Neurotransmitters can be positively or negatively charged, and therefore initiate a change in the neuron's resting membrane potential. Once there is a strong enough concentration of neurotransmitter, an action potential, or 'spike' is initiated along the axon of the neuron. This electrical signal then causes the release of neurotransmitter into the synapses at the end of the neuron. Thus, the neuron is responsible for converting chemical signals from neurotransmitters to electrical signals that allow communication of information throughout the entire nervous system.

In 1775, Charles LeRoy reported applying several electrical shocks to a blind patient's head, initiating visual percepts in the patient. However, the mechanisms for these sensations were not

understood until the mid-1800s, when several scientists discovered that they could replicate neurons' electrochemical communication process by delivering electrical current to a neuron via an electrode near the cell to initiate action potentials (for a detailed history of this discovery, see Doty, 1969). Today, devices known as brain-machine interfaces or neuroprosthetics use electricity delivered via implanted microelectrodes to stimulate neural activity in the hopes of treating diseases and injuries to the nervous system (e.g. Parkinson's, amyotriphic lateral sclerosis (ALS), spinal cord injuries, and stroke. For a review see Burns, Adeli, & Buford, 2014), psychiatric disorders such as depression (Williams & Okun, 2013), chronic pain (Boccard, Pereira, & Aziz, 2015), and restoration of sensory function such as audition (Zeng, 2004), and sight (P. M. Lewis, Ackland, Lowery, & Rosenfeld, 2015).

Sections 1.6.3 through 1.6.6 review the different approaches to stimulating the visual pathway to restore vision. Each method is reviewed, and challenges are discussed. **Figure 1.3** presents a summary of different prosthesis designs.

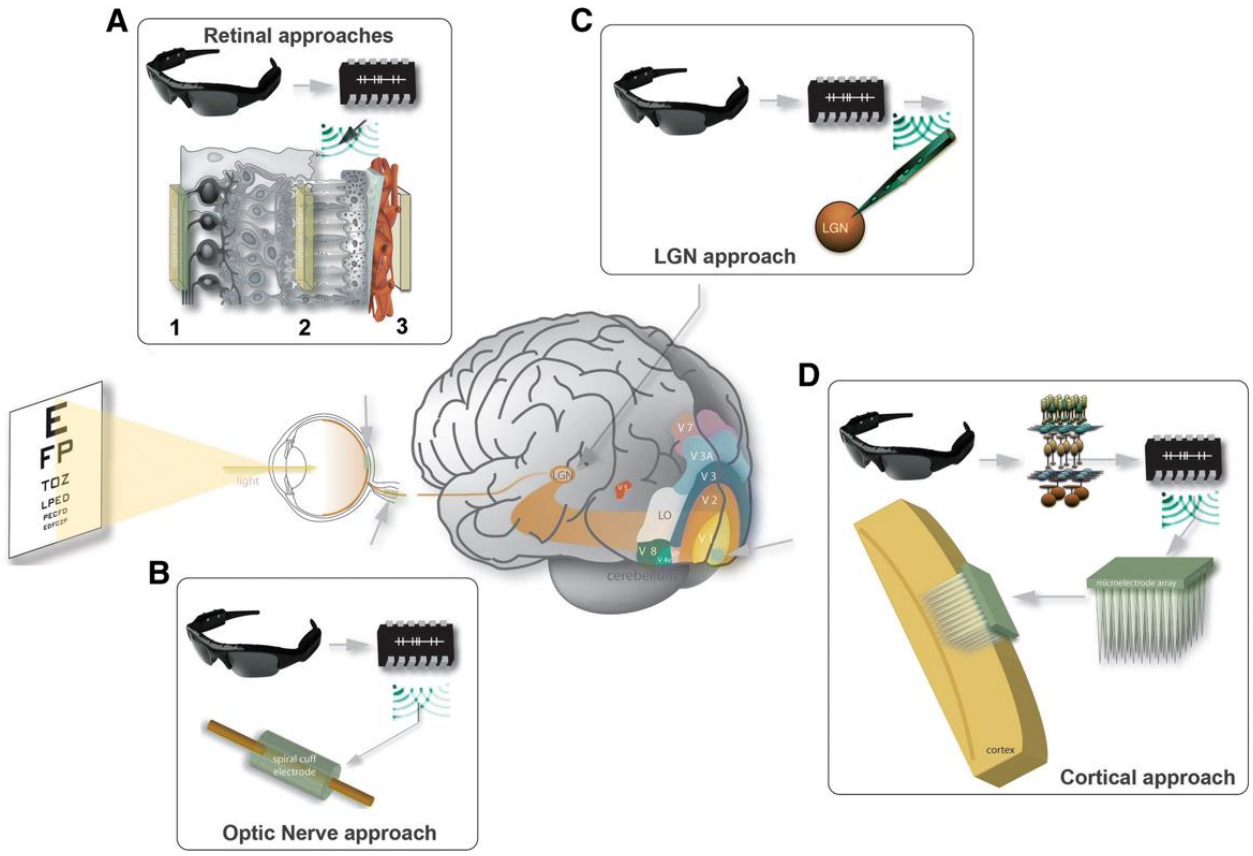


Figure 1.3. Summary of different approaches to vision restoration using microelectrode stimulating arrays in different areas of the visual pathway. In each method, patients are fitted with a pair of glasses with a head-mounted camera which processes incoming visual imagery and wirelessly sends it to a receiver which converts the images into electrical pulses that are transmitted to the stimulating array. A) Retinal prosthesis (outlined in section 1.6.3). Areas 1, 2, and 3 represent different possible locations of the stimulating array within the retinal space. B) Optic nerve prosthesis (outlined in section 1.6.4). C) LGN prosthesis (outlined in section 1.6.5). D) Cortical prosthesis (outlined in section 1.6.6). The microelectrode array is placed near area V1. Figure from Fernandez (2018).

1.6.3 *Retinal Prosthesis*

Retinal stimulation devices consist of a camera that processes incoming visual information, a processing unit that analyzes and converts the visual image into electrical pulses, and a microelectrode array placed on one of three locations in the retina. Device placement includes epiretinal (on top of the ganglion cell layer), subretinal (occupying the space of the degenerated photoreceptors), or suprachoroidal (between the choroid and the sclera) locations. **Figure 1.4** depicts the organizational flow of visual information from the photoreceptors to the ganglion cells, with information on the placement of the three types of retinal devices.

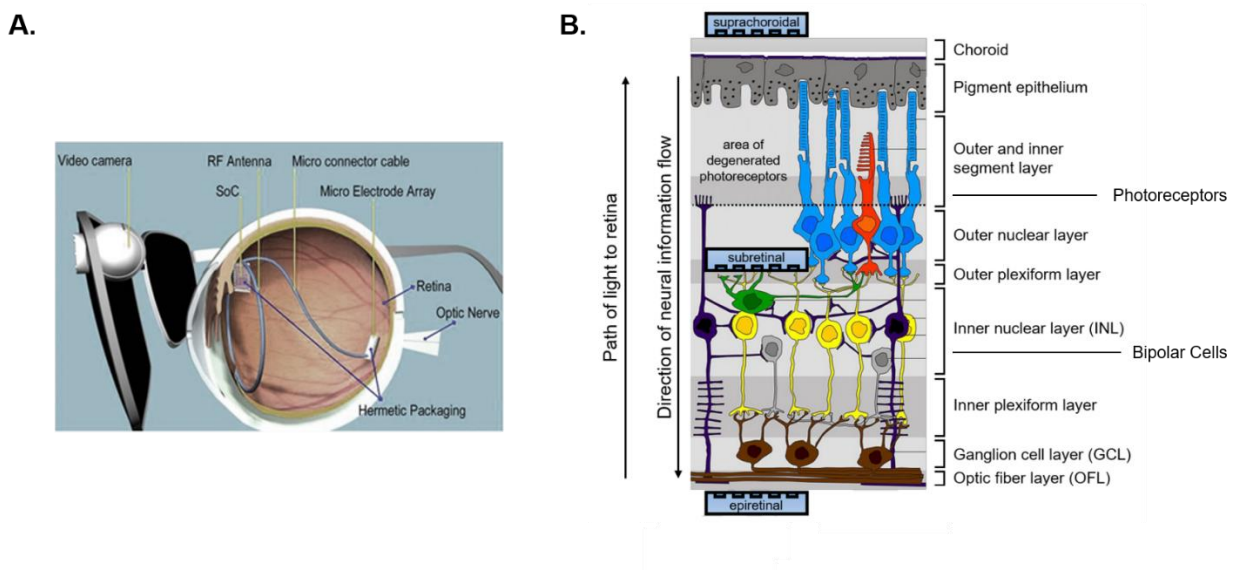


Figure 1.4. A. Schematic of a retinal implant in the eye. The array is tacked onto one of three locations on the retina, and a coil is placed around the sclera which keeps the array in place. An antenna is embedded within the device which aids in the wireless transmission of image information from the video camera to the electronic processing unit (SoC). The electrode array is wrapped in hermetic packaging to prevent direct exposure to retinal vasculature. Adapted from image provided by Dr. Wentai Liu, UCLA. **B.** Diagram of the retina and its various layers.

Implants are either placed in the epiretina area on top of the ganglion cell leyer, in the subretinal space near degenerated photoreceptors, or just outside the retina in on top of the choroid. Adapted from (Beyeler, Boynton, Fine, & Rokem, 2017).

Epiretinal Stimulation. Epiretinal stimulation bypasses the inner retina to position electrodes on top of the ganglion cell layer. The attractiveness of this device is the relative ease of implantation and removal of the device, and the ability to activate RGCs directly, bypassing retinal reorganization that occurs as a result of photoreceptor disease.

The Argus II system (Second Sight Medical Products Inc., Sylmar, CA) underwent clinical trials beginning in 2007, and implanted a 60-channel array on the macular region of the inner retina in 29 patients with RP, and one with choroideremia (Ho et al., 2015; Humayun et al., 2012). A receiving transmission coil and electronics case were fixed outside the eye, attached to a band circumventing the sclera. Patients wore glasses that contain a camera which converts visual imagery to electrical signals on a receiver coil attached to the glasses, and wirelessly transmits it to the implanted receiver coil. The receiver coil then sends the electrical signals to the implanted array which converts the signals into pulses that are emitted by individual electrodes to stimulate inner retinal cells. A three-year follow up study showed that while the device was well tolerated and patients performed better on visual tests with the system on than off, the best recorded logMAR from a single patient was a mere 1.8 (20/1262 Snellen acuity), and the mean visual was 2.5 logMAR (20/26,325 Snellen acuity). The five-year Argus II follow up showed almost no improvement over time in square localization, motion discrimination, or visual acuity (da Cruz et al., 2016). To date, the Argus II is the only epiretinal prosthetic device with United States FDA approval.

Between 2016 and 2017, ten RP patients were implanted with the Intelligent Retinal Implant System II from Pixium Vision (Muqit et al., 2019). The device is similar to the Argus II, except that it consists of a 150-electrode array. While mean visual acuity with the device switched on was not reported, there were significant improvements in patients' performance on square localization and direction-of-motion tests.

As opposed to the Argus II and IRIS II, the EPIRET3 contains all the device components fixed intraocularly, and avoids the need for a physical transscleral cable connecting the visual processor to the array, reducing the risk of infection (Mokwa et al., 2008). The EPIRET3 array consists of only 25 penetrating electrodes in the ganglion cell layer. A feasibility clinical trial in six patients consisted of implantation of the device for one month. Implantations and explantations were successful with minimal complications, and results showed very low stimulation thresholds, with successful detection of phosphene ranging in size and shape.

Other epiretinal devices have been designed and conceptualized with differing array sizes and electrode density ranging from 256 (K. Chen et al., 2010) to 512 (Monge et al., 2013), but none have moved into clinical trials. However, it is not clear whether increasing the size of the array and density of electrodes will lead to a perceptual advantage, both because the underlying physiology of the retina and the rigidity of the device. A flat electrode array being fit to the curved retinal surface results in only a certain number of electrodes on the array being fixed directly on the retina, while others will be further away. This creates issues with stimulation thresholds for each electrode, which can ultimately decrease the resolution of the device.

Recently, the field of epiretinal devices has moved toward foldable arrays made of elastic materials such as polydimethylsiloxane (PDMS) (Ferlauto et al., 2018). Further, the use of photovoltaic

electrodes which converting incoming light from the pupil into electrical impulses, removes the need for an image processing camera, as the electrodes themselves are photosensitive.

Significant challenges with epiretinal stimulation create a major barrier to the success of these devices. Stimulation of the ganglion cells bypasses retinal interneurons, each of which substantially contribute to normal retinal processing. Thus, placing a device at the end of the retinal neural pathway may compromise a patient's percept.

Another challenge concerns the inability to avoid stimulating the RGC axonal fibers that form the optic nerve. The desired placement of a stimulating electrode is directly over the soma of a neuron to produce a small, circular spot of light known as a phosphene. However, in most cases, patients report elongated, 'oval' like phosphenes (Fine & Boynton, 2015; Picaud & Sahel, 2014; Tsai et al., 2012). Recent work has shown that modifying stimulation parameters such as increasing the pulse duration and frequency (i.e. the number of pulses delivered in a given timeframe) by orders of magnitude than in current devices may help reduce this effect (Beyeler, Boynton, Fine, & Rokem, 2019; Tsai et al., 2012; Weitz et al., 2015).

Subretinal Stimulation. These devices are placed in the area of the damaged photoreceptors, in order to take advantage of the significant retinal processing that occurs before delivery of a signal to the RGC layer. It is presumed that stimulating the retinal pathway as 'early' as possible will help create a more realistic percept for patients. Groups such as Retina Implant AG (Reutlingen, Germany), Pixium Vision (Paris, France), Boston Retinal Implant Project (Boston, MA, USA) and Optobionics (Glen Ellyn, IL, USA) have worked towards development of a subretinal prosthesis (for a review see Bloch, Luo, & da Cruz, 2019).

The Alpha IMS (Retina Implant AG, Reutlingen, Germany) was granted European regulatory approval for the treatment of RP in 2013. The stimulating array consists of 1,500 micro-photodiodes that each process incoming light and convert it into an electrical signal, connected to a power cable that leads to a coil which magnetically connects to the back of the ear (Stingl et al., 2013). The amplitude of the electrical signal is based on the brightness of the incoming light, which processes the visual scene at 5-7 Hz. In a clinical trial (NCT01024803) of nine patients, all reported light perception, but remaining visual tests were mixed (Stingl et al., 2015). Visual acuity was only able to be recorded in two subjects, and ranged from logMAR 1.43 (Snellen acuity of 2/546) to logMAR 2 (Snellen acuity of 20/2000). Most patients were able to localize objects, and some were able to read select letters. Five out of nine subjects reported the implant was useful in improving their daily activities.

Subretinal implantation of a photovoltaic device (PRIMA, Pixium Vision, France) was tested in a clinical trial of 5 AMD patients (NCT03333954) between 2017 and 2018, and yielded promising results . The device was placed in the scotoma (blind spot) of AMD patients with some residual peripheral vision in the affected and fellow eye. In addition, the device was equipped with the ability to magnify incoming visual imagery by up to a factor of 8. The combination of residual vision and prosthetic vision with the patient's preferred image magnification factor led to staggering increases in Snellen acuity (up to 20/63) post implantation (Palanker, Le Mer, Mohand-Said, & Sahel, 2022).

A significant issue with subretinal implantation is the extreme difficulty of surgery due to the formation of the glial seal, which separates the neural retina from the RPE, after significant cone photoreceptor death (Marc, Jones, Watt, & Strettoi, 2003). In late-stage disease, microglia

migration to the subretinal space renders this area nearly inaccessible due to the risk of major damage to the choroid, potentially harming the patient (Marc et al., 2003).

Although the idea of taking advantage of the entire retinal code through implantation of a subretinal device is attractive, a notable consequence of photoreceptor death is the initiation of a completely remodeled inner retina. This includes neuronal cell death, migration of bipolar, amacrine, and horizontal cells to foreign retinal layers, and the formation of neurites on inner retinal cells that ultimately lead to new and foreign synaptic connections (Marc et al., 2003). Development of future subretinal devices should consider the significant remodeling that occurs in the degenerated retina. One potential solution may be to implant these devices while patients still have some residual vision (Palanker et al., 2022) and before the retina completes the final stages of remodeling. However, feasibility of this strategy remains questionable due to ethical reasons, and because diagnosis of photoreceptor disease can happen after the inner retina has already undergone significant remodeling.

Suprachoroidal Stimulation. The advantage of implanting suprachoroidal device is a much simpler surgery, due to the avoidance of penetrating the retina. Sinclair et al. (2016) developed a technique where the array is inserted and essentially ‘sandwiched’ between the sclera and the choroid, which avoids the need for surgical tacks as in epiretinal and subretinal devices.

In 2012, Bionic Vision Australia tested a device with 33 stimulating electrodes in three patients, marking the first phase 1 human clinical trial of a suprachoroidal prosthesis (Ayton et al., 2014). All three patients experienced significant postoperative hemorrhaging, which ultimately resolved without intervention. Phosphenes were perceived by all three patients, and acuity testing in one patient revealed a logMAR acuity of 2.62 (Snellen acuity 20/8397), whereas acuity could not be measured in the patient with the device off. Bionic Vision Australia is currently developing a next

generation device that incorporates more electrodes and different stimulation patterns with the goal of improving resolution.

Beyond severe postoperative risks, suprachoroidal placement of a prosthetic is likely located too far away from surviving retinal neurons, necessitating higher and potentially harmful levels of stimulation to elicit phosphenes. Further, placing the array further from the retina results in electrical current having to travel a further distance to stimulate neurons, resulting in greater current spread compared to epiretinal and subretinal devices, leading to decreased resolution (Bloch et al., 2019).

1.6.4 *Optic Nerve Prosthesis*

The optic nerve contains the axons of RGCs and is considered a viable stimulation strategy for vision restoration. Animal (Lu et al., 2013) and human (Delbeke, Oozeer, & Veraart, 2003) studies have shown that stimulation of the optic nerve roughly maintains the visuotopic organization of the retina and cortex. Further, optic nerve prosthesis is attractive when there is significant damage to the retina that renders the site inaccessible.

Devices that that target the optic nerve have been tested in humans using penetrating electrodes (Brelén, Duret, Gérard, Delbeke, & Veraart, 2005) and a non-penetrating cuff wrapped around the perimeter of the optic nerve (Veraart et al., 1998). One patient implanted with a 4-contact non-penetrating cuff perceived phosphenes, however, changing stimulation parameters of the same contact site resulted in completely different phosphene size, shape, and brightness (Veraart et al., 1998). The penetrating device elicited a more reliable perception of phosphenes, however, patients needed extensive training to make use of their perceptions (Delbeke et al., 2003). While accessing the optic nerve is both safer and simpler than accessing retinal or cortical areas, the bundling of

RGC axons in the formation of the optic nerve results in many fibers being wrapped on top of others, so that stimulating certain areas of the visual field is difficult, and resulting increased stimulation thresholds for many fibers.

1.6.5 *LGN Prosthesis*

A large majority of the ganglion cell axon fibers of the optic nerve synapse in the lateral geniculate nucleus (LGN) of the thalamus, which then relays signals to the primary visual cortex. The LGN is retinotopically organized, contains neurons with simple receptive fields, and is entirely accessible with existing implant techniques (Pezaris & Reid, 2007). Thus, for patients where the retina or optic nerve implantation is not feasible, as in conditions such as glaucoma, diabetic retinopathy or ocular trauma, the LGN has been considered as a potential implantation site (Panetsos, Sanchez-Jimenez, Cerio, Diaz-Guemes, & Sanchez, 2011; Pezaris & Reid, 2007).

While no groups have ever brought any devices to a human clinical trial, results of animal studies have shown that electrical stimulation of the LGN elicits responses in primary visual cortex that are similar to those evoked by very simple natural stimuli (Panetsos et al., 2011). Further, these implants would potentially take advantage of the increased representation of the central visual field in the LGN, allowing for higher spatial resolution. The promising evidence in animals suggest studies in humans are needed to determine the safety, mechanical stability, reliability and efficacy of such devices.

1.6.6 *Cortical Prosthesis*

The first stage of visual cortical processing takes place in the primary visual cortex (V1) and extends through the occipital lobe, which consists at least four to five higher visual processing areas. In comparison to the retina or LGN, the size and ease of access of this region are key factors

that drive developers toward cortical placement of a prosthetic. Cortical prosthesis placement can also take advantage of cortical magnification: the overrepresentation of cells dedicated to processing the central visual field (Horton & Hoyt, 1991). Although some portions of V1 are inaccessible due to its positioning in the calcarine sulcus, its overall surface area combined with cortical magnification and relatively simple receptive fields make vision restoration via this method ideal when the retina or optic nerve is damaged due to injury or disease (Najarpour Foroushani, Pack, & Sawan, 2018).

Brindley and Lewin (1968) tested the first prototype cortical prosthesis which consisted of an 80-electrode array placed on the surface of the occipital pole of a 52-year-old patient with bare light perception. Stimulation of electrodes resulted in the consistent perception of phosphenes. Some electrodes invoked multiple phosphenes, which was likely due to current spread from the need for high levels of electrical stimulation because of the distance from the implant to the target cells. Simultaneous stimulation of multiple electrodes resulted in the perception of simple visual patterns. Similar results were found when a 64-electrode array was implanted on the occipital pole of two blind patients (Dobelle & Mladejovsky, 1974).

The field of cortical prosthetics advanced in the 1990s when it was discovered that intracortical microstimulation – implantation of microelectrodes with drastically smaller tip sizes within the occipital tissue – could potentially improve resolution with stimulation currents at least 2 orders of magnitude less than surface-penetrating electrodes (Schmidt et al., 1996).

While no devices have received regulatory approval, several clinical trials are underway (Second Sight Orion VCP system, NCT03344848; ICVP System, NCT04634383; CORTIVIS, NCT02983370). However, these devices will likely suffer from several challenges in recreating visual percepts, and one reason could be due to the small number of electrodes on each array (the

100-channel CORTIVIS system has the highest electrode density). Simulation of pixel density in visual cortex reveals estimates of required electrodes on the order of 625 to 1024 to achieve visual acuity of 20/30 (Cha, Horch, & Normann, 1992). Recently, Chen, Wang, Fernandez, & Roelfsema (2020) investigated visual percepts of macaques in response to stimulation of a 1024 micro-electrode system consisting of 16 connected Utah arrays (Rousche & Normann, 1998). Fourteen arrays were placed across the surface of V1, and two in V4, which allowed for recording of V1 activity driven by electrical stimulation for a less time-consuming threshold estimation process. Macaques were able locate phosphenes with great accuracy, even perceiving shapes and motion when successive electrode stimulation patterns were used. Advances in microelectronics in addition to clever placement of multiple stimulating arrays throughout the visual cortex are progressing the field of cortical prosthetics forward.

1.7 Challenges in Vision Restoration

Each vision restoration method suffers from its own benefits and drawbacks, and selection of which type of device to implant will depend on patient specific needs. As discussed above, all devices require a nontrivial surgery, and do not yet provide adequate restoration of useful vision. Other limiting factors include biocompatibility issues such as acute inflammatory responses resulting from implantation trauma, gradual electrode signal degradation due to the formation of a glial scar around the device, and loss of neural tissue due to chronic electrical stimulation (for a review, see Polikov, Tresco, & Reichert, 2005). Further, the misalignment between the patient's head mounted camera and their eye gaze position results in a decreased ability to spatially locate phosphenes (Sabbah et al., 2014). This results in patients needing to use back-and-forth head

scanning while constantly fixing their eyes in the direction of the head mounted camera to localize objects.

Assuming that the factors described above can be resolved with advances in device hardware and surgical techniques , the inability for a visual prosthesis to restore even moderate forms of natural vision still remains. Even if devices are extremely high resolution, simply activating patterns of electrodes will likely never recreate normal vision. Each of the 60 functionally distinct cell types in the retina require different spatiotemporal electrical stimulation patterns to be activated in a way that mimics normal retinal cell firing. In order to recreate useful vision, prosthetic devices will need to incorporate knowledge of the physiology of the retina and its neural code into electrical stimulation patterns in order to increase efficacy of the device.

1.7.1 *Interactions with the device and the underlying physiology of the retina*

Simultaneous activation of ON- and OFF-pathways. Bipolar cells receive input from photoreceptors that is parsed into multiple different streams of visual information, depending on the receiving BC's on- or off-center receptive field properties. At least 20-30 different types of ganglion cells receive input from their upstream BCs and further discretize this information into multiple parallel streams of encoded information about the visual features of a retinal image and transmit this information to the brain through unique spiking patterns (Dacey, 2004). Current electronic implant and optogenetic technologies are not equipped with small enough electrodes, or the spatiotemporal resolution needed to activate different RGC cell types in a way that mimics cell firing in a biologically natural fashion. This simultaneous and indiscriminate activation of multiple parallel processing streams results in a severely degraded perceptual image.

One of the major issues resulting from indiscriminate activation of cell types, is the simultaneous stimulation of on- and off- visual pathways. These two pathways are complementary to one another: when a bright stimulus activates cells in the on-pathway, the corresponding off-pathway is suppressed, and vice versa (see **Figure 1.2**). However, this is not to say that suppression of the off-pathway when the on-pathway is active indicates a "lack" of information transmission to the brain – rather, it indicates that the function of the on-pathway operates in concert with the off-pathway, in that the response of one is directly dependent on the other (Fine & Boynton, 2015). Thus, turning both pathways 'on' through simultaneous electrical stimulation will result in the delivery of conflicting signals to the brain.

While compacting the array with more electrodes or decreasing electrode size might result in a small increase in resolution, a more effective strategy to address indiscriminate cell activation may be to discover optimal stimulation parameters that specifically target individual RGC types. Several groups have shown that modulating the amplitude, frequency, and pulse duration of electrical stimulation can differentially activate some on- and off-ganglion cell types (for a review see Tong, Meffin, Garrett, & Ibbotson, 2020). For example, Twyford, Cai, & Fried (2014) applied high frequency stimulation to transient on- and off-RGCs that resulted in preferential activation of on-cells, and suppression of off-cells, suggesting that a single stimulation pattern may recreate the naturally opposite firing patterns of on- and off-RGCs. Further, work towards a bidirectional device that records responses of multiple cell types to different visual stimuli and then recreates those firing patterns based on that cell type's preferred electrical stimulation pattern could aid reproducing more biologically natural vision (Shah & Chichilnisky, 2020).

Axonal stimulation. The goal of an electronic retinal device is to stimulate the cell bodies of RGCs to initiate action potentials naturally. However, the organization of the ganglion cell layer

consists of a web of ganglion cell bodies entangled with axons that travel across the entire retina to form the optic nerve. Placement of an epiretinal device on top of the ganglion cell layer makes it impossible to avoid placement of electrodes over axonal fibers. Stimulation of an electrode placed over axonal fibers can cause activation of cell bodies distal to the stimulation site, and cause unwanted activation of groups of cells (Beyeler, 2019; Esler et al., 2018). The perceptual consequences of activating RGC axons are streaky, arc-shaped phosphenes, as opposed to focal stimulation that would lead to a circular phosphene (Beyeler, et al., 2019; Fine & Boynton, 2015) which degrades visual acuity.

It is reasonable to assume that a different electrode array placement strategy, such as in the subretinal or suprachoroidal space, may alleviate the problem of RGC axonal stimulation. However, some patients implanted with a subretinal device have reported perceiving arc-like or semi-circle shaped phosphenes, especially two or more electrodes are simultaneously activated (Wilke et al., 2011). It is likely that the combined electrical activity from multiple active electrodes has the potential to travel through multiple layers of the retina to activate the ganglion cells and their axons. Thus the problem of decreased resolution due to axonal stimulation does not only exist for epiretinal devices.

Many groups have explored addressing the issue of axonal activation by combining alternative stimulation parameters with different electrode placement configuration to achieve focal activation of ganglion cells (Beyeler, 2019; Esler et al., 2018; Italiano, Guo, Lovell, & Tsai, 2022). For example, Beyeler (2019) investigated thresholds in different regions of the ganglion cell soma and defined a ‘dynamic range’ of currents that would activate the soma rather than the passing axons. Further, it was discovered that increasing the pulse duration led to reduced activation of passing axons. Italiano et al. (2022) extended this idea by testing various electrode configurations and

found that a hexapolar arrangement of electrodes resulted in focused current that frequently activated single RGCs, limiting activation of nearby axons. Finally, placement of the electrode array during surgery based on modeling of individual patients' optic fiber layout has the potential to reduce axonal activation by up to 55% (Beyeler, et al., 2019).

Replicating the neural code. As described above, the retina processesde visual input into multiple streams of information, and decodes this information into patterns of action potentials that are sent to the brain. These action potentials are translated into a language that the brain expects and understands. Thus, recreating retinal signal processing will require functional characterization of each cell type. Specifically, a clear understanding of the visual features a cell type responds to, and the pattern of electrical stimulation needed to activate each cell type. Replication of the neural retinal code will also require, at baseline, precise knowledge of the asynchronous patterns of activation in populations of RGCs in response to common visual features, and ultimately dynamic visual scenes (Zheng, Jia, Yu, Liu, & Huang, 2021).

Current prosthetic devices do not transform a visual image into readable patterns of electrical stimulation, limiting patients' visual perception. As discussed above, stimulating electrodes indiscriminately activate multiple different RGC types simultaneously, which leads to an incoherent perceptual experience. The most obvious example of this issue is described above, in which simultaneous activation of on- and off-cell pathways leads to a degraded perceptual image. However, while simultaneous activation of on- and off-pathways may degrade visual perception, it may be possible for the brain make sense of these distorted inputs over time (Esquenazi, Meier, Beyeler, Boynton, & Fine, 2021). Indeed, the perceptual consequences of a missing or degraded on-bipolar pathway, as in the case of individuals with Schubert-Bornshein congenital stationary

night blindness type 1 (Bijveld et al., 2013), appear to be relatively minor (Zeitz, Robson, & Audo, 2015).

Advances in the ability for electronic visual prosthetics to recreate or at least 'mimic' the neural code will come from a better understanding of the retina. According to a recent report, only about half of the different types of RGCs have been categorized and functionally characterized, although significant progress is being made on the remaining types (Sanes & Masland, 2015). Devices will also need the ability to quickly locate and characterize multiple RGC types in a patient-specific manner. Some groups have approached this issue through the development of a bi-directional system that detects neural activity and refines stimulation parameters based on information gathered from the neural activity (Guo et al., 2018; Montes et al., 2019; Shah & Chichilnisky, 2020) . However, these approaches are all still being tested *in-vitro*, and primarily in healthy retinas (Montes et al., 2019; Shah & Chichilnisky, 2020). It remains to be seen how this approach will work in human patients with advanced disease, whose retinal architecture is significantly altered over time. Finally, changes in RGC responses to electrical stimulation in the degenerated retina will need to be investigated, as there is evidence that patients in advanced disease stages report the inability to perceive any visual sensation when stimulating electrodes are activated (Wilke et al., 2011).

1.8 Visual Cortical Plasticity

The cortex is equipped with a staggering ability to react to change as a result of experience. The term cortical plasticity is referred to as the changing and shaping of synaptic connections within the brain as a result of learning, injury, or sensory experience. In 1949, Donald Hebb proposed that the repeated firing of a pre-synaptic cell onto a nearby post-synaptic cell results in an increased

probability that those two cells will fire together and create an efficient, strong connection. For example, when an infant learns to walk, new synaptic connections are created in the cortex. As the skill of walking is mastered, the corresponding neural circuitry exhibits use-dependent changes: the size of neuronal cell bodies increase, dendritic spines grow in number and size, and neurons create more synapses (Rosenzweig & Bennett, 1996).

The developing brain is tasked with refining and shaping networks of neurons as a result of early sensory experience. The period during development in which this optimization of networks takes place is referred to as the critical period, when plasticity is at its peak. Following the critical period, the developing brain stabilizes in activity, and plasticity declines. Normal development during the critical period is directly dependent on experience, and disruptions to sensory systems may cause irreversible damage. Seminal work by Hubel and Wiesel (1963) showed that kittens who were deprived of visual input in one eye from birth exhibited significant atrophy of cells in the LGN, resulting in almost complete blindness in the affected eye. Kittens who were reared with some visual experience displayed similar cortical and behavioral changes, albeit to a lesser extent. In humans, disruptions in alignment of the two eyes during childhood can result in amblyopia, a condition that often causes a lifelong deficit in visual acuity (Lewis & Maurer, 2009).

The temporal window of the critical period differs across sensory modalities. In rats, receptive field tuning for whiskers has a critical period that onsets much earlier than the auditory tonotopic mapping critical period (Hensch, 2004). Critical periods also operate on different timescales with some occurring on scales which last days, while others occur on scales which last years. For example, the human language critical period is roughly twelve years (Hensch, 2004).

While it was previously thought that the human brain is relatively stable following development, dramatic evidence has pointed to the contrary. Indeed, the critical period is the time when cortical

plasticity is at its peak. Exiting the critical period results in decreased neural excitation, and refinement of established neural architecture. However, the brain remains highly sensitive to injury and sensory experience well into adulthood. Cortical maps have the ability to reorganize leading to structural modifications such as an increase in cortical thickness (see Rosenzweig & Bennett, 1996 for a review; Wenger et al., 2012), and the creation of new horizontal and long range connections (Darian-Smith & Gilbert, 1995). Consequently, we can think of the brain as retaining some level plasticity in a 'sensitive period', in which deprivation of a normal sensory experience will not result in permanent deficits, that allows for the potential to form new, long-lasting connections and pathways (Park & Fine, 2020). It follows then that sensitive periods outlast the critical period, but it is unknown how long each sensory modality retains a vast sensitivity to change. While it is typical for plasticity to decelerate after development, evidence shows that it does not completely diminish in adulthood (Fu, Kaneko, Tang, Alvarez-Buylla, & Stryker, 2015; Rosenzweig & Bennett, 1996; Sato & Stryker, 2008; Tschetter et al., 2013). For example, adults can recover from amblyopia, a condition leading to decreased visual acuity, resulting from abnormal visual input during the critical period.

Recent pharmacological and sensory training evidence suggests that the sensitive period has the potential to be reopened at many stages across the lifespan. This might be achieved by placing the adult brain, which is generally balanced in excitatory and inhibitory input, into a more plastic state through the modulation of excitatory neurotransmitters such as norepinephrine, serotonin, dopamine (Bavelier, Levi, Li, Dan, & Hensch, 2010). Commercially available medications such as serotonin reuptake inhibitors (SSRI) increase the amount of serotonin available in presynaptic cells, and are thought to place the brain in a 'plastic' state. Acute administration of SSRIs has been shown to improve visual function in rats with poor visual acuity (Maya Vetencourt et al., 2008),

and has resulted in memory enhancement in humans during a depressive episode (Levkovitz, Caftori, Avital, & Richter-Levin, 2002). These results highlight the possibility for the brain to rapidly respond to sensory impairments, and even improve on cognitive skills long after the close of the critical period.

There are several open questions concerning cortical plasticity in the visual system. How much plasticity does the visual system retain in adulthood? Are there differences in plasticity across different stages of the visual hierarchy? Rather than examining plasticity through a temporal lens by distinguishing between critical, sensitive, and periods of stability, are there distinct visual conditions or contexts (e.g. following injury, changing environmental conditions, or rehabilitative training) that make the visual system plastic? Given these open questions, and a strong consensus that most, if not all areas of the brain are plastic during development, evidence below regarding plasticity across the visual hierarchy will be discussed as it pertains to plasticity in adulthood.

1.8.1 *Plasticity in the primary visual cortex*

The primary visual cortex, also referred to as V1, is the first area of the occipital cortex to process retinal input received by the subcortical LGN. V1 is also typically referred to as the striate cortex, because of its anatomically segregated striping that organizes distinctly different visual features of retinogeniculate input into six different layers. In layer 4 of V1, cells are organized in a columnar fashion by eye of origin known as ocular dominance columns (Hubel & Wiesel, 1959), and the column's receptive field orientation preference, known as orientation columns (Hubel & Wiesel, 1962). Each column processes a specific portion of space and orientation information in the retinal visual field, and thus creates a topographic map between the visual world and the cortex. Beyond layer 4, cells can be activated binocularly, with the corresponding monocular receptive fields

exhibiting significant overlap, and displaying a summative effect in response strength in binocular regions of the cortex (Hubel & Wiesel, 1962).

Evidence suggests that primary visual cortex is particularly stable in adulthood, and not very amenable to change. However plasticity in V1 may aid in continued specialization of neurons to maintain responses to small changes in visual features (Espinosa & Stryker, 2012), and is thus important to study. One way that plasticity in the adult striate cortex has been studied is by examining changes in receptive field sizes and visual field maps after applying a lesion in the retina. These studies have produced several conflicting reports of changes in V1 (reviewed by Wandell & Smirnakis, 2009).

Kaas et al. (1990) lesioned 5 to 10 degrees of a single retina in adult cats, but found no discernable changes in V1 receptive fields. Yet, when injury was added by removal of the non-lesioned eye, within six months a substantial number of neurons in the blind spot created by the retinal lesion (referred to as the lesion projection zone, LPZ) developed new receptive fields representing retinal locations at the borders of the LPZ. Several other studies have recorded similar evidence by observing physiological changes in V1 after lesioning a region of the retina (Calford et al., 2000; Calford, Schmid, & Rosa, 1999; Schmid, Rosa, Calford, & Ambler, 1996). However, functional magnetic resonance imaging (fMRI) revealed results that countered claims of V1 plasticity. Smirnakis et al. (2005) used fMRI imaging to examine potential shifts in cortical activity around the margins of the LPZ following binocular retinal lesions. Results showed no significant change in activity over time near the borders of the LPZ, and electrophysiology mapping of receptive fields substantiated the fMRI results. Additionally, other reports of retinal lesions find little to no evidence of topographic reorganization (Baseler et al., 2011a; Horton & Hocking, 1998).

The conflicting reports of V1 plasticity following retinal lesions may point to methodological differences. The failure to standardize a measurement for cortical reorganization in V1 will naturally lead to an incoherent understanding of plasticity. Smirnakis et al. (2005) points out that the discrepancy between electrophysiology and fMRI results could indicate that reorganization following lesioning only occurs in relatively few, unevenly distributed single neurons that are recorded in physiology experiments. fMRI measures the pooled activity of neuronal populations, and will thus fail to capture changes in single neurons. Further, studying changes in the LPZ might be confounded by reduced inflammation from retinal recovery following a lesion. This highlights the importance of measuring changes on adequate and standardized timescales.

Finally, the studies reviewed here only contest the notion of V1 plasticity on relatively short timescales. Indeed, changes in V1 as a result of plastic mechanisms might take years to develop (Baker, Peli, Knouf, & Kanwisher, 2005), and may also depend on the extent of damage to the retina (Sunness, Liu, & Yantis, 2004). Simply exposing the adult mouse visual system to binocular experience after monocular deprivation at birth does not result in a full recovery of cortical functions (Espinosa & Stryker, 2012). However, some recovery does occur, and an open question remains about how much more recovery could be witnessed in early visual cortex could be on longer timescales.

1.8.2 *Visual cortical plasticity beyond V1*

A vast amount of visual information is processed in primary visual cortex and sent to higher-order areas in the brain for further processing. As low-level inputs in V1 are transformed, representations become increasingly complex, creating a hierarchy of information that begins with basic features such as orientation and spatial frequency, and ends with representation of stimuli in motion,

objects, and faces. In monkey, roughly 30 different cortical areas have been identified. Physiological techniques allow for the analysis of markers that designate one area of the cortex as anatomically distinct from one another. While many of these techniques cannot be performed on humans, fMRI evidence has allowed for the identification of several visual areas beyond V1, referred to as extrastriate areas (Grill-Spector & Malach, 2004; Wandell, 1991), which substantiate anatomical findings. Extrastriate regions include areas V2, V3, V4/V8/VO, MT/hMT+/V5, LOC (lateral occipital cortex), VOC (ventral occipital cortex), and FFA (fusiform face area) (Grill-Spector & Malach, 2004).

Areas V2 and V3 are located near the primary visual cortex and maintain retinotopic organization similar to V1. Cells in V2 receive much of their information from the output of V1 neurons, and respond to subjective contours or "perceived edges" (Peterhans & von der Heydt, 1991), and angled stimuli (Boynton & Hegdé, 2004). Area V3 neurons have been shown to exhibit motion selectivity (Smith, Greenlee, Singh, Kraemer, & Hennig, 1998; Tootell et al., 1997). Color and pattern selectivity has been documented in area V4 (also referred to as area V8 or VO) (Gallant, Braun, & Van Essen, 1993; Roe et al., 2012). It should be noted that the specific function of these earlier extrastriate areas remains somewhat elusive. However, cells in area hMT+ unequivocally respond to many aspects of visual motion (see Grill-Spector & Malach, 2004 for a review), and LOC is a centrally object selective area (Grill-Spector, Kourtzi, & Kanwisher, 2001).

Given the variability of functionality within and between extrastriate areas and a lack of conclusive evidence about the specific functionality of each, there is likely a significant amount of overlap in neuronal processing of visual features along the visual hierarchy. Thus, it is more difficult to make conclusions in lesion studies of extrastriate areas because it is unclear whether there is a functional reorganization of neurons in the LPZ, or whether the lesion results in extrastriate areas assuming

some of the function of the damaged region (Rodman & Moore, 1997). For example, Yamasaki and Wurtz (1991) demonstrated that small lesions of area MT resulted in very little or lasting damage to motion perception. However, when area MT was removed and lesions were applied to other areas that aid in the perception of motion, recovery was incomplete at least 7 months post-injury, and progressed on a prolonged timescale. This work does not preclude the potential of extrastriate areas to exhibit plasticity, rather, plasticity in higher visual areas may be inherently different than that of V1, with changes in function involving the recruitment of areas with similar stimulus processing capability to increase response sensitivity and functionally replace the damaged region.

Work by Yang and Maunsell (2004) directly elucidated the effects of visual training on neuronal responses in area V4 of monkey. After training on an orientation discrimination task, corresponding V4 receptive fields of the trained retinal location exhibited heightened responses and more precise orientation tuning curves than neurons in untrained locations. This study was one of the first to examine neuronal receptive field changes in extrastriate areas of adult animals in response to visual training. Moreover, it appears that altered visual experience even in the absence of explicit training (i.e. exposure to a distorted visual world) has rapid effects on higher order visual neurons (Li & DiCarlo, 2008). As information travels along the visual hierarchy, receptive fields become much more tolerant to stimulus variations such as size or retinal position, a term coined as “stimulus invariance”. Li & DiCarlo (2008) tested a clever paradigm whereby an object freely moved around a screen, but was intermittently swapped with an unexpected object at a new retinal location. After several repetitions of this paradigm, neurons in the inferior temporal cortex, an area that responds to objects and typically exhibits stimulus invariance, began to show

selectivity for distinct objects at specific retinal locations, revealing changes in the receptive field properties.

These studies, in addition to others (e.g. Op De Beck, Baker, DiCarlo, & Kanwisher, 2006) demonstrate the potential for extrastriate areas to reorganize as a function of injury, visual training, and altered visual experience. While it is important to understand the potential for the underlying circuitry of the visual system for plasticity, equally important questions exist about the short and long-term behavioral outcomes of these neuronal changes. Further, as many physiology experiments investigating plasticity cannot be conducted in humans, plasticity as a function of perceptual learning can elucidate the potential of the visual system to transiently or permanently alter connections based on training and experience.

1.9 Perceptual Learning

Visual perceptual learning is defined as an increase in perceptual performance as a result of practice or training, and is a behavioral indicator of cortical plasticity. The ability to learn to discriminate small variations in a stimulus with increasing precision represents a change in cortical optimization at various stages of the visual system as a result of neuronal, cognitive, attentional, and motivational factors. Perceptual learning studies, rather than the lesioning and electrophysiology work discussed above, present a unique way to study plasticity in the adult visual system. Work to understand the mechanisms and conditions under which learning takes place has been documented back to the late 1800s (James, 1890), and continues to be a heavily studied subject in vision research.

Perceptual learning has a variety of clinical and real-world applications. Much research has focused on techniques to harness the effects of visual training to improve visual abilities in adults

with abnormal vision. Amblyopia, for example, is a condition that results in poor visual acuity because of abnormal visual experience during development. Several studies have explored the potential for adults with amblyopia to engage in visual training that improves real world performance in the affected eye (for a review, see Astle, Webb, & Mcgraw, 2011) including increases in contrast sensitivity (Chung, Li, & Levi, 2006; Polat, Ma-Naim, Belkin, & Sagi, 2004) and visual acuity (Chen, Chen, Fu, Chien, & Lu, 2008). Another clinical application of perceptual learning is improving skills of elderly adults, with the aim of translating those skills into quality of life improvements, such training to improve visuo-motor skills and prevent injuries (Bower, Watanabe, & Andersen, 2013).

A typical learning procedure involves an observer repeatedly performing a task in regular intervals over an extended period of time. A psychophysical task such as discriminating oriented lines, detecting low contrast stimuli, or identifying textures and objects is used, however, more dynamic tasks such as playing video games (Achtman, Green, & Bavelier, 2008), have recently gained popularity. Performance, such as accuracy or reaction time for the trained (termed task-relevant perceptual learning), and sometimes untrained feature (termed task-irrelevant perceptual learning) is monitored over time. In virtually all cases, performance on an initially difficult task improves, although learning can be slow and depend on the number of training sessions and the type of skill being trained (Seitz, 2017).

A central question in the perceptual learning literature is whether visual training reflects neuronal changes in V1, or at a later stage along the cortical hierarchy. Changes in the primary visual cortex would suggest that the most fundamental building blocks of vision are capable of reorganization in adulthood, whereas changes at later stages would likely reflect the combined involvement of visual and decision making centers of the brain. Although there are a host of perceptual learning

studies aimed at addressing the cortical locus of perceptual learning, there is little consensus on the topic.

1.9.1 *Evidence for learning-induced changes in primary visual cortex*

Perceptual learning is often specific to the trained retinal location, stimulus feature, and sometimes, eye of origin. This lack of learning transfer results in at least two important implications of perceptual learning. On the one hand, learning that is not generalizable to different contexts presents little applicability in rehabilitative settings where the goal of learning is to improve outcomes of patients with visual abnormalities. On the other hand, this kind of specificity is thought to provide clues into the cortical locus of learning. Specificity of learning has been attributed to changes in area V1 because learning can be strongly monocular (Karni & Sagi, 1991, 1993), and finely tuned for low level stimulus features (Fahle, Edelman, & Poggio, 1995; Schoups, Vogels, Qian, & Orban, 2001; Vaina, Belliveau, Des Roziers, & Zeffiro, 1998) across a narrow spatial range (Crist, Kapadia, Westheimer, & Gilbert, 1997; Fiorentini & Berardi, 1980; Shiu & Pashler, 1992). These are hallmark features of neurons in the primary visual cortex, and researchers have attributed the lack of transfer to indicate changes in the first stage of the visual hierarchy (but see Mollon & Danilova(1996), suggesting that specificity of learning might be attributed to changes in neurons that were specific for the trained task, which does not preclude changes in higher cortical areas).

The hypothesis that learning results in changes in primary visual cortex has also been supported by observing increased blood-oxygen-level-dependent (BOLD) signals in human fMRI studies. For example, Furmanski, Schluppeck, & Engel (2004) measured BOLD responses in V1 before and after training on low-contrast orientation patterns, and found an increase in responses for the

trained, but not untrained orientations. Measurement of electrical activity in visual parts of the brain showed that the amplitudes of waves known to be associated with early visual processing were markedly altered for some parts of the visual field (Pourtois, Rauss, Vuilleumier, & Schwartz, 2008) after training. Further, electrophysiological recordings of neurons in monkey V1 resulted in long-term changes to orientation tuning curves as a result of learning in an orientation identification task (Schoups et al., 2001).

1.9.2 *Evidence for learning-induced changes in higher-order cortical areas*

Specificity of learning for features that the primary visual cortex typically responds to is not a consistent finding in perceptual learning studies. Ahissar & Hochstein (1997) found that learning transferred to different orientations and retinal locations in a simple orientation detection task. Training on a three-line bisection task transferred to nearby retinal locations, but with a range that is greater than the receptive field size of neurons in V1, suggesting that changes involved interactions between multiple regions of V1, or higher-order visual areas (Crist et al., 1997; Gilbert, Sigman, & Crist, 2001). Further, lack of interocular transfer, a central tenet of the claim that V1 is altered as a result of training, has been challenged by studies that show learning transfer across the two eyes (Ball & Sekuler, 1987; Beard, Levi, & Reich, 1995).

Neuroimaging studies of perceptual learning in adult monkeys has also challenged the involvement of purely visual areas in perceptual learning. Training monkeys on a motion discrimination task did not result in any discernable retuning of neurons in the motion processing center of the cortex (area MT), but rather in changed the sensitivity of neurons in the lateral intraparietal area that is involved in decision making for eye movements (Law & Gold, 2008). These results, in addition to

the idea that specificity of learning may not be exclusively indicative of changes in V1 call the early-change theory into question.

To explain the conflicting evidence of primary visual cortex involvement in perceptual learning, an alternate model has been proposed. Rather than changes in neuronal retuning, it is proposed that learning is a result of selective *reweighting* of neuronal channels in the visual cortex, and regions of the brain involved in decision making, see **Figure 1.5** (Dosher & Lu, 1998; Sasaki, Nanez, & Watanabe, 2009).

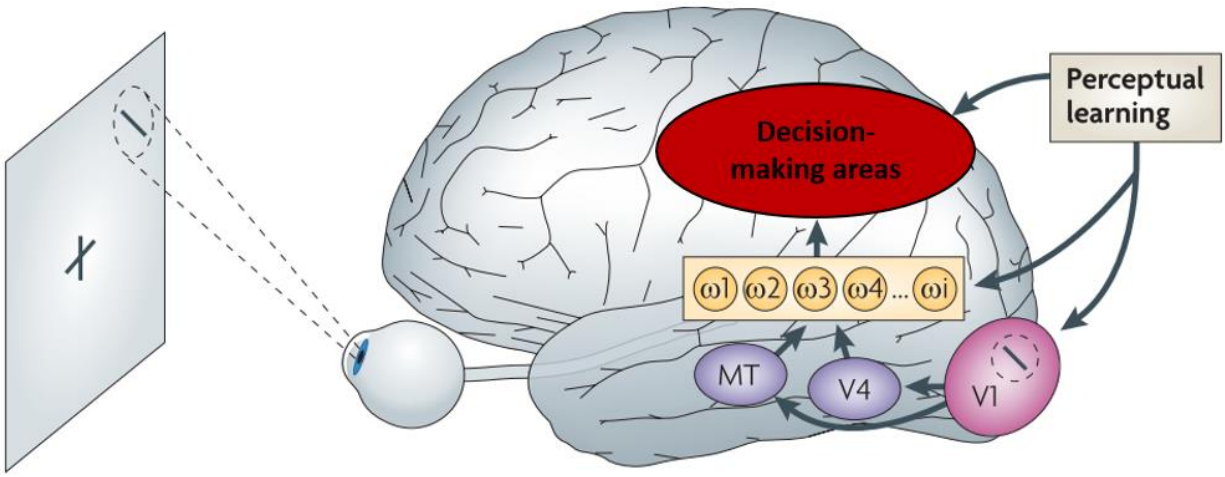


Figure 1.5. Dosher & Lu (1998) theory of perceptual learning. Task-specific reweighting of connections between early visual areas, such as V1, and higher order decision making centers of the brain results in learning as a result of visual training. Figure adapted from Sasaki et al. (2009).

Dosher & Lu (1998) investigated perceptual learning in a peripheral orientation discrimination task with varying levels of stimulus noise overlaid on an oriented Gabor patch. Improvements in the absolute and contrast thresholds of the stimulus were attributed to reductions in internal noise of the visual system, and external noise of the stimulus respectively. It was reasoned that

perception is moderated by a series of spatial frequency and orientation channels, and perceptual learning strengthens task-appropriate channels by reducing the weight of irrelevant channels. The strengthening of task-appropriate channels likely occurs in areas of the brain associated with attention and decision making.

The two theories of the mechanisms of perceptual learning described here are not likely to be mutually exclusive. Fine & Jacobs (2002) point out that the reweighting of neuronal connections will certainly involve cortical changes, but that these modifications are dependent on the context of training conditions. Indeed, very little learning is witnessed in tasks that train on low-level stimulus features, and more learning is associated with tasks that involve simultaneous discrimination along multiple stimulus features or complex stimuli (Fine & Jacobs, 2002). Thus, cortical retuning or reweighting might be selective for the appropriate stimulus and task conditions (Ahissar & Hochstein, 1997).

1.10 Thesis Outline

In this thesis, I outline an approach to approximate cortical plasticity and perceptual learning in patients with electronic sight restoration technologies. Using sighted participants as 'virtual patients', I describe an approach to mimic the visual distortion of simultaneous stimulation of on- and off-cells (see **section 1.7.1**). By mimicking this distortion, it is possible to study the potential of sighted participants to learn to adapt to the abnormal visual input created by electronically or optogenetically stimulating the retina or visual cortex.

In Chapter 1, I introduce a Fourier filtering method of an image and its contrast-reversed complement to produce two images which, when viewed by participants, induces a similar distortion to simultaneous on- and off-cell stimulation caused by sight restoration technologies. I

then investigate the potential for sighted participants to learn to recognize objects using this distorted input in a standard psychophysical task. Mechanisms of learning are also investigated using pre- and post-transfer of learning tests.

In Chapter 2, I designed a ‘gamified’ version of the object discrimination task to ask whether learning to discriminate the distorted visual input described in Chapter 1 generalizes to new objects in a different context. I also introduce modifications to pre- and post-testing that answer key questions which arose from the results of Chapter 1. The results of these two studies are then discussed in the context of the plasticity potential of patients with electronic and optogenetic implants, in addition to insights from learning strategies during training that inform basic science perspectives on visual perceptual learning.

Chapter 2. LEARNING TO SEE AGAIN: PERCEPTUAL LEARNING OF SIMULATED ABNORMAL ON- OFF- CELL POPULATION RESPONSES IN SIGHTED INDIVIDUALS

2.1 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 et al., 2014; Ferlauto et al., 2018; Fujikado et al., 2016b; Genovesi-Ebert et al., 2014; Hornig, 2017; Lorach et al., 2015; Daniel 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; Murphey, Maunsell, Beauchamp, & Yoshor, 2009) or developing implants for cortical stimulation (X. Chen et al., 2020; Troyk, 2017). Gene therapy has been approved for Leber congenital amaurosis, a photoreceptor disorder, and over 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 two years (Liu, Fattah, & Degenaar, 2020). Multiple other technologies such as stem-cell transplantation are also in development (Cuevas, Parmar, & Sowden, 2019; Garita-Hernandez 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 will differ 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 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, in order to generate a firing pattern whose predicted percept closely resembled the original white noise image. While 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 CSNB (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 logMAR (Snellen 20/40) (Zeitz et al., 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 CSNB1 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 vs. 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 tenfold drop 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). Retraining with a visual stimulus then interfered with detecting electrical stimulation, suggesting that optimal detection of electrical stimulation required a long-term reconfiguration of the decoding of neuronal responses within V1. Adaptation to electrical stimulation was slow (over 10,000 trials), and both training and learning was 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, et al., 2017; Ghose, Yang, & Maunsell, 2002). The reason for this is still unclear, but one possible explanation is that, compared to 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 to 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 half of the total combination of spatial frequencies and orientations in I were passed through. Convolving 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 2D 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.

While it is impossible to recreate the effects of electronic retinal or cortical stimulation in sighted individuals, the dichoptic stimuli described above create a visual stimulus that similarly 'scrambles' the cortical input. **Figure 2.1** shows simulated cortical responses (see Methods) for three example stimuli: a natural binocular image (**Panel A**), monocular electronic retinal stimulation (**Panel B**), and the filtered dichoptic images used in our experiment (**Panel C**). The upper images of panels A-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 panels A-C show cortical responses for ~32,500 typical V1 cells, distributed evenly over 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. **Panels D-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 (**Panels A**, **B** and **D**). Similarly, the population responses produced by our filtered dichoptic stimuli are also only weakly correlated with naturalistic population responses (**Panels A**, **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, while others were anti-correlated. Nonetheless, these simulations show that 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 (**Panels B**, **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 = ~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 > ~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 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.

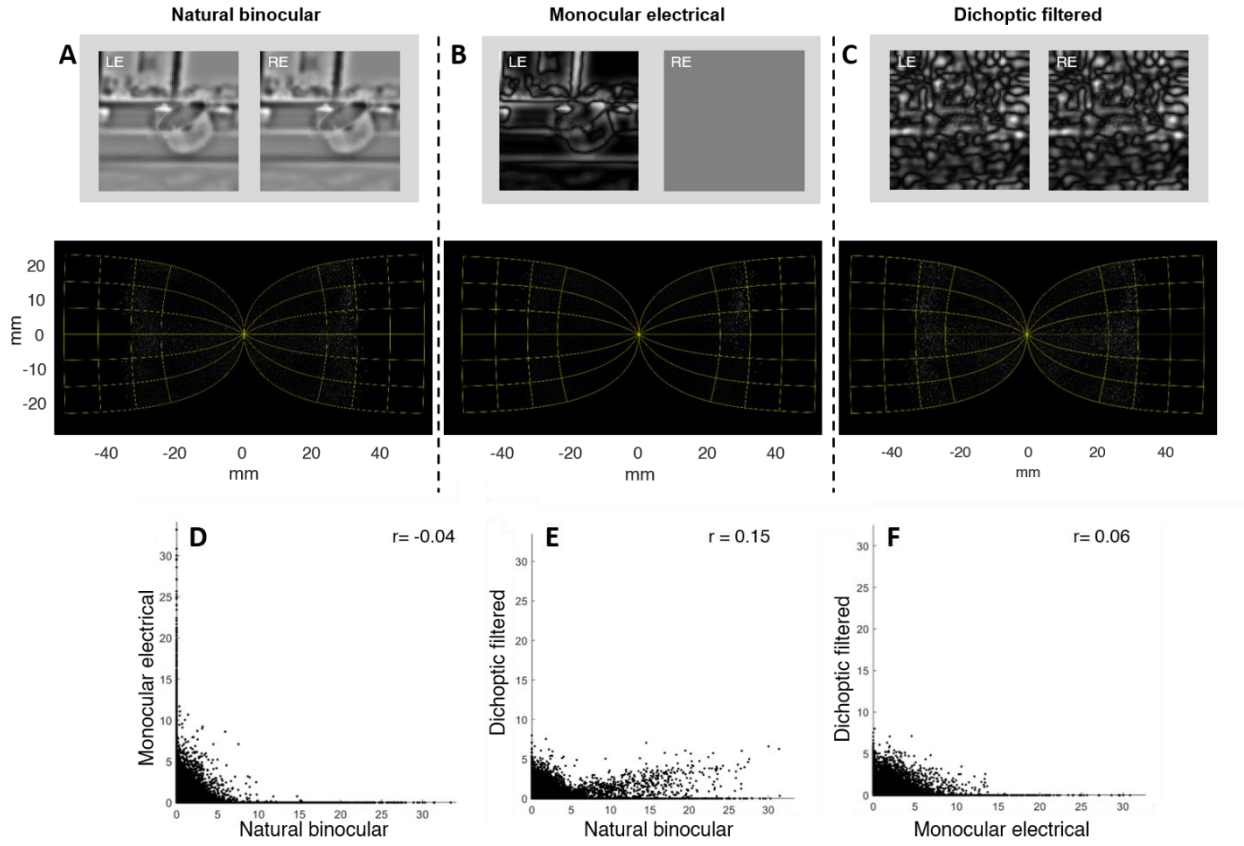


Figure 2.1. A-C) Examples of simulated cortical responses for three stimuli: the original binocular image, monocular electrical retinal stimulation, and the dichoptic filtered images used in our experiment. The 'cortical input image' for each panel is shown as an inset. For all types of stimulation, 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.

Participants in our study showed significant improvement in object discrimination over 14, one-hour training sessions. Transfer of learning in pre- 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.

2.2 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.

2.2.1 *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 at least 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.

2.2.2 *Stimulus and Procedure*

Stimuli were presented dichoptically using a custom-built stereoscope which consisted of two cold mirrors mounted on posts, rotated at a 45 degree 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 x 1440 pixel resolution projected to each cold mirror. Each monitor spanned 28.9 degrees, and these

monitors were the only light source in the room. Each stimulus spanned 768 x 768 pixels (8.84 degrees). Stimuli had a mean luminance of 132 cd/m^2, and were presented on a mid-grey background with luminance 80 cd/m^2.

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.

Stimuli were created by manipulating the spatial frequency and orientation information of naturalistic scenes in the Fourier domain, see **Figure 2.2 and Figure 2.3**.

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 2.4B**). To minimize image-specific learning, objects could take on one of six possible logarithmically spaced sizes ranging from 22.2% (1.96 degrees) to 66.7% (5.90 degrees) of the original object (768 x 768 pixels or 8.84 degrees). Each object was randomly located within the scene, and was rotated by up to 30 degrees in either direction. Thus, there were over 13 thousand 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 2.2**). 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.

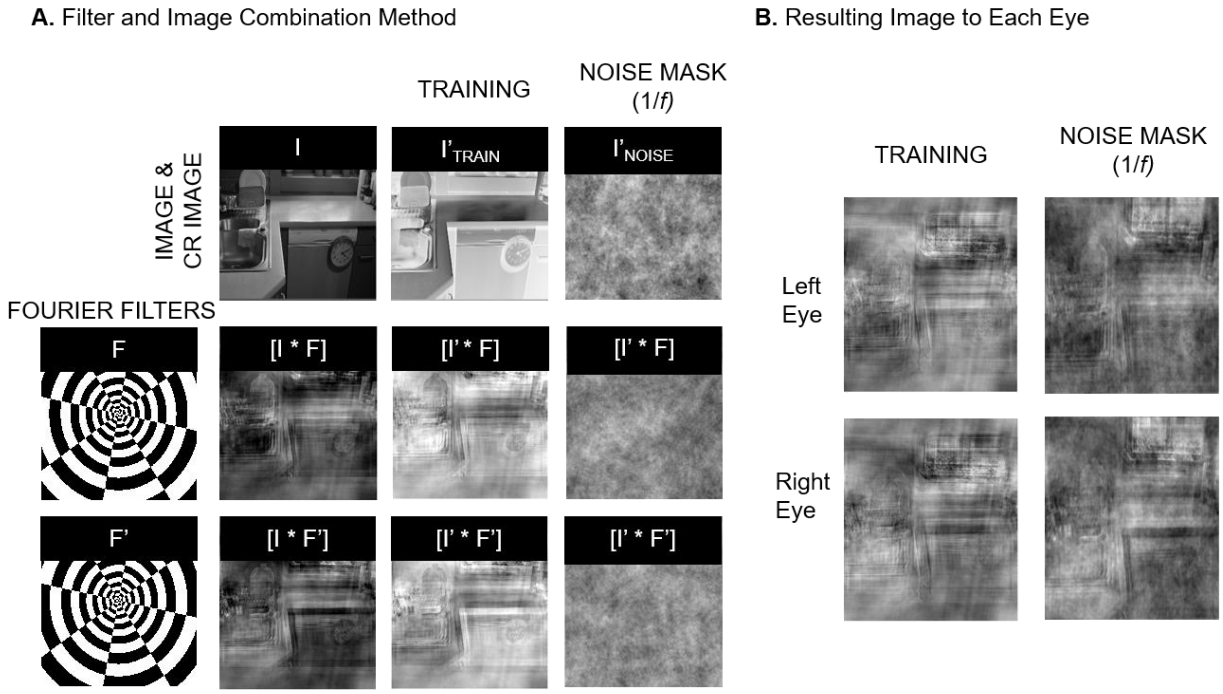


Figure 2.2. Example of filtering for dichoptic presentation. **A)** The two upper-left panels show an example scene (I), and the contrast-reversed version of that scene (I'). The upper-right panel shows the noise mask 1/f NOISE. The leftmost panels show two filters: **F** and **F′**. Filters images represent 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.

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 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_ϑ, was defined as: $F_\vartheta = \alpha_0 . \alpha$, where α is the orientation of the Fourier image, and α_0 defines the number of radial spokes.

The radial checkerboard filter in was built with sharp edges in the Fourier domain, which leads to image artifacts – often seen as 'ringing' in the spatial domain (see **Figure 2.3**). 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 (RMS) contrast of ~1/3 that of the original images, see **Figure 2.3B**) and was unlikely to be the primary cause of masking. Pilot data (not shown) from 3 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_ϑ) and spatial frequency (*n*), and radial spokes (α_0) of our filters, resulted in strong 'striping'; (horizontal striping can be seen in [I*F] and [I*F'] in **Figure 2.2**). 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 $F_{\square}$, with one being the negative of the other: F = Ff × F_9, and F' = $-Ff$ × 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 2.2** 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 2.2A** show original (left), contrast-reversed (middle) and 1/f images (right). The bottom two panels (**Figure 2.2A**) show the 4 examples of possible filtering: I * F, I * F′, I'*F, and I'* F′ (where * denotes 2D convolution) for both the training and 1/f noise stimuli.

In the training paradigm, we presented the left eye (**Figure 2.2B, Figure 2.3**) the sum of two filtered images, [I * F′] + [I' * F], such that half 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 half the information from 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.

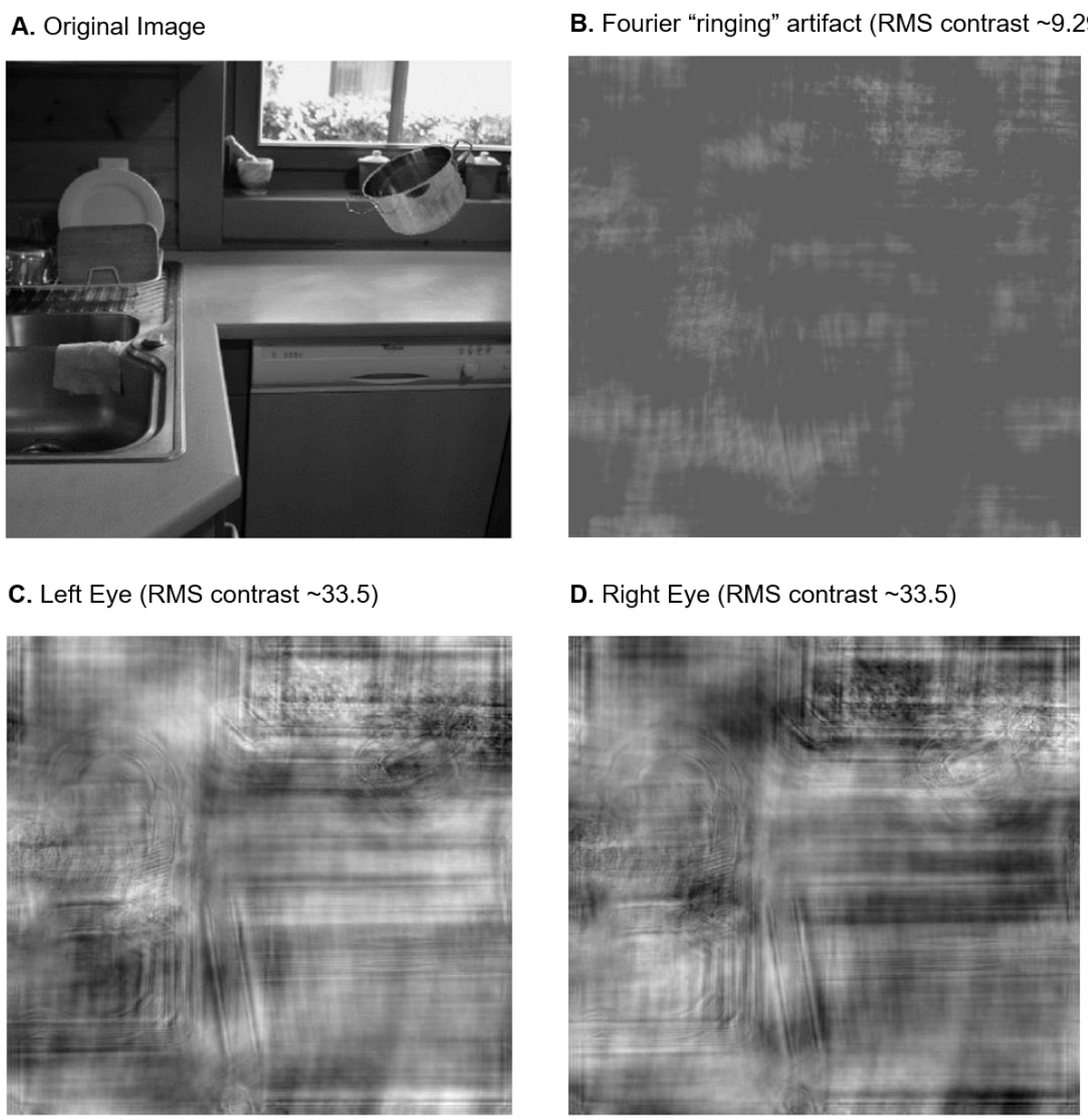


Figure 2.3. A) Unfiltered image with a pot (upper right corner) overlaid as part of the object discrimination task. **B)** The Fourier artifact. **C, D)** Example of the resulting dichoptic stimuli after the filter was applied to the top image (see Filtering section of the main text).

Task. A brief fixation cue (0.5 s) began each trial (**Figure 2.4A**). After a 0.5 s pause, a word cue told the participants what the target object was (e.g. “cup”, “clock”). Following the word cue, a

scene with an overlaid object was displayed for up to 2s, 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 2s trial. The field of view drifted to the right or left, at a rate that was uniformly distributed between 0.21 and 0.52 degrees/s. The image also expanded or contracted at a maximum rate of 0.35 degrees/s.

Participants performed a two-alternative forced-choice (2AFC) 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.

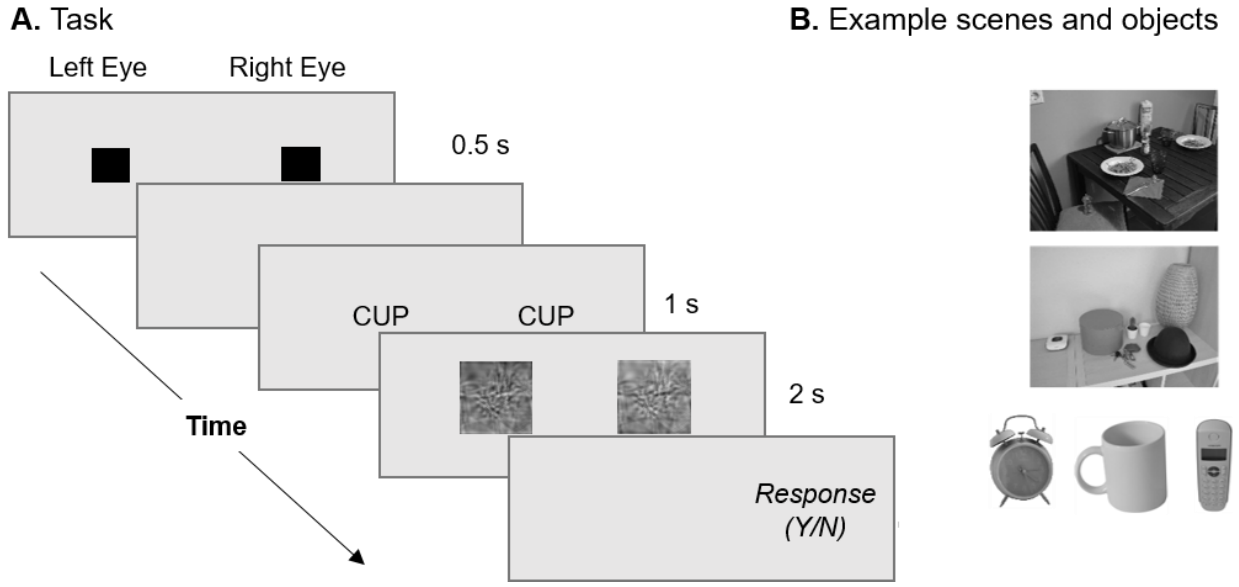


Figure 2.4. A) The object discrimination task. On each trial, the participant reported whether or not the cued object was present within the scene. **B)** Examples of unfiltered scenes and objects from the SCEGRAM database.

2.2.3 *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 3 sessions contained 100 trials of the training stimulus set, and 100 trials of each of the three (monocular, filter-switched, 1/f noise) pre- and post-test conditions, for a total of 400 trials. The remaining 14 sessions consisted of 400 trials of the training set.

2.2.4 *Pre- 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’].

2.2.5 *Statistical Analyses*

Hits, misses, correct rejections and false alarms from conditions within each session of 2AFC trials were converted into d-prime (*d’*) units (D. M. 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 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 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 4 levels consisting of all types of pre/post-tests, and (2) 'time', which assessed the progression of learning between pre and 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- and post-test conditions, compared to the trained stimulus.

2.2.6 *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 2.1**).

We modeled retinal receptive fields as a bank of circular center-surround difference of Gaussian filters (with 4 sizes), with centers fixed at 2x 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, Nanduri, 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 cortex, ocular dominance columns and orientation pinwheels were simulated based on Rojer and Schwartz (Rojer & 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 2.1**), we projected the entire scene onto the retina, assumed participants were fixating centrally, and assumed no neural noise.

When modeling monocular vs. binocular performance on the task (**Figure 2.7**) we first calculated noise-free cortical response to the filtered target ('POT') as a 1 x *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 ~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 2.7**) by simulating 1,000 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 vs. monocular presentations.

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

2.3 RESULTS

2.3.1 *Performance improvements over time for the trained stimulus set*

Figure 2.5 shows *d'* values for each session, for each individual participant. Panel A 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% 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 pre- 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 2.5B shows individual differences in the rate of learning. For each participant the x-axis represents the testing session and the y-axis (Learning Index) 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 5 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).

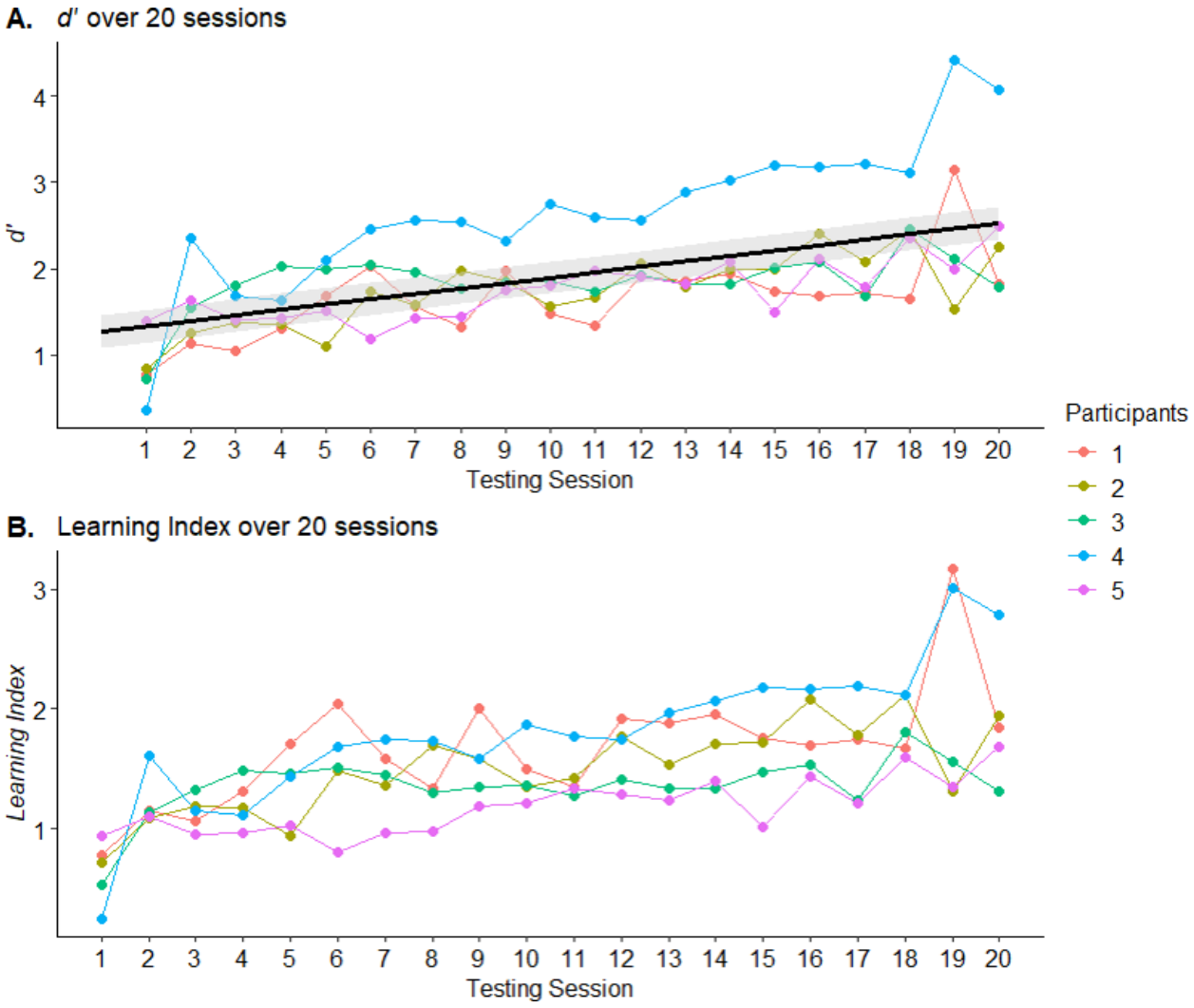


Figure 2.5. **A)** *d'* scores for participants in the trained stimulus set. The regression line for data pooled across participants (black) is overlaid on individual participant scores. **B)** The rate of learning for all subjects, normalized to the average *d'* of the three pre-test sessions. A Learning Index > 1 shows better performance than the average *d'* of the participant's three pre-test sessions, a Learning Index < 1 shows worse performance than the average *d'* of the three pre-test sessions.

2.3.2 *Pre and Post-tests*

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

Results of a full model assessing the four stimulus conditions (*trained, monocular, filter-switched,* and *1/f noise*) in the pre- and post-testing phase 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- 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- and post-test conditions compared to 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 to performance in the training condition during the pre-test ($M_{\text{Training}} = 1.292$, $SD_{\text{Training}} = 0.500$) and post-test ($M_{\text{Training}} = 2.510$, $SD_{\text{Training}} = 0.848$), see **Figure 2.6**.

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 [$F(1,52) = 1.455$, $p = 0.232$, $\eta_p^2 = 0.030$] or interaction between condition and time [$F(1,52) = 0.004$, $p = 0.951$, $\eta_p^2 < 0.001$]. Thus, performance was very similar for the monocular and training condition, in both pre- ($M_{\text{monoc}} = 1.139$, $SD_{\text{monoc}} = 0.447$) 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- to post-test [$F(1,52) = 73.821$, $p < .001$, $\eta_p^2 = 0.590$]. There was no significant main effect of condition [$F(1,52) = 0.002$, $p = 0.963$, $\eta_p^2 < 0.001$] or interaction between condition and time [$F(1,52) = 0.275$, $p = 0.602$, $\eta_p^2 < 0.001$].

1/f noise

1/f noise was a less effective mask than contrast-reversed filtered information; participants had relatively high *d'* for both pre- ($M_{1/f\ \text{Noise Pre}} = 2.305$, $SD_{1/f\ \text{Noise Pre}} = 0.740$), 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 2.6**), with a significant improvement in performance between pre- 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- to post-test [$F(1,55) = 4.463$, $p = .039$, $\eta_p^2 = 0.08$].

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

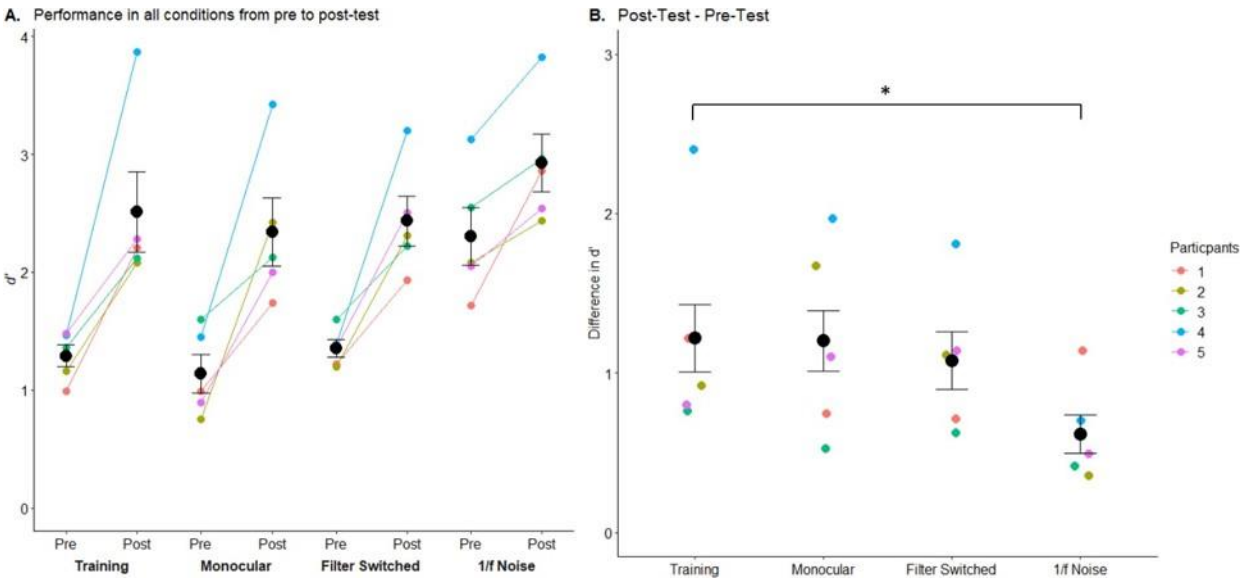


Figure 2.6. A) *d'* scores for each pre/post-test condition. Each pre- and post-test d' is calculated as the average of each participant's three runs in each test. Large black data points represent the average *d'* (across subjects) for that condition. Error bars represent standard error of the mean 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 to 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 to the 1/f noise condition.

2.4 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 to 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 pre-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 to 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).

2.4.1 *Sighted participants vs. prosthetic patients*

One limitation of this study is that there are, of course, major differences between our training protocol in sighted participants and the experience of prosthetic patients. One major difference is that prosthetic patients have access to distorted information for much more than 1 hour/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/pixelated input is a factor.

A second difference is that for prosthetic patients the alternative to distorted input is no input, whereas our virtual patients spend most of their day with normal vision. This 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 co-exist 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.

2.4.2 *Fast vs. 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 et al., 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 to learning rates for low-level properties, such as perceptual judgements 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).

2.4.3 *Monocular vs. 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 drop 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 vs. 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 2.7** shows simulated histograms representing these distributions for both binocular and monocular simulations. The *d'* values were consistently, but only very slightly (less than 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.

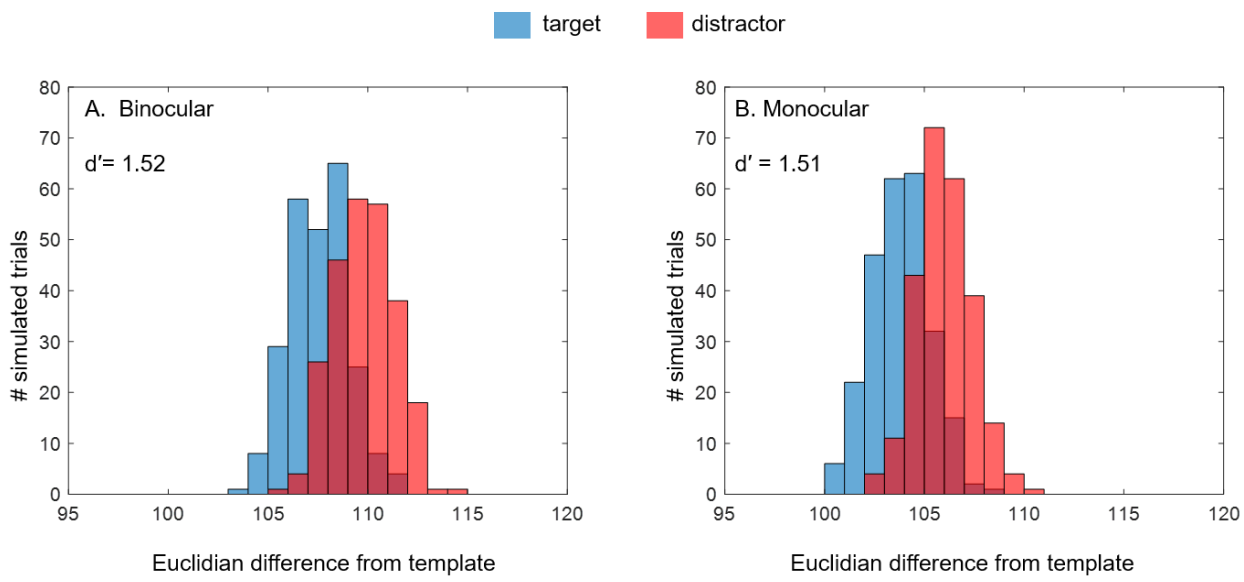


Figure 2.7. Example histograms showing the Euclidian norm of the difference between the noise-free cortical response to a target object ('POT') and noisy cortical responses to the target or a distractor ('CLOCK'). 1,000 trials were simulated for each condition. Note that, unlike most "signal" and "noise" representations, a small Euclidian norm represents 'good performance', where the cortical response is similar to the perceptual template.

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.

2.4.4 *1/f noise*

Performance was initially much better in the 1/f noise condition. As described in **Figure 2.2**, in the training 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 below, 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 2.8**). 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 2.8**).

We did not see full transfer of learning to the 1/f noise condition. One possibility is that this 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 ($M_{1/f\ Noise\ Post}$ = 2.925; **Figure 2.6**). 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.

2.4.5 *Models of learning*

Three possible learning models are schematized in **Figure 2.8**. 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- 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 reduced 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.

While 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.

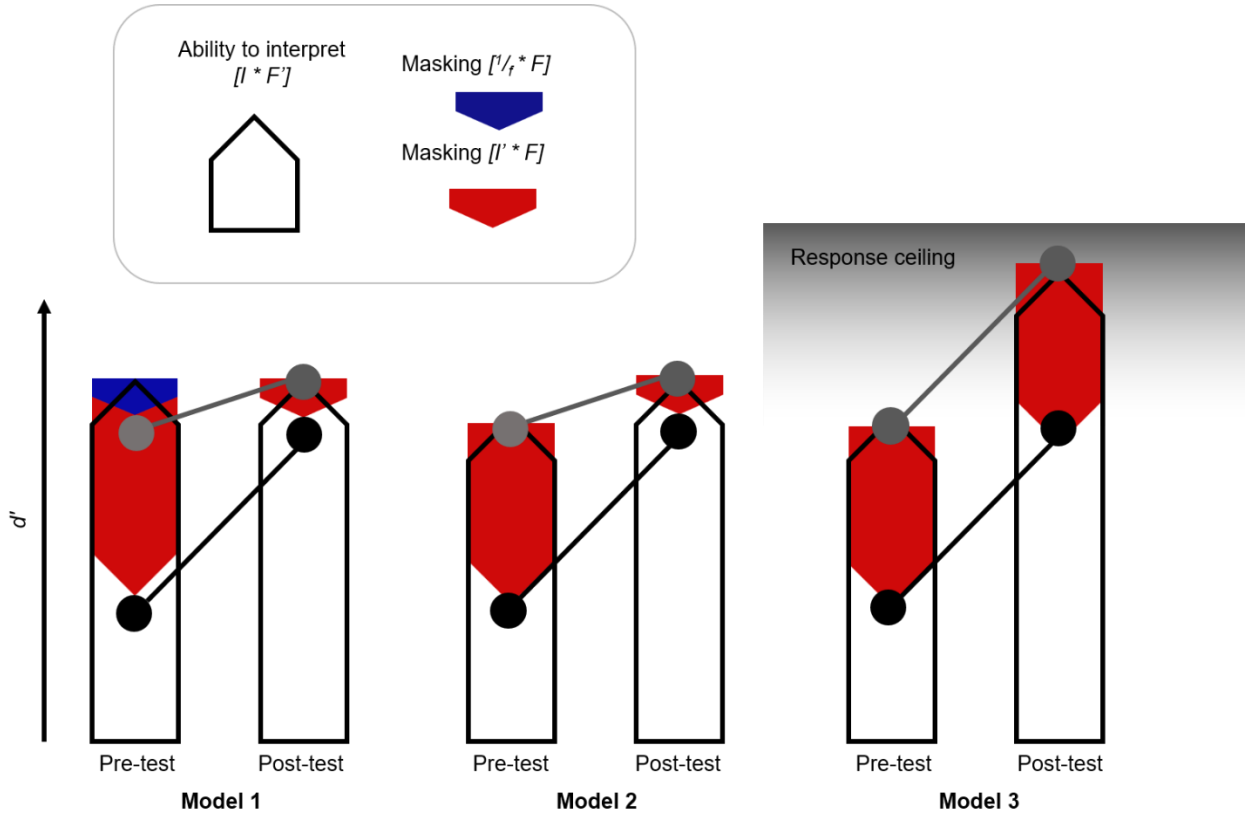


Figure 2.8. Schematic of possible sources of learning.

2.5 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 (Baseler et al., 2002, 2011b; Corinna Darian-Smith & Gilbert, 1994; Hiroshi et al., 2015; Masuda, Dumoulin, Nakadomari, & Wandell, 2008; Smirnakis et al., 2005), within one eye (Lunghi, Berchicci, Morrone, & Di Russo, 2015; Lunghi, Burr, & Morrone, 2011), or by removing orientation or spatial frequency information (Georgeson & Sullivan, 1975; Haak et al., 2014; Webster, Georgeson, & 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.

Chapter 3. PERCEPTUAL LEARNING OF INPUT THAT SIMULATES ABNORMAL ON-OFF-CELL POPULATION RESPONSES IS CONTEXT SPECIFIC: INSIGHTS FROM VIDEO GAME TRAINING

3.1 INTRODUCTION

In **Chapter 2**, we tested the hypothesis that plasticity might compensate for abnormal on- and off-cell population responses in visual cortex (Esquenazi et al., 2021). We dichoptically presented stimuli to sighted participants, which were designed to elicited abnormal cortical population responses that were somewhat similar to the distorted input caused by sight restoration technologies. Our sighted participants exhibited remarkable performance improvements in an object discrimination task with distorted input over fourteen training sessions. These results provided initial evidence that it may be possible for patients to adapt to the unnatural on- and off-cell population responses likely to be produced by electronic and optogenetic sight recovery technologies.

Here, we extend our previous research in three ways. First, we wondered if it was possible that a more engaging and dynamic training regimen – playing a video game – might speed up the rate of learning witnessed previously. Second, we asked whether training to discriminate objects with a video game would transfer to a novel set of objects in a discrimination task. Finally, we developed additional pre- and post-tests to examine the mechanisms underlying learning.

A standard visual training paradigm consists of naïve observers repeatedly viewing a stimulus and making a judgment about it. Learning of any given stimulus dimension (e.g. contrast, orientation,

spatial frequency, and direction of motion) occurs over multiple trials and sessions, and almost always involves improvement in sensitivity (for a review see Fine & Jacobs, 2002). However improvements are typically restricted to the trained stimulus (Crist et al., 1997; Fahle, 2005; Karni & Sagi, 1991) and task parameters (Ahissar & Hochstein, 1997; Jeter, Dosher, Petrov, & Lu, 2009). A second disadvantage of these paradigms is that they are inherently monotonous and time consuming: several hours of engagement in a task occasionally leads to only meager improvements for the trained stimulus (e.g. Matthews, Liu, Geesaman, & Qian, 1999). So, merely repeating a fatigue-inducing task over multiple hours may not be the most effective form of training.

A good example of how recovery from visual loss might exploit more engaging training models is in the treatment amblyopia in children, who are generally unlikely to have the motivation and attention to comply with traditional paradigms (Hussain, Astle, Webb, & McGraw, 2014). 'Gamification' of perceptual learning using video game play as a training tool has proven to be very useful for visual training. Playing action video games which feature fast-moving objects, a high degree of perceptual, cognitive and motor load, unpredictable events, and significant peripheral processing (Green, Li, & Bavelier, 2010) has resulted in improvements in visual skills such as spatial resolution (Green & Bavelier, 2019), contrast sensitivity (R. Li, Polat, Makous, & Bavelier, 2009), visual search (Castel, Pratt, & Drummond, 2005), selective attention (Green & Bavelier, 2003) and object tracking (Trick, Jaspers-Fayer, & Sethi, 2005). The cortical mechanisms of improvement are thought to be two-fold: for one, action video games result in a highly engaged reward system through activation of dopaminergic pathways (Bavelier & Green, 2019). Two, action video games actively engage attentional pathways that aid in the selection of task-relevant information (Bavelier & Green, 2019). Activation of these reward and attentional

systems seem to place the brain into a plastic state, and subsequently enhances the ability to refine and perfect perceptual skills.

In the current study, we gamified our original object discrimination training task by creating a video game that resembled the popular game Fruit Ninja™. Participants were required to interact with fast-moving objects that unpredictably appeared on screen, focusing attention on identifying and tracking targets, while ignoring known distractors. We predicted that playing a video game with distorted input in real time might increase the rate of learning compared to that in our previous study, which relied on a traditional perceptual learning methodology (Esquenazi et al., 2021). Further, given the large body of evidence that game play enhances performance on a wide range of psychophysical tasks, we predicted that game play in this context might *transfer* to a novel set of objects in the object discrimination task used in the original experiment.

After completion of a pre-test, participants played our custom-coded video game for 25 hours. Every five hours of video-game play, participants completed an object discrimination task, which served as a measure of learning transfer. A separate control group of participants did not play video games, and instead completed the object discrimination task on roughly the same timeline as those in the video game group. Results reveal that while performance in the video game improved over time, only a very small amount of learning transferred to the object discrimination task. This evidence suggests that learning to interpret the unnatural on- and off-cell population responses likely to be produced by electronic sight recovery technologies may be very context specific. Pre- and post-training results further provide an indication of what mechanisms underlie improvements in the object recognition task.

3.2 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.

3.2.1 *Participants*

Twenty naïve observers (16 males) aged 19 to 32 years (M = 25) were recruited through word-of-mouth at the University of Washington. Participants were separated into two groups: gaming (N = 10) or control (N = 10). Binocular and monocular visual acuity was assessed using the same methods described in **Section 2.2.1**.

3.2.2 *Stimulus and Procedure*

Stimuli were presented dichoptically using methods previously described in Esquenazi et al. (2021). A custom-built stereoscope, consisting of two cold mirrors mounted on posts, rotated at a 45-degree angle to capture input from two separate 32" LED monitors and reflect it separately into each eye. Each monitor spanned 28.9 degrees, and these monitors were the only light source in the room. Each stimulus spanned 768 x 768 pixels (8.84 degrees). Monitors had a mean luminance of 80 cd/m^2. The stimuli had a mean luminance of 132 cd/m^2. Participants completed a nonius task at the beginning of each session to align the screens and account for any fixation disparities.

The apparatus consisted of two computers. The first computer ran the video game. A KONA 4-Channel HDMI Capture Card was used to stream the HDMI output of the gaming computer to the stimulus processing computer at 30 Hz. Onboard the stimulus computer, the capture card directly passed each frame to a high-powered NVIDIA Quadro RTX 6000 graphics processing unit (GPU) without the involvement of the central processing unit, preserving the 30 Hz stream rate with

minimal lag (~1 frame). The GPU filtered each frame (as described below) using the Compute Unified Device Architecture (CUDA), a parallel programming framework developed by NVIDIA.

3.2.3 *Test Stimuli for Object Recognition Task*

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 objects (e.g. clock, rolling pin, dish soap) were overlaid onto each scene. Each phase of the study (pre-test, test for learning transfer, and post-test) consisted of a different set of 45 objects, for a total of 135 unique objects. To minimize image-specific learning, objects could take on one of six possible logarithmically spaced sizes ranging from 22.2% (1.96 degrees) to 66.7% (5.90 degrees) of the original object (768 x 768 pixels or 8.84 degrees). Each object was randomly located within the scene and was rotated by up to 30 degrees in either direction. Thus, there were over 13 thousand unique images in each object-scene set.

3.2.4 *Filtering*

Our filtering procedure was exactly the same as described in **Section 2.2.2**. Briefly, binarized radial checkerboard filters in Fourier space were used to present separate spatial frequency and orientation information to each eye (**Figure 2.2**).

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 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_ϑ, was defined as: $F_\vartheta = \alpha_0 . \alpha$, where α is the orientation of the Fourier image, and α_0 defines the number of radial spokes.

The final filters were the product of F_f and $F_{\square}$, with one being the negative of the other: F = Ff × F_9, and F' = $-Ff$ × F_9, Lastly, 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.

As described in section 2.2.2, the monocular input contains half the information from the original image, and half the information from 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.

3.2.5 *Video Game*

Participants in the gaming group played a custom-coded game using Godot open-source software (Linietsky, 2014). The game was modeled after the well-known video game Fruit Ninja™, in which fruits (apple, banana, etc.) are hurled onto the screen at various speeds. Players slice the fruits as they appear on screen with a blade controlled via touch screen, while avoiding slicing distracting bombs that would end the game.

The goal of our game was similar to Fruit Ninja: identify objects (such as 'skateboard' and 'baseball') launched on screen, while avoiding distractor objects. Prior to the start of each round, players were given the name of an object to avoid, while being tasked with slashing as many

objects, except the named object, as possible. The video game object set was different than that of the object sets in every other phase of the study.

In any given round, multiple objects were launched onto the screen. Of forty-five possible objects, a unique set of five possible objects was randomly selected before the beginning of each round. Object sizes varied depending on the gamer's level: as the player advanced, the size of the object decreased logarithmically between ~2-6 degrees (see Stimulus and Procedure). Each object was launched into the scene from the bottom of the screen at a random location, and was rotated by up to 30 degrees clockwise, or counterclockwise.

Players always started each session with three lives at level one. At the beginning of each round, participants received directions to avoid one object (distractor object) and slash as many other objects as possible (target objects). While participants did not know what target objects would appear on screen, they always knew the distractor object prior to the start of a round. Objects could be destroyed by clicking and dragging the computer mouse through the body of the object to create a slashing motion. If the player slashed a target object, they received positive auditory feedback and gained a 'Target Object Point'. If the player slashed a distractor object, they received negative auditory feedback and gained a 'Distractor Object Point'. Players automatically advanced to the next level if they received 25 Target Object Points before dying. Receiving five Distractor Object Points caused the level to restart, and the player lost a life. When three lives were lost, the game automatically restarted at level one, and players received a 'Game Over' notification.

The difficulty of the game varied within and between levels. As any given round progressed, objects were launched onto screen with increasing velocity. When the player advanced a level, the size of the objects decreased. Once the player reached level 4, they were given two distractor objects to avoid, and thus had just four target objects to slash. In total, there were six levels of

difficulty. If players beat the game in any given training session, they were instructed to restart the game at level one.

3.2.6 *Object Discrimination Task*

Transfer of learning to an object discrimination task was assessed every five video game sessions. The object discrimination task was exactly as described in **section 2.2.2**. A brief fixation cue (0.5 s) began each trial (**Figure 2.4A**). After a 0.5 s pause, a word cue told the participants what the target object was (e.g. “cup”, “clock”). Following the word cue, a scene with an overlaid object was displayed for up to 2s, or until the participant responded with a key press.

Participants performed a two-alternative forced-choice (2AFC) object discrimination task, judging whether 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.

3.2.7 *Learning Protocol*

The study was organized into three phases: pre-test, training, and post-test. Participants in the control and gaming groups completed three sessions each of the pre- and post-tests, which were identical. Participants in the gaming group completed 25 video game sessions, each lasting one hour. During gaming sessions, participants played the video game for at least 45 minutes, with 15 minutes allotted for a break. Video game sessions were conducted on separate days with an average of 3 sessions per week.

Every five sessions of gaming, participants paused game play and completed one session of the object discrimination task, which served as a test for learning transfer. Participants in the control

group did not play video games, and instead completed five sessions of the object discrimination task. The control group completed the object discrimination task every seven to ten days, so that each experimental group completed the object discrimination task in roughly the same time-frame.

3.2.8 *Pre- and Post-test Conditions*

The pre- and post-test consisted of the same object discrimination task used in the learning transfer test, each with a different set of 45 unique objects overlaid onto each scene. There were five different possible conditions.

Training. The same conditions present during the training phase (see section 2.2.2). The left eye received [I * F′] + [I’ * F], and the right eye received [I * F] + [I’ * F′].

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.

Monocular object. The object was presented to one eye only (randomly interleaved across trials), with the other eye containing the filtered background scene without the object.

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’]. The contrast of the 1/f noise pattern was either matched to the original contrast-reversed image ($N_{control} = 5$, $N_{gaming} = 7$), or double that of the contrast reversed image ($N_{control} = 5$, $N_{gaming} = 3$). We tested these two different contrast levels to understand the effect of the strength of the 1/f noise masking on learning.

3.2.9 *Statistical Analyses*

3.2.9.1 Gaming

Performance improvements during gaming were measured by calculating the average number of times per level that a participant was forced to restart the game because they died three times. Thus, learning was quantified as the average "Game Over Rate" per level as a function of session.

3.2.9.2 Object Discrimination Task

Hits, misses, correct rejections, and false alarms from conditions within each session 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).

3.2.9.3 Transfer of learning

A two-way linear mixed effects regression analysis was carried out to examine performance (in *d'*) on the five sessions of the object discrimination task completed by both the control and gaming groups. Participants were treated as a random factor and the session number was treated as a fixed factor.

3.2.9.4 Performance in pre-test and post-test conditions.

Differences between the monocular, filter-switched, and 1/f noise pre- and post-test conditions compared to the trained stimulus were carried out as three separate linear mixed effects models, with participants entered as a random effect. In each model, two additional factors of time, which examined the progression of learning from pre- to post-test, and group, which examined differences between those in the gaming and control groups were added.

3.3 RESULTS

3.3.1 *Video game play*

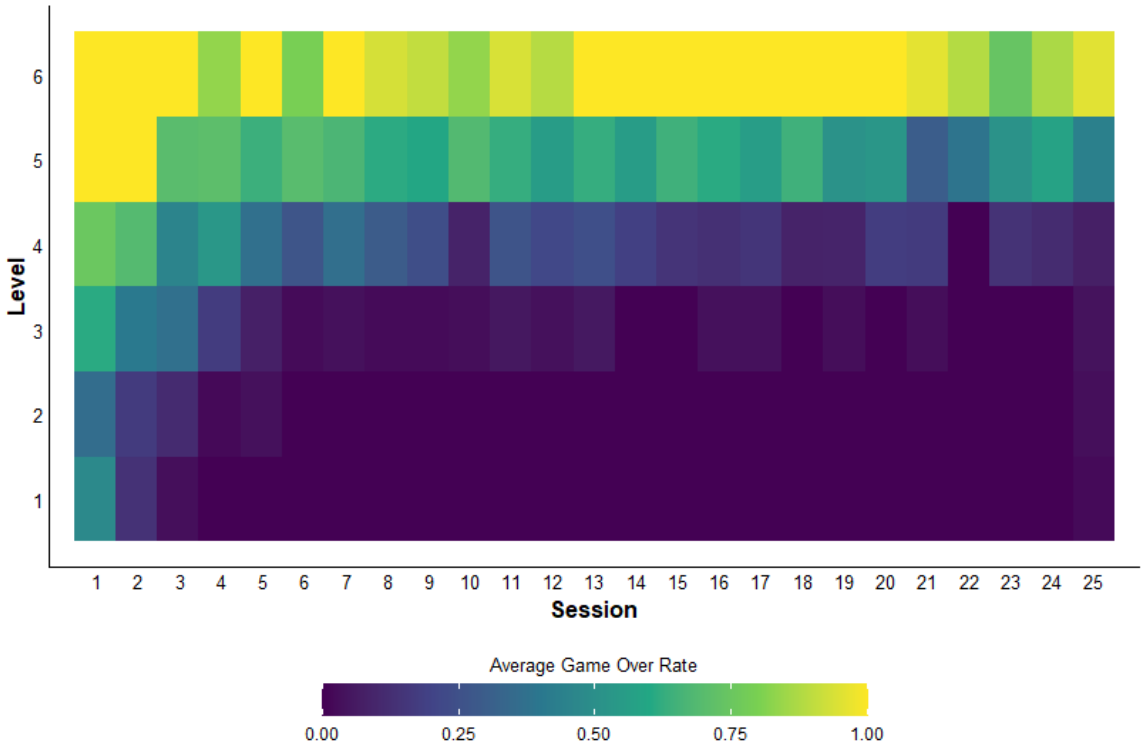


Figure 3.1. The average 'game over' rate (i.e. the number of times gamers lost three lives and had to restart the game), for participants in each level, across all 25 gaming sessions. Yellow and green colors indicate a higher game over rate, while blue colors indicate a lower game over rate.

The number of times a participant died during game play and was forced to restart the game was used as a metric for learning (i.e. "Game Over" rate). **Figure 3.1** shows the game over rate averaged over each participant for each level, in each session. An average game over rate of zero would indicate that participants had entirely mastered that level. **Figure 3.1** shows that participants mastered levels 1-3 fairly quickly. By session 11, the average game over rate for level 4 was consistently less than 0.25. While training did improve performance for levels 5 and 6, average

game over rates remained high (M_{L5} = 0.61; M_{L6} = 0.94). The average game over rate for level 5 declined over time: sessions 13 – 25 resulted in an average game over rate of M = 0.53. However, the game over rate for level 6 consistently remained above 0.9, suggesting that participant performance asymptotes at level 6.

3.3.2 *Transfer of learning to the object discrimination task.*

Figure 3.2 shows *d'* (Panel A) and normalized (Panel B) *d'* scores for participants in both the gaming and control groups, across the five object discrimination sessions. Participant intercepts, which reflect *d'* at session 1, ranged from 0.391 to 2.472 with a mean value M = 1.308, SD = 0.513.

When assessing the raw data, we noticed a considerable difference in *d'* between the control (M = 1.374, SD = 0.473) and gaming (M = 1.828, SD = 0.593) groups on the first object discrimination session. Up to this point, both groups had been exposed to the same pre-testing procedure and, in theory, should not have a statistically significant difference in *d'* scores. To test this, we conducted a Welch's two sample t-test on *d'* scores between the two groups, restricting the dataset to session 1 of the object discrimination task. Group means were marginally significant $t(17) = 1.896$, p = 0.075. This difference could be due to the differing level of commitment between the two experimental groups. Participants in the control group were only required to come into the lab for 11 sessions, as opposed to the gaming group who came in for 36 sessions. Further, the experience of gaming may have been more motivating than a traditional psychophysics experiment.

Another explanation may be that the gaming group was comprised mostly of graduate students in the UW psychology department (N = 8 in the gaming group vs. N = 2 in the control group). In addition to having far more availability to complete a 36 session study, the graduate students in our experiment may have been more motivated and attentive due to their familiarity with the

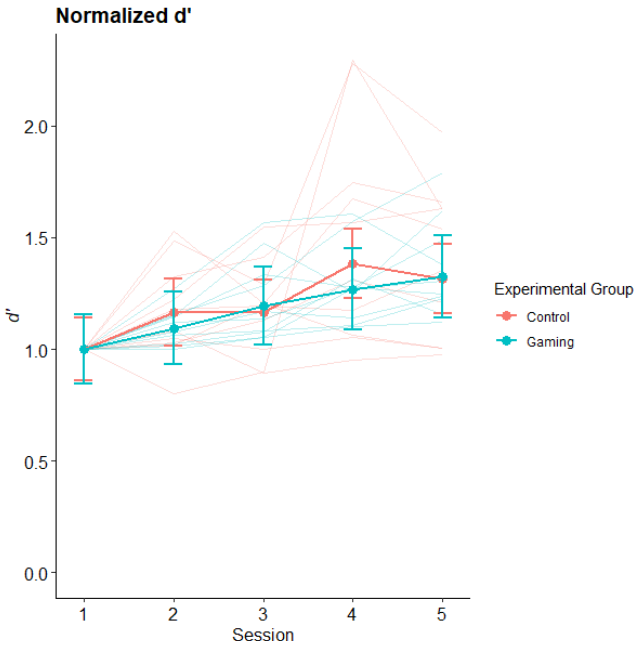


Figure 3.2. Normalized *d'* scores for participants in both the gaming and control groups across the five object discrimination sessions. Thick colored lines represent group means, while thin lines represent individual performance. Error bars reflect the standard error of the mean.

research process and knowledge of the importance of providing "good data" during a psychophysical task. To test this, we assessed a linear model, predicting *d'* scores, using experimental group and graduate student status (i.e. a binary variable for whether the participant was a graduate student at UW) as predictors. Results of an ANOVA showed that the graduate student status predictor approached significance $F(1,16) = 4.274$, $p = 0.055$, $\eta_p^2 = 0.21$. Regardless of whether graduate students were in the control ($M = 1.932$, $SD = 0.267$) or gaming group ($M = 1.921$, $SD = 0.629$), they consistently outperformed undergraduate students in both the control ($M = 1.23$, $SD = 0.408$) and gaming ($M = 1.456$, $SD = 0.216$) groups.

In the main model, there was not a significant difference in slope estimates between experimental groups $m = 0.389$, $SE = 0.255$, 95% confidence interval (CI) = -0.107 – 0.885, $t(22) = 1.524$, $p = 0.142$, $\eta_p^2 = 0.09$, showing that on average, participants in each group performed roughly the same in the five object discrimination tests.

The slope estimate for the session number variable, an indicator of learning rate, was $m = 0.116$, $SE = 0.019$, 95% CI = 0.078 – 0.154, $t(78) = 6.009$, $p < .001$, $\eta_p^2 = 0.55$, indicating that *d'* improved significantly over time, regardless of experimental group. Over the course of 5 sessions, the

predicted *d'* increased from 1.424 to 1.889, a percentage increase of 133%. The interaction between experimental group and session number was not significant $m = 0.035$, $SE = 0.027$, 95% CI = -0.018 – 0.089, $t(78) = 1.284$, $p = 0.203$, $\eta_p^2 = 0.02$, showing that the slopes between the control and gaming groups did not differ significantly over the 5 object discrimination sessions.

3.3.3 *Pre and Post-Test Conditions*

The purpose of the pre-test and post-tests was to elucidate the learning mechanisms that formed the basis of performance improvements over the course of training. We also looked to see whether video game training resulted in the development of different learning strategies, compared to those who did not play video games (despite similar overall levels of learning).

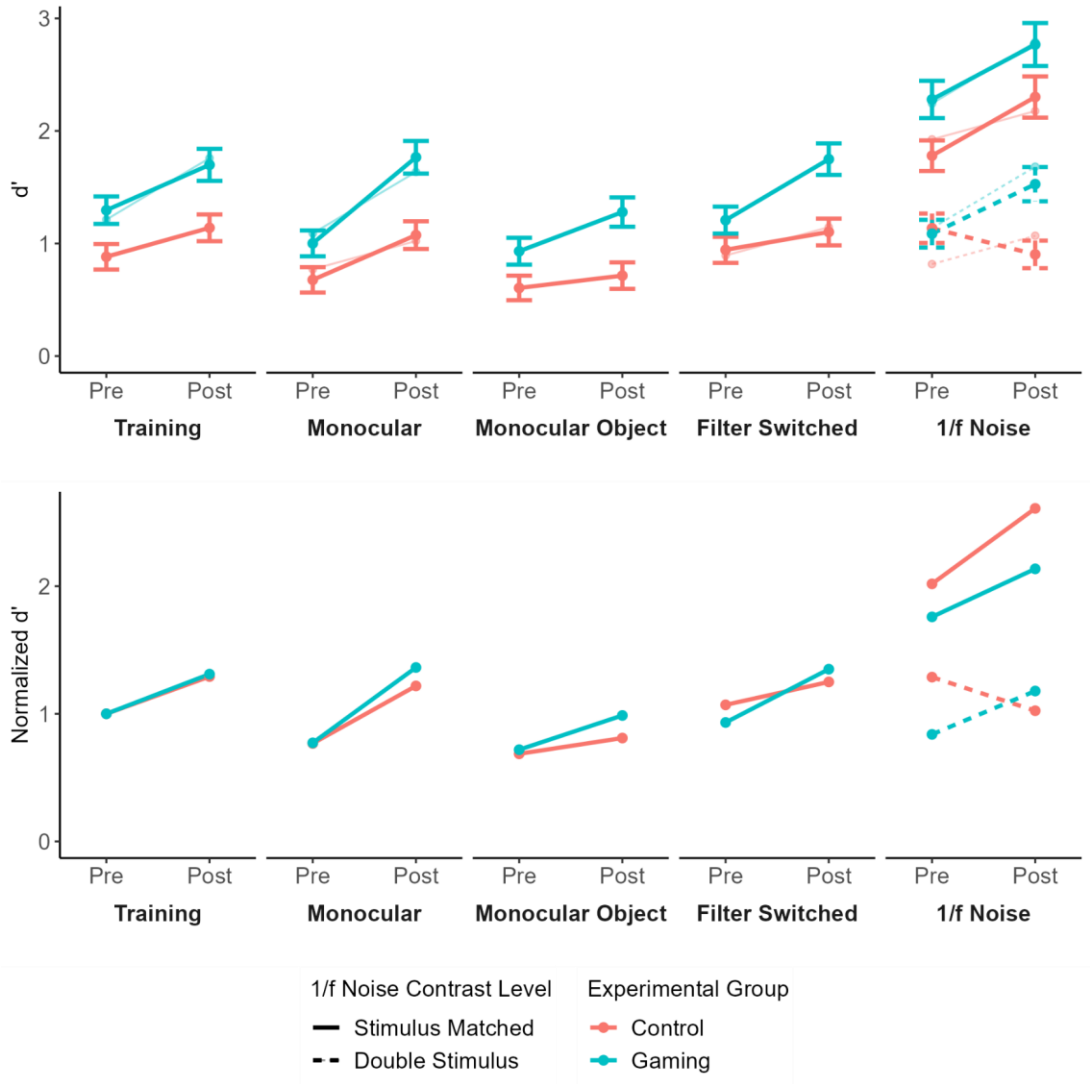


Figure 3.3. Top. Mean *d'* scores for gaming and control participants. Error bars represent standard error of the mean. Faintly colored lines represent average predicted *d'* a model evaluating interocular suppression and masking strength effects (see section 3.3.8 for model description and results). **Bottom.** Mean *d'* scores when each condition was normalized to pre-test performance in the training condition. Scores greater than a *d'* of 1 indicate better performance in that condition, compared to the training condition in the pre-test. A *d'* score below 1 indicates reduced performance in that condition, compared to the training condition in the pre-test.

Analyses were conducted on for each pre- and post-test condition as separate models. **Figure 3.3** displays group *d'* means for each condition from pre- to post-test for both gaming and non-gaming groups.

Because our goal was to test transfer of learning from the trained condition, performance in each condition was compared to performance in the training condition in the pre-test (control group: $M = 0.882$, $SD = 0.391$; gaming group: $M = 1.296$, $SD = 0.473$) and the post-test (control group: $M = 1.139$, $SD = 0.394$; gaming group: $M = 1.699$, $SD = 0.653$). For each condition analyses were conducted using a linear mixed effects model with participant entered as a random effect. We also assessed main effects of experimental group (i.e. control and gaming), pre- and post-test performance, and the interactions between each factor.

3.3.4 *Monocular Presentation*

Comparisons of performance between the trained and monocular stimulus sets revealed a significant main effect of time: participants improved in performance from pre- to post-test $F(1,54) = 45.810$, $p < 0.001$, $\eta_p^2 = 0.46$, and significant main effect of group: performance in the gaming group ($M = 1.440$, $SD = 0.622$) was superior to that of the control group ($M = 0.943$, $SD = 0.432$), $F(1,18) = 7.227$, $p = 0.015$, $\eta_p^2 = 0.29$.

There was a marginally significant effect of condition, whereby participants performed better when presented with the trained dichoptic images ($M = 1.25$, $SD = 0.558$), compared with monocular presentation, where there was only one of the two dichoptic images presented during a trial ($M = 1.13$, $SD = 0.618$), $F(1,54) = 3.409$, $p = 0.070$, $\eta_p^2 = 0.06$.

There was a marginally significant interaction between experimental group and time $F(1,54) = 3.632$, $p = 0.062$, $\eta_p^2 = 0.06$: the difference in *d'* scores from pre- to post-test was larger for the gaming ($M_{diff} = 0.58$) compared to the control group ($M_{diff} = 0.33$).

There was also a marginally significant interaction between condition and time, $F(1,54) = 3.475$, $p = 0.068$, $\eta_p^2 = 0.06$: there was a difference in *d'* scores before and after training in the monocular condition ($M_{diff} = 0.581$), compared to the dichoptic training condition ($M_{diff} = 0.330$). These results indicate that performance was slightly better when the participant viewed the dichoptic images, compared to monocular viewing.

3.3.5 *Filter-switched*

Comparisons of performance between the trained stimulus set and when the filters were switched across the left and right eyes revealed a significant main effect of time $F(1,54) = 26.790$, $p < 0.001$, $\eta_p^2 = 0.33$, indicating a significant increase in *d'* scores from pre- to post-test in both conditions, regardless of experimental group. There was also a main effect of group, whereby the gaming group ($M = 1.49$, $SD = 0.562$) outperformed the control group ($M = 1.02$, $SD = 0.404$) overall $F(1,18) = 7.081$, $p = 0.016$. There was no main effect of condition, $F(1,54) = 0.003$, $p = 0.960$, $\eta_p^2 < 0.001$.

Results revealed a significant interaction between experimental group and time, $F(1,54) = 4.048$, $p = 0.0492$: the gaming group's performance increased more from pre- to post-test ($M_{diff} = 0.47$) compared to the control group ($M_{diff} = 0.207$). There were no other significant interaction effects. Thus, performance in the training and filter-switched conditions was very similar from pre- to post-test for both the gaming and control groups.

3.3.6 *1/f noise*

We compared performance on differing contrast levels of the *1/f* stimulus (stimulus-matched contrast, or stimulus doubled contrast) to performance on the trained stimulus in a single model. There was a significant main effect of time whereby performance significantly increased from pre- (M = 1.40, SD = 0.658) to post-test (M = 1.73, SD = 0.813). Further, performance in the gaming group (M = 1.83, SD = 0.761) was significantly better in comparison to that of the control group (M = 1.28, SD = 0.645), $F(1,20) = 5.571$, $p = 0.029$, $\eta_p^2 = 0.22$.

As expected, there was a significant main effect of condition, indicating different performance between the trained stimulus and both contrast levels of *1/f* noise $F(1,55) = 52.029$, $p < 0.001$, $\eta_p^2 = 0.65$.

We conducted two additional planned comparisons comparing each *1/f* noise level to the trained stimulus. There was a significant difference in performance between the stimulus-matched contrast level of the *1/f* noise stimulus and the trained stimulus, $F(1,49) = 81.379$, $p < 0.001$, $\eta_p^2 = 0.62$, with participants performing better in the stimulus-matched *1/f* noise condition (M = 2.32, SD = 0.612) compared to the training condition (M = 1.25, SD = 0.558). As expected (the noise level was set, based on pilot data in attempts to equate performance) there was no significant difference in performance between the stimulus-doubled contrast level of the *1/f* noise condition (M = 1.14, SD = 0.470) and the trained stimulus $F(1, 41) = 0.077$, $p = 0.782$, $\eta_p^2 < 0.001$.

3.3.7 *Monocular Object*

Comparisons of performance between the trained stimulus set and when the object was only present in one of the two eyes revealed a significant main effect of time $F(1,54) = 19.485$, $p < 0.001$, $\eta_p^2 = 0.27$, indicating a significant increase in *d'* scores from pre- to post-test in both

conditions, regardless of experimental group. There was also a main effect of group, whereby the gaming group (M = 1.301, SD = 0.58) significantly outperformed the control group (M = 0.835, SD = 0.400), $F(1,18) = 7.418$, $p = 0.014$, $\eta_p^2 = 0.29$. Finally, there was a main effect of condition: performance significantly decreased when the object was only present in one of the two eyes (M = 0.883, SD = 0.477), versus when the object was present in both eyes in the training condition (M = 1.25, SD = 0.558), $F(1,54) = 34.455$, $p < 0.001$, $\eta_p^2 = 0.39$. There were no significant interactions between condition and time, condition and group, or time and experimental group.

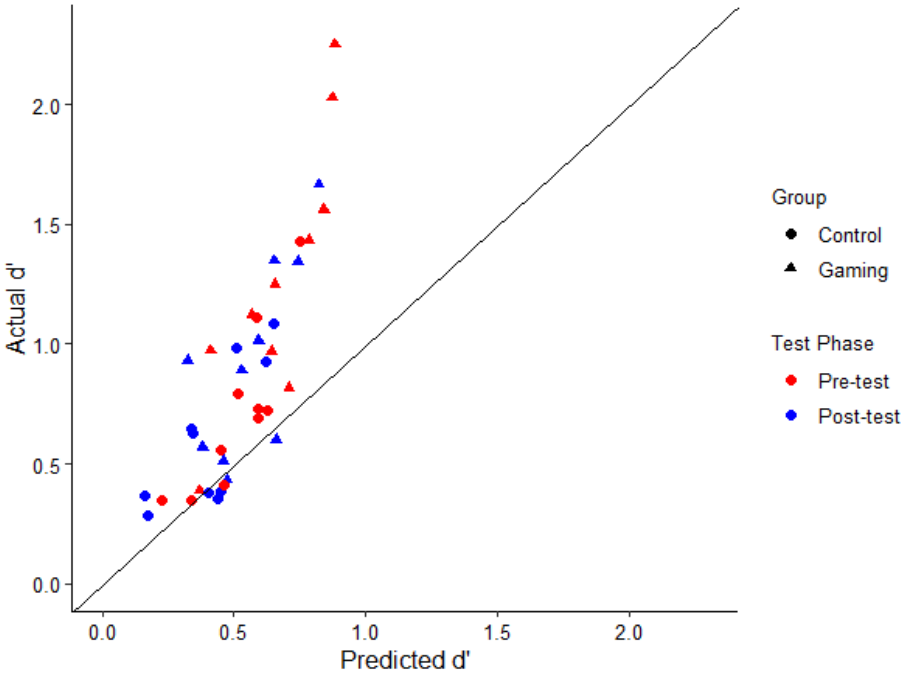


Figure 3.4. Predicted *d'* scores in the monocular object condition compared to participants actual *d'* in the monocular condition. The black line represents performance if participants actual *d'* matched that of the predicted *d'*, with an intercept equal to zero and a slope equal to one.

The primary purpose of the monocular object condition was to assess whether participants were suppressing one eye, either consistently throughout a session or on a trial by trial basis. The object was presented to one eye only (randomly interleaved across trials), with the other eye containing the filtered background scene without the object.

If participants were suppressing one eye throughout a trial then, in the monocular object they would perform at chance on those trials where they suppressed the eye that contained the object. Predicted *d'* scores were calculated by taking the midpoint between participants' performance in the dichoptic training condition, and chance performance (50%). As an example, if a participant performed with a 75% accuracy rate in the dichoptic training condition, their predicted performance in the monocular object condition, if they were suppressing one eye, would be 62.5%. We calculated the predicted *d'* score for each participant, and compared it to their actual performance in the monocular object condition using a binomial test.

Figure 3.4 shows participants predicted *d'* in the monocular object condition compared to each participant's actual *d'* in both the pre- and post-test monocular object conditions. Data points fall consistently above the line, suggesting that participants were not simply suppressing on eye throughout a trial.

In one-tailed binomial tests using a Monte Carlo simulation ($n_{samples}$ = 1000), we compared the number of observations falling above the predicted d' value for both the pre (N = 17) and the post (N = 16) test to the total number of observations (N = 20). Each test generated a significant p value (p_{pre} = 0.007; p_{post} = .023), showing that *d'* values were larger than the performance value that would be expected if they were suppressing one eye throughout individual trials.

3.3.8 *Modeling of interocular suppression and masking strength effects*

Finally, we used a linear regression model to test the effects of interocular and within eye masking. The 'interocular' predictor consisted of the monocular condition as one level, and the filter-switched, 1/f noise, and training conditions combined into a second level. The 'mask level' factor consisted of three levels: 1) whether the mask in each eye contained the trained image masking

information (i.e. [I' * F] & [I' * F']), 2) whether the *1/f* noise contrast in the *1/*f condition was matched to the trained stimulus, or 3) whether the *1/f* noise contrast was double that of the trained stimulus mask. Time (pre vs. post-test) and group (gaming vs. control) were additional factors. The model also assessed the interaction between experimental group and time.

The model accounted for 55.2% of the variance in the data ($R^2 = 0.552$), and had a *RMSE* value of 0.485, indicating the model fit the data well. The transparent lines in the top panel of **Figure 3.3** represent model predictions for each pre- and post-test condition. Most model predictions are close to data values (with the exception of the stimulus-matched pre-test *1/f* noise condition). Note, that there are no predictions for the monocular object conditions, as this condition was designed to test a different learning strategy.

The model intercept was significantly different from zero, $b = 0.893$, $SE = 0.084$, $p < 0.001$. As expected, there was a significant effect of group: with the gaming group performing significantly better compared to controls, $t(151) = 2.99$, $SE = 0.107$, $p = 0.003$. There was also a significant main effect of time, in that learning significantly increased from pre- to post-test $t(151) = 2.35$, $SE = 0.107$, $p = 0.020$. The interaction between experimental group and time approached significance, indicating that the gaming group learned slightly more over time, $t(151) = 1.96$, $SE = 0.150$, $p = 0.052$.

There was no effect of monocular vs. dichoptic stimulation on performance $t(151) = -1.339$, $SE = 0.092$, $p = 0.182$. Mask strength did have an effect: performance was better for the *1/f* noise condition at low contrast ($t(151) = 9.357$, $SE = 0.110$, $p < 0.001$) than for the other two conditions, for which performance was equivalent ($t(151) = -0.557$, $SE = 0.137$, $p = 0.578$)

To evaluate whether or not participants' status as a graduate student had a different effect on these results, we ran a second model with graduate student status rather than gaming/non-gaming as the group predictor. This model resulted in slightly more explained variance, $R^2 = 0.62$ but a slightly lower *RMSE* of 0.43, compared to the model that used training condition as the grouping variable.

The model intercept was significantly different from zero, $b = 0.820$, *SE* 0.078, $p < 0.001$ showing that there was a significant difference in performance between graduate and undergraduate students, in that graduate students performed significantly better overall, compared to undergraduates, $t(151) = -4.744$, $SE = 0.098$, $p < 0.001$. All other results (the effects of having an interocular stimulus and the effect of mask strength) were the same as the model described above. There was a marginally significant interaction between graduate status and learning over time, with graduate students learning significantly more over time compared to undergraduates, $t(151) = -1.829$, $SE = 0.138$, $p = 0.069$.

These two models suggest that participants employed a learning strategy that involved discounting the stimulus mask within each eye. While graduate students did learn more compared to undergraduates, this effect was marginally significant, and is likely associated with a slightly larger effect of learning in the gaming group, considering the gaming group was made up of 70% graduate students.

3.3.9 *Modeling of a stimulus detection using Probability Summation*

These results, described above, show better performance in the dichoptic training condition, compared to both the monocular and monocular object conditions (**Figure 3.3**). There are a few possibilities why performance might be better in the dichoptic condition.

The simplest explanation is that participants might be suppressing one eye throughout the experiment. However, there are two reasons to reject this model. First, this model predicts identical performance in monocular and dichoptic conditions, whereas, as shown in **Panel A** of **Figure 3.5**, performance is slightly better in the dichoptic compared to the monocular condition. It is also possible, as shown in **Figure 3.5, Panel E**, to calculate estimated probability correct in the monocular object condition, given dichoptic performance. This was done by calculating the underlying detection probability corresponding to each dichoptic *d'* in a 2 alternative forced-choice task. One then (based on the assumption that participants only attend to the eye containing the object in half the trials) halves this detection probability, and can then recalculate the expected *d'*. Performance in the monocular object condition is far better than what would be expected if participants were suppressing one eye.

A second possibility is that participants may be monitoring the stimulus in each eye independently during dichoptic viewing, and performance is enhanced in dichoptic training conditions due to probability summation (this model assumes that there is unlimited attentional capacity when viewing the dichoptic stimulus). Probability summation was estimated using a function which took in the probability of detecting the stimulus in the monocular object condition when the object was in the left and right eye separately, and the dichoptic conditions. We then took this input and found the underlying detection probability that minimized the mean squared error between the predicted and obtained *d'* for both monocular conditions and the dichoptic condition, as shown in **Figure 3.5**, **Panels D and F**. Predicted *d'* scores predict performance in each of these conditions well, suggesting that probability summation model is a plausible explanation for detecting the object in each trial.

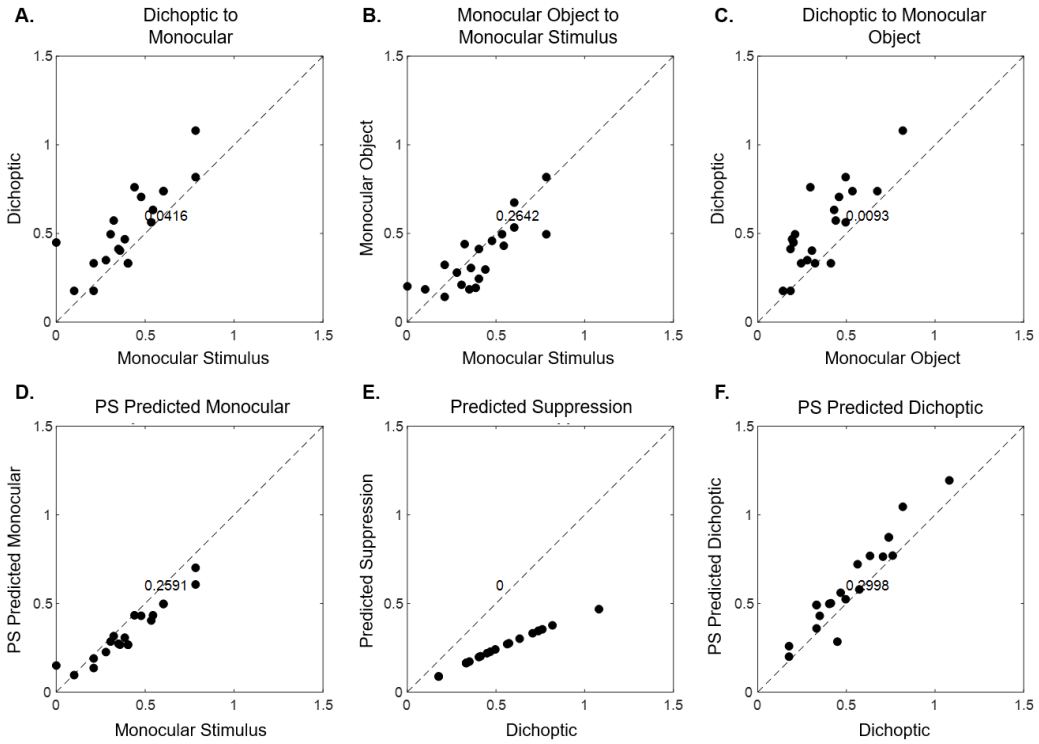


Figure 3.5. A). Comparison of participants actual *d'* in the dichoptic, compared to the monocular condition. **B)** Comparison of participants actual *d'* in the Monocular object, compared to the monocular stimulus condition (where there was only one of two dichoptic images presented during each trial). **C)** Comparison of participants actual *d'* in the dichoptic, compared to the monocular object condition. **D)** Results of the probability summation model for monocular stimulus detection. The correlation between the probability summation model and the actual data is shown on the dotted line. **E)** Using the detection probability corresponding to each individual's dichoptic performance, we calculated the underlying stimulus detection probability. We then assumed that this detection probability would be halved if observers were only processing one eye in a given trial. We then converted this lower detection probability into percent correct and corresponding *d'* value. **F)** Results of the probability summation model for dichoptic stimulus detection. The

correlation between the probability summation model and the actual data is shown on the dotted line.

3.4 DISCUSSION

While our dichoptic images did not simultaneously stimulate on- and off- pathways per se, participants viewed input that likely caused a similar disruption to early population responses, as described previously. We previously established that it is possible to learn to discriminate everyday objects using this distorted input in a standard object discrimination task (Esquenazi et al., 2021). The work here extends our previous work by asking whether playing an interactive video game with the distorted input provides an enhanced learning experience, compared to a traditional object recognition task. Specifically, we asked whether gaming might lead to generalizable learning that transfers to a novel set of objects, and investigated possible mechanisms of learning using additional pre- and post- transfer tests not employed previously.

3.4.1 *Transfer of learning to a novel set of objects*

While we expected control participants to display some improvement in the object discrimination task due to repeated stimulus exposure, we predicted that gamers would outperform controls. Results show that while both groups improved over time, any effects of video game training were very small. Both experimental groups performed roughly the same over all five sessions of the object discrimination task (see **Figure 3.2**). These results indicate that while it is possible to learn to compensate for abnormal early visual population responses, learning may be specific to the trained stimuli, and the context in which training takes place.

Specificity of learning, or a low degree of transfer to novel tasks, is typical in tasks that train on low-level stimulus features such as orientation and spatial frequency (Fiorentini & Berardi, 1980).

Training can lead to learning that is so specific that it fails to transfer to novel retinal locations (Karni & Sagi, 1991), and in some cases, to an untrained eye (for an extensive review on learning specificity in perceptual learning, see Fahle, Poggio, & Poggio, 2002). This is not to say that learning that does not transfer is not useful learning. Rather, specific learning may indicate that cells which are highly tuned for low-level stimulus features have responded to training, which implicates cortical changes at a site that is relatively early along the visual hierarchy. However, it is also possible that training can result in changes to neuronal tuning in cells that were most useful in decoding the trained task parameters, a theory that does not preclude the involvement of higher visual areas in stimulus specific learning (Mollon & Danilova, 1996).

The extent of learning transfer is also influenced by the difficulty of training conditions. The reverse hierarchy theory of perceptual learning (Ahissar & Hochstein, 1997) states that for learning to transfer, the training task should be relatively easy so that later visual areas with less specific receptive fields are responsible for any initial learning. Learning at early cortical sites will take place as task difficulty begins to increase, with changes in cortical tuning cascading down the visual hierarchy in a top-down manner. Thus, if a task starts off difficult, it will likely target changes in tuning at an early cortical site, and learning will fail to transfer. In an update to this theory, it is proposed that the *transfer* task, rather than the training task determines the degree of transfer (Jeter et al., 2009). In order for transfer to occur, the stimulus tested at transfer should be easily discriminable. According to this model, generalizable perceptual learning is associated with an attentional system in high-level cortical locations that, when activated, unlocks learning in lower areas of visual processing.

Recent studies using action video games as a training tool show a remarkably high rate of transfer to a wide range of visual tasks (Dye, Green, & Bavelier, 2009; Green & Bavelier, 2019; Green &

Bavelier, 2003, 2012; Green et al., 2010; Li et al., 2009). While mechanisms of generalizable learning as a result of video game training are still under investigation, it has been proposed that the high transfer rate to standard psychophysical tasks might be attributed to shared visual demands between the transfer task and the trained video game. For example, Oei & Patterson (2015) tested learning transfer in different action video games that had varying levels of demand on speed of decision making, multiple object tracking, attentional switch, selective attention, and visual search. The level of video game 'demand' predicted performance in the transfer tasks: the game that required the highest demand in attentional processing resulted in the greatest transfer to measures of attentional blink and attention to multiple targets, but not visual search or auditory detection, skills that were not highly required to play the game.

Initially, we hypothesized that training on our object video game would transfer quite well to a discrimination task involving the same type of stimuli. Our video game and transfer task contained stimuli that varied in size, location, and rotation, all while becoming progressively more difficult, conditions that all foster generalizable learning (Green et al., 2010). Further, participants appeared to master the game from levels 1-4 by at least session 15, indicating substantial, and potentially transferrable learning during training (see **Figure 3.1**). However, it remains possible that our training and testing tasks may have not contained enough common elements to produce substantial transfer. Our video game was engaging, and required identification of multiple objects quickly and simultaneously rotating in motion simultaneously on a consistent background. However, the object discrimination transfer task did not contain stimuli in motion (although the entire image did pan across the screen on each trial), required discrimination of one object on multiple different backgrounds, and like many other psychophysics tasks, was monotonous and not mentally stimulating. Perhaps a more engaging transfer task that required discrimination of objects in

motion might have resulted in a higher degree of transfer as a result of video game play. A second possibility, predicted by the reverse hierarchy of learning, is that our task may have been too perceptually difficult to result in transfer.

A third possible reason why we did not witness learning transfer from our video game to the object discrimination task might have been because our game did not contain enough elements of an action video game. While no two video games are the same, some categories of games appear to enhance perceptual skills more than others. It appears that the largest contributing factor of learning transfer following video game play is the speed and accuracy with which decisions need to be made in order to succeed in the game (Achtman et al., 2008). First- and third-person shooter games contain high stakes, fast-paced, and unpredictable events which require precisely intact visuo-motor, and cognitive systems to succeed. Video games in categories such as sports and racing (e.g. FIFA, NBA2k), strategy (e.g. Roller Coaster Tycoon, Sims), and puzzle or card games can be visually complex, require quick reallocation of attention, and demand the tracking of multiple objects, but do not typically lead to increased visual skills (see Achtman et al., (2008) for a review). This is likely because the visual demands by these games are not nearly as intense as those required by fast-paced action video games.

Our video game appears to align most closely with Tetris, a puzzle video game that requires fast visuo-motor control, the need to track multiple objects and spatially attend to the entire screen, just as in our video game. However, games like Tetris are devoid of the need to "quickly aim at targets", as in first-person shooter games. Indeed, training with Tetris has indeed been seen to transfer to some other visual tasks (Belchior et al., 2013; Okagaki & Frensch, 1994), though not nearly to the degree that action video games do (Achtman et al., 2008; C. S. Green & Bavelier, 2019; S. Green & Bavelier, 2012; R. Li et al., 2009). Pilot testing with our filters revealed that it was nearly

impossible to play a filtered version of any commercially available action video game (Mortal Combat II) in a way that was engaging or comprehensible to participants. Thus, the need for a less visually demanding game was vital in our experiment.

3.4.2 *Transfer of learning tests reveal a within-eye strategy to discount masking information*

We previously examined transfer of learning across three conditions: monocular presentation of the stimulus, switching the filters across the two eyes, and replacing the contrast-reversed image information with 1/f noise (Esquenazi et al., 2021). Learning completely transferred to the monocular and filter-switched conditions, while less learning transferred to the *1/f* noise condition. It was concluded that the source of learning was attributed to participants learning to discount within-eye masking information during training. However, some open questions still remained, such as whether or not the lack of transfer for the *1/f* noise mask was due to it being ineffective as a mask, given that there was high performance in that condition before and after training. Finally, it was unclear whether participants were suppressing one eye immediately upon viewing our stimuli – our monocular condition only tested whether participants *learned* to suppress one eye during training. Here, we employed the same transfer of learning tests used previously, and added two additional transfer conditions. Finally, it is worth a reminder to the reader that the transfer of learning from pre- to post-test in **Chapter 2** was found with a consistent object set across all phases of the study, while in this study, there was a different set of objects in every study phase. Therefore the results here might have been different if the object set remained consistent across the pre-test, training, and post-test phases of the study.

The first modification to the pre- and post-tests was the addition of a condition where the contrast of the *1/f* noise was double that of the stimulus. Doubling the contrast of the *1/f* noise made this

condition a more effective mask. Transfer of learning in this condition was complete, suggesting that the lack of transfer in the previous experiment was indeed due to the mask being less effective, rather than a failure of learning to transfer to *1/f* noise mask.

To understand whether participants suppressed one eye upon immediately viewing the stimulus, we designed a condition where the target object was only present in the scene in one of the two eyes. We reasoned that if participants were suppressing one eye to perform the task, they would only see the object in their non-suppressed eye, and thus performance in this condition would be worse than that in the training condition. However, results show that this is not the case. Performance in the monocular object condition was greater than what would be expected if participants consistently suppressed one eye to view the stimulus. The results of this condition provide further evidence that within a trial participants are processing and interpreting information from both eyes. One possibility is that participants do suppress, but that the trials were long enough (2s) that the suppressed eye switched within each trial. An alternative possibility is that participants used probability summation to decode the stimuli (**Figure 3.5**, section 3.3.9).

Probability summation predicts that the images projected to each eye are monitored by separated mechanisms, and that the better performance in the dichoptic condition is attributed to the participant having a greater chance of detecting the target object because they are monitoring the two eyes separately. However, close inspection of the data suggests that the probability summation model systematically under-predicts performance in the monocular condition and over-predicts performance in the dichoptic condition (**Figure 3.5**, **Panels D and F**). This systematic deviation in the model suggests that a model that included a small amount of additive summation across the two (Kingdom, Baldwin, & Schmidtmann, 2015) eyes would also provide a good fit for these data. A model using purely probability summation in the post-test results in further under- and over-

predictions, providing additional support for a combination of additive and summative probability estimation in each trial decision.

Results of all other pre- and post-tests were mostly replicated here. Learning transferred when the filters were switched across the two eyes, and participants performed better when the [I'*F] and [I'*F'] masking was replaced with *1/f* noise when the contrast of the noise was matched to the stimulus contrast. One difference between the results here and those previously was the small main effect of condition, showing slightly better performance in the training condition (M = 1.25, SD = 0.558) compared to the monocular condition (M = 1.13, SD = 0.618). In addition, there was a moderately significant interaction between condition and time that indicated more learning over time in the training condition compared to the monocular condition. However, the effect sizes for both of these results are small ($\eta_p^2 = 0.06$ for both results), so these results should be interpreted with caution. The similarity in *d'* likely points to a statistical anomaly in the models, rather than a true, meaningful difference in learning between the two conditions.

Our mixed effects models tell a story about performance in each condition before and after training, however, we provide a simple linear model here that explains learning strategies, and whether these strategies differed by experimental group. There was no significant difference in performance when all conditions that contained an interocular mask were combined and compared to the monocoular condition where there was only one mask present at any given time. This provides additional evidence that counters a suppression theory. Further, the contrast of the mask only affected performance when the contrast strength was low. These results indicate that training aided in participants ability to interpret the [I * F'] stimulus, while learning to discount the [I' * F] masking. Performance decrements, but not the complete inability to interpret the stimulus when the *1/f* noise contrast was high were similarly due to improvements in interpreting the [I * F']

stimulus. This model was fit to data from our previous experiment, and results were similar (Esquenazi et al., 2021; data not shown). One obvious next step would be to test participants on a stimulus where [I * F'] was in one eye, and 1/f noise was in the other. Our model here predicts that performance would be similar to that of the training condition.

The gaming group consistently outperformed the control group in all of the transfer of learning tests, even prior to training. This is possibly due in part to a selection bias (see limitation section below), whereby gamers were initially more invested in the study. Regardless of gamers' superior initial performance, our model indicates that gamers *learned* more over time, albeit the interaction between experimental group and time only approached statistical significance (m = 0.294, p = 0.052). This suggests that gaming with the filtered stimulus may have provided a small advantage in learning. However, the effect size is small, and it is possible that this effect could be better explained by gamers' increased overall motivation in the study.

3.4.3 *Limitations*

One major limitation of these results is the noticeable difference in performance between the gaming and control groups in all phases of the study, including the pre-test. One possibility is that performance might have been impacted by gamers overall engagement in the study. From the start, the gaming group committed a significant amount of time to the experiment (36 hours total), and had knowledge of which experimental group they were in. This commitment, in addition to a more enriching training experience could have resulted in significantly more interest in the study as a whole, and this increased commitment might have created a more attentive and alert participant group.

Further examination revealed that a more plausible reason for a difference in pre-training performance between gamers and controls might have been the students' graduate status. Graduate students tended to outperform undergraduate students, regardless of experimental group, raising an important question about performance differences in psychophysics experiments between two populations who present, at least superficially, as similar. The graduates in our study were mainly comprised of students in the psychology department, the thus were well educated in the field of psychology and at the very least, knowledgeable about the importance of staying attentive during an experiment, despite a dull and repetitive task. It is then possible that the graduate students in our study, could be considered "sophisticated" subjects (Eyman, 1966), and thus may have possessed higher motivation to perform well, a factor known to affect psychophysical judgements (Galanter, 1962).

3.4.4 *Conclusions*

Our method of presenting conflicting spatial frequency and orientation information to each eye represents a novel way to study the distortions caused by electronic sight restoration technologies, by creating 'virtual patients'.

While gaming with our filters did not appear to provide a substantial improvement in the ability to discriminate objects, participants significantly improved in both the discrimination task, and the video game, suggesting that learning to adapt to unnatural on- and off-cell population responses is likely possible, albeit context specific. Further, this indicates that retinal and cortical prosthesis patients' learning process will likely be orders of magnitude longer than that of patients learning to interpret the input provided by a cochlear implant (Fallon et al., 2008), where significant learning can take place in hours or days.

Finally, the learning strategy employed by participants here reveals evidence of a cognitive decision mechanism that involves processing information from both eyes when decoding a stimulus with very low interpretability, but choosing the "best fit" according to which eye provides the most discriminable visual information. This is in line with theories that suggest perceptual learning stems from learning to reweight the appropriate visual channels for the task at hand, and prune irrelevant ones (Dosher & Lu, 1998). The work here extends this idea and suggests that although participants perceived one coherent image during viewing, training may have aided in the ability to decouple the images in both eyes and place a decision weight on each eye on a trial by trial basis. This conclusion can be, in part, be explained by a probability summation model which predicts that participants superior performance when the object was present in both eyes is attributable to an increased chance of detecting the object in one of the two eyes.

Chapter 4. SUMMARY AND CONCLUSIONS

This thesis asked several questions about visual cortical plasticity and perceptual learning in the context of patients with sight restoration technologies. Is it possible to modify the fundamental building blocks of perception in adulthood? How much, if any, learning occurs when sighted participants are presented with visual input that presents a similar decoding challenge to that of simultaneous on- and off-cell stimulation by electronic sight restoration technologies? Can training with this distorted input in a highly dynamic and interactive environment boost the initial rate of learning, and does learning transfer to novel stimuli in a different context? Finally, what can training with distorted visual input that creates unnatural population responses in early visual cortex tell us about mechanisms of cortical plasticity and perceptual learning? The following results were found:

1. Modeling of cortical responses suggests that dichoptic presentation of conflicting spatial frequency and orientation information to sighted individuals resulted in a distortion of population responses in early visual processing that represents a reasonable proxy to the cortical distortion of simultaneous stimulation of on- and off-cells in electronic sight restoration technologies (see **Figure 2.1**, and **Section Chapter 2**).

2. Our ‘virtual patients’ exhibited remarkable improvements in an object discrimination task with the filtered input over multiple training sessions (**Figure 2.5** and **Section 2.4**).

3. Performance improved over time when participants were tasked with playing an interactive video game with the distorted input (**Error! Reference source not found.**, and **Section** Error! Reference source not found.), however, there was little evidence of generalization

of learning to a novel set of stimuli in a transfer of learning task (**Error! Reference source not found.**).

4. Participants likely did not learn to improve perception of the distorted image information through monocular suppression, but rather, processed input from both eyes and 1) discounted irrelevant masking information from the contrast-reversed image and 2) Chose the eye with the most relevant information about the stimulus to perform the task at hand, a result that could be explained using a simple probability summation model (see **Figure 2.6**, **Error! Reference source not found., Figure 3.5** and **Sections 2.4**, **3.4**).

4.1 Insights into Perceptual Learning

The work in this thesis presents an opportunity to examine the cortical mechanisms that explain perceptual learning. As reviewed in **Section 1.9**, there is substantial disagreement about *what* changes occur in the brain as a result of learning, *where* those changes take place, and exactly *how* the brain facilitates improved performance on a wide array of visual tasks. Opinions are divided based on two ideas: 1) perceptual learning results from changes in neuronal representation, a an experience-based cortical retuning theory, and 2) learning is a result of task-specific weights being placed on neuronal channels activated during the task, a selective channel reweighting theory. Other theories have been proposed, specifically those which involve contributions by attentional networks which may mediate the learning process (Szpiro & Carrasco, 2015; Tsushima & Watanabe, 2009; however, attention to trained feature does not indicate that unnatended features cannot be learned, Seitz & Watanabe, 2005).

Here, we did not find any evidence of monocular specificity: participants in both studies performed equally well when the trained stimuli were switched across the two eyes. These results provide

evidence against an early retuning theory model to explain perceptual learning. Learning that transfers across the two eyes has been found when low precision stimuli (e.g. stimuli that vary in retinal location, size, and orientation) are used during training (Ahissar & Hochstein, 1997) or the task tested at transfer (Jeter et al., 2009), suggesting that learning which transfers across the two eyes may not be the result of purely early cortical changes. The specific task parameters in our experiments likely explain the interocular transfer, but the learning here, which generalized across the two eyes likely does not preclude the involvement of V1 specific tuning changes.

Recently, the field of perceptual learning made a substantial advance towards understanding the cortical mechanisms involved in learning. Jia et al. (2020) used ultra-high-field fMRI to examine the different layers in V1, which are associated with different forms of visual processing. The superficial layers of V1 contain the dendrites of neurons that project to visual decision making areas such as the intraparietal sulcus (IPS), while the middle and deeper layers are responsible for interpreting input connections from the retina and LGN. Jia et al (2020) found that training in a classic orientation discrimination task altered connections in the superficial output layers of V1 and the middle, input layers of the IPS. It was also concluded that feedforward connectivity between these two sub-lamina was responsible for a bulk of the changes. While the possibility of changes involving horizontal connections in V1 cannot be ruled out (Seitz, 2020), this study, along with others (Ahissar & Hochstein, 1997; Dosher & Lu, 1998), supports a perceptual learning model that involves multiple brain areas associated with the visual demands of the task, and decision areas that refine perceptual readouts to improve on the task over time.

While we did not employ fMRI techniques to elucidate the cortical locus of learning in our study, it is reasonable to assume that our data can be explained by a model that involves multiple areas of the brain. Our stimuli created a disruption in V1 population responses, but the demands of object

discrimination task may have led to alterations of object specific representations in cortical areas responsible for processing objects (e.g. area LOC). One future direction of this research would be to use fMRI and multivoxel pattern classification analysis before and after training to examine how the distorted input described here affects neural representations of the trained objects.

Finally, the pre-and post-testing results used to examine learning strategies also point to the involvement of areas outside of early visual cortex that are involved in decision making. Participants in both experiments were able to process the conflicting information across the two eyes, but subsequently make a correct decision in the task based on monitoring both eyes. Thus, success in decoding abnormal population responses in V1 might involve higher order cognition to solve the task which contains the stimuli that result in such distortions.

4.2 Insights into True Patients

A central question of this thesis was whether it is possible for patients to adapt to prosthetic vision, which severely distorts physiologically normal vision. Answers to this question will be crucial for the development of next generation devices. Given the limited number of patients, long term post implantation monitoring of experience-dependent learning is still scarce. Repeated use of an epiretinal device over a period of 10 weeks resulted in a patient's ability locate the direction of motion of objects with increasing accuracy (Humayun et al., 2003). When patients were followed for one year after implantation of a suprachoroidal device, small but significant improvements in a line drawing test were obtained beginning roughly 9 months post-surgery (Fujikado et al., 2016a). The longest patient follow up was 10 years after implantation of the first generation epiretinal Argus device, but improvements in visual function measured by square localization, moving bar tasks, and orientation and mobility tasks were mixed (Ho et al., 2015). Importantly,

none of these findings are the result of any concerted effort to perceptually *train* patients to improve visual skills.

The results provided here present evidence that patients may be able to overcome distortions to early visual population responses in V1, but that learning to use a prosthetic device may require rigorous effort in the form of repetitive and consistent visual training. With the limited number of patients ($N \cong 500$ globally; Ayton et al., 2020), the work here might provide a framework to understand the extent of learning possible in other distortions created by sight restoration technologies such as "streaky percepts" caused by axonal stimulation (Fine & Boynton, 2015) .

There are very few, if any, established visual training paradigms for patients that harness the beneficial effects of perceptual learning. Thus, the work here can be considered a stepping stone to understanding plasticity of patients in two ways:

1. In the absence of access to real patients, recreating other spatial and temporal distortions (see **Section 1.7**) and measuring learning using methods described in this thesis can aid in the construction of a model of plasticity and learning in prosthetic vision. By isolating each visual distortion (e.g. the effects of axonal stimulation) and testing perceptual learning of that distortion, it may be possible to understand plasticity in response to isolated distortions, and estimate plasticity potential across all known distortions.
2. Design of an effective rehabilitation program that incorporates perceptual learning protocols to improve prosthetic vision. Playing a video game with the distorted input resulted in very little additional learning compared to a standard, repetitive psychophysical task. This supports the likelihood that learning to discriminate stimuli in prosthetic vision will be highly specific and context dependent, improve very gradually, and may not require

a dynamic training environment. Thus, while prosthetic vision will never recreate normal vision, it is possible that continued, effortful use (in a recent study, patients reported not turning on their device on a daily basis (Erickson-davis & Korzybska, 2020) combined with structured learning might alter patients' representations of visual stimuli to include prosthetic representations, and aid in a completely new form of vision (Roth & Zohary, 2015; Striem-amit, Cohen, Dehaene, & Amedi, 2012).

4.2.1 *Effects of Ageing*

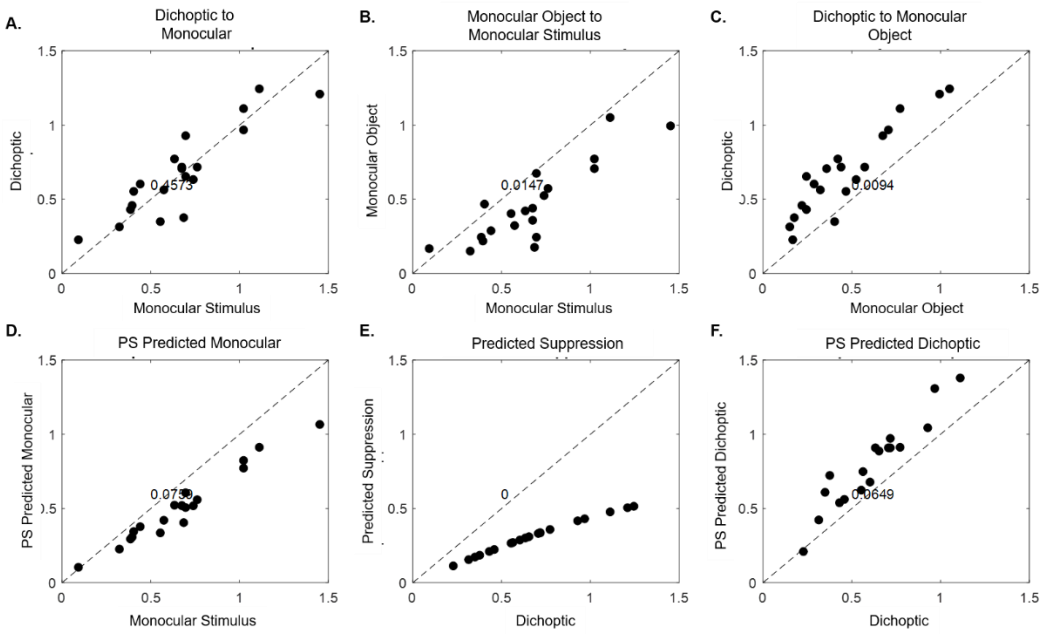
It is well known that the plasticity potential of the brain declines with age. As the brain ages, cellular metabolic activity slows down, cells begin to lose myelin, genes undergo mutations, and dendritic branches decrease, all factors that influence plasticity (Burke & Barnes, 2006). Behaviorally, these changes can result in decreased visual abilities. For example, increased reaction times in response to moving objects as one ages (Mendelson & Wells, 2002). One glaring limitation of the research here is that learning and plasticity was examined in healthy, young adults. Prosthesis patients tend to be much older: some patients are well into their 70s. This presents a challenge to extrapolating the potential of plasticity in actual patients from the results obtained here.

Recent evidence showing increased cortical plasticity in the somatosensory (Pellicciari, Miniussi, Rossini, & Gennaro, 2009) motor cortices (Blais, Martin, Albaret, & Tallet, 2014) call the hypothesis of plasticity in the elderly into question. While it is true that plasticity potential may decrease, this is not to say that it is not completely absent. Repetitive electrical stimulation delivered through transcranial magnetic stimulation (TMS) over the somatosensory cortex resulted in increased excitability of neurons, placing the elderly brain into a plastic state (Pellicciari et al.,

2009). Training a neural network that modeled visual cortical activity based on behavioral data from humans revealed that the ageing brain still retains the capacity to increase connectivity in visual areas through training, albeit learning was slower and required more training iterations in the model (Koprinkova-hristova et al., 2020).

It can be expected that learning to adapt to prosthetic vision will be slow, supported by the rate of learning witnessed here, and by evidence of decreased plasticity in the elderly. However, it is clear that the ageing brain retains a remarkable degree of plasticity, and simulating distortions of electronic sight restoration technologies and other aspects of prosthetic vision (Kasowski & Beyeler, 2022) can provide clues into the patient experience. However, the potential for cortical plasticity may not necessarily indicate that this plasticity can be behaviorally exploited.

APPENDIX A



Probability Summation results for Post-Test. A). Comparison of participants actual *d'* in the dichoptic, compared to the monocular condition. **B)** Comparison of participants actual *d'* in the Monocular object, compared to the monocular stimulus condition (where there was only one of two dichoptic images presented during each trial). **C)** Comparison of participants actual *d'* in the dichoptic, compared to the monocular object condition. **D)** Results of the probability summation model for monocular stimulus detection. **E)** Using the detection probability corresponding to each individual's dichoptic performance, we calculated the underlying stimulus detection probability. We then assumed that this detection probability would be halved if observers were only processing one eye in a given trial. We then converted this lower detection probability into percent correct and corresponding *d'* value. **F)** Results of the probability summation model for dichoptic stimulus detection.

BIBLIOGRAPHY

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.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*. https://doi.org/10.1038/387401a0

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). https://doi.org/10.1167/tvst.2.4.1

Astle, A. T., Webb, B. S., & Mcgraw, P. V. (2011). Can perceptual learning be used to treat amblyopia beyond the critical period of visual development? *Ophthalmic and Physiological Optics*, *31*(6), 564–573. https://doi.org/10.1111/j.1475-1313.2011.00873.x

Ayton, L. N., Barnes, N., Dagnelie, G., Fujikado, T., Goetz, G., Hornig, R., … Petoe, M. A. (2020). An update on retinal prostheses. *Clin Neurophysiol*, *131*(6), 1383–1398. https://doi.org/10.1016/j.clinph.2019.11.029.An

Ayton, L. N., Blamey, P. J., Guymer, R. H., Luu, C. D., Nayagam, D. A. X., Sinclair, N. C., … Allen, P. J. (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

Baker, C. I., Peli, E., Knouf, N., & Kanwisher, N. G. (2005). Reorganization of visual processing in macular degeneration. *Journal of Neuroscience*, *25*(3), 614–618. https://doi.org/10.1523/JNEUROSCI.3476-04.2005

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953–965.

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. (2011a). 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

Baseler, H. A., Gouws, A., Haak, K. V., Racey, C., Crossland, M. D., Tufail, A., … Morland, A. B. (2011b). 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

Bavelier, D., & Green, C. S. (2019). Enhancing Attentional Control: Lessons from Action Video Games. *Neuron*, *104*(1), 147–163. https://doi.org/10.1016/j.neuron.2019.09.031

Bavelier, D., Levi, D. M., Li, R. W., Dan, Y., & Hensch, T. K. (2010). Removing brakes on adult brain plasticity: From molecular to behavioral interventions. *Journal of Neuroscience*, *30*(45), 14964–14971. https://doi.org/10.1523/JNEUROSCI.4812-10.2010

Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual Learning in Parafoveal Vision. *Vision Research*, *35*(12), 1679–1690.

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

Belchior, P., Marsiske, M., Sisco, S. M., Yam, A., Bavelier, D., Ball, K., & Mann, W. C. (2013). Video game training to improve selective visual attention in older adults. *Computers in Human Behavior*, *29*(4), 1318–1324. https://doi.org/10.1016/j.chb.2013.01.034

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

Berry, M. H., Holt, A., Salari, A., Veit, J., Visel, M., Levitz, J., … Isacoff, E. Y. (2019). Restoration of high-sensitivity and adapting vision with a cone opsin. *Nature Communications*, *10*(1), 1–12. https://doi.org/10.1038/s41467-019-09124-x

Bessant, D. A. R., Ali, R. R., & Bhattacharya, S. S. (2001). Molecular genetics and prospects for therapy of the inherited retinal dystrophies. *Current Opinion in Genetics and Development*,

11(3), 307–316. https://doi.org/10.1016/S0959-437X(00)00195-7

Beyeler, M. (2019). Biophysical model of axonal stimulation in epiretinal visual prostheses. *International IEEE/EMBS Conference on Neural Engineering, NER*, *2019-March*, 348–351. https://doi.org/10.1109/NER.2019.8716969

Beyeler, M., Boynton, G. M., Fine, I., & Rokem, A. (2017). pulse2percept : A Python-based simulation framework for bionic vision, (Scipy), 81–88.

Beyeler, M., Boynton, G. M., Fine, I., & Rokem, A. (2019). Model-Based Recommendations for Optimal Surgical Placement of Epiretinal Implants. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 394–402). https://doi.org/10.1007/978-3-030-32254-0_44

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

Blais, M., Martin, E., Albaret, J., & Tallet, J. (2014). Preservation of perceptual integration improves temporal stability of bimanual coordination in the elderly : An evidence of age-

related brain plasticity. *Behavioural Brain Research*, *275*, 34–42. https://doi.org/10.1016/j.bbr.2014.08.043

Bloch, E., Luo, Y., & da Cruz, L. (2019). Advances in retinal prosthesis systems. *Therapeutic Advances in Ophthalmology*, *11*, 251584141881750. https://doi.org/10.1177/2515841418817501

Bloomfield, S. A. (1994). Orientation-sensitive amacrine and ganglion cells in the rabbit retina. *Journal of Neurophysiology*, *71*(5), 1672–1691. https://doi.org/10.1152/jn.1994.71.5.1672

Boccard, S. G. J., Pereira, E. A. C., & Aziz, T. Z. (2015). Deep brain stimulation for chronic pain. *Journal of Clinical Neuroscience*, *22*(10), 1537–1543. https://doi.org/10.1016/j.jocn.2015.04.005

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

Bower, J. D., Watanabe, T., & Andersen, G. J. (2013). Perceptual Learning and Aging: Improved Performance for Low-Contrast Motion Discrimination. *Frontiers in Psychology*, *4*(February), 1–7. https://doi.org/10.3389/fpsyg.2013.00066

Boynton, G. M., & Hegdé, J. (2004). Visual cortex: The continuing puzzle of area V2. *Current Biology*, *14*(13), 523–524. https://doi.org/10.1016/j.cub.2004.06.044

Brelén, M. E., Duret, F., Gérard, B., Delbeke, J., & Veraart, C. (2005). Creating a meaningful visual perception in blind volunteers by optic nerve stimulation. *Journal of Neural Engineering*, *2*(1). https://doi.org/10.1088/1741-2560/2/1/004

Brindley, G. S., & Lewin, W. S. (1968). The sensations produced by electrical stimulation of the visual cortex. *The Journal of Physiology*, *196*, 479–493.

Broadgate, S., Yu, J., Downes, S., & Halford, S. (2017). Unravelling the genetics of inherited retinal dystrophies: past present and future. *Progress in Retinal and Eye Research*, (59), 53--96.

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*, *7*(1), 30–40. https://doi.org/10.1038/nrn1809

Burns, A., Adeli, H., & Buford, J. A. (2014). Brain-computer interface after nervous system injury. *Neuroscientist*, *20*(6), 639–651. https://doi.org/10.1177/1073858414549015

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

Calford, M. B., Wang, C., Taglianetti, V., Waleszczyk, W. J., Burke, W., & Dreher, B. (2000). Plasticity in adult cat visual cortex (area 17) following circumscribed monocular lesions of all retinal layers. *Journal of Physiology*, *524*(2), 587–602. https://doi.org/10.1111/j.1469-7793.2000.t01-1-00587.x

Calford, Michael B., Schmid, L. M., & Rosa, M. G. P. (1999). Monocular focal retinal lesions induce short-term topographic plasticity in adult cat visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, *266*(1418), 499–507. https://doi.org/10.1098/rspb.1999.0665

Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica*, *119*(2), 217–230. https://doi.org/10.1016/j.actpsy.2005.02.004

Cha, K., Horch, K., & Normann, R. A. (1992). Simulation of a phosphene-based visual field: Visual acuity in a pixelized vision system. *Annals of Biomedical Engineering*, *20*, 439–449. https://doi.org/10.1007/BF02368135

Chen, K., Yang, Z., Hoang, L., Weiland, J., Humayun, M., & Liu, W. (2010). An integrated 256-channel epiretinal prosthesis. *IEEE Journal of Solid-State Circuits*, *45*(9), 1946–1956. https://doi.org/10.1109/JSSC.2010.2055371

Chen, P. L., Chen, J. T., Fu, J. J., Chien, K. H., & Lu, D. W. (2008). A pilot study of anisometropic amblyopia improved in adults and children by perceptual learning: An alternative treatment to patching. *Ophthalmic and Physiological Optics*, *28*(5), 422–428. https://doi.org/10.1111/j.1475-1313.2008.00588.x

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

Chung, S. T. L., Li, R. W., & Levi, D. M. (2006). Identification of contrast-defined letters benefits from perceptual learning in adults with amblyopia. *Vision Research*, *46*(22), 3853–3861. https://doi.org/10.1016/j.visres.2006.06.014

Cibis, G. W., & Fitzgerald, K. M. (2001). The negative ERG is not synonymous with nightblindness. *Transactions of the American Ophthalmological Society*, *99*, 171–176.

Crist, R. E., Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1997). Perceptual learning of spatial localization: Specificity for orientation, position, and context. *Journal of Neurophysiology*, *78*(6), 2889–2894. https://doi.org/10.1152/jn.1997.78.6.2889

Cuevas, E., Parmar, P., & Sowden, J. C. (2019). Restoring Vision Using Stem Cells and Transplantation. *Retinal Degenerative Diseases*, 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

Dacey, D. (2004). 20 origins of perception: Retinal ganglion cell diversity and the creation of parallel visual pathways.

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. (1995). Topographic reorganization in the striate cortex of the adult cat and monkey is cortically mediated. *Journal of Neuroscience*, *15*(3 I), 1631–1647. https://doi.org/10.1523/jneurosci.15-03-01631.1995

Darian-Smith, Corinna, & 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

de Jong, P. (2006). Age-Related Macular Degeneration. *New England Journal of Medicine*, 1474–1485.

Delbeke, J., Oozeer, M., & Veraart, C. (2003). Position, size and luminosity of phosphenes generated by direct optic nerve stimulation. *Vision Research*, *43*(9), 1091–1102. https://doi.org/10.1016/S0042-6989(03)00013-0

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

Dobelle, W. H., & Mladejovsky, M. G. (1974). Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. *Journal of Physiology*, *243*, 553–576.

Donald O. Hebb. (1949). *Organization of Bahavior. A neuropsychological theory*. New York: Wiley.

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

Doty, R. (1969). Electrical Stimulation of the Brain in Behavioral Context. *Annual Review of Psychology*, *20*(1), 289–320. https://doi.org/10.1016/S0140-6736(51)92044-2

Dowling, J. (1987). *The Retina: An approachable part of the brain*. Harvard University Press.

Dye, M. W. G., Green, S., & Bavelier, D. (2009). Increasing speed of processing with action video games. *Current Directions in Psychological Science*, *18*(6), 321–326. https://doi.org/10.1111/j.1467-8721.2009.01660.x

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. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2020.02.03.932905

Eshraghian, J. K., Cho, K., Baek, S., Kim, J. H., & Eshraghian, K. (2017). Biological modeling of vertebrate retina: Rod cell to bipolar cell. *40th International Conference on Telecommunications and Signal Processing*, *2017-Janua*, 391–394. https://doi.org/10.1109/TSP.2017.8076012

Esler, T. B., Kerr, R. R., Tahayori, B., Grayden, D. B., Meffin, H., & Burkitt, A. N. (2018). Minimizing activation of overlying axons with epiretinal stimulation: The role of fiber orientation and electrode configuration. *PLoS ONE*, *13*(3), 1–27. https://doi.org/10.1371/journal.pone.0193598

Espinosa, J. S., & Stryker, M. P. (2012). Development and Plasticity of the Primary Visual

Cortex. *Neuron*, *75*(2), 230–249. https://doi.org/10.1016/j.neuron.2012.06.009

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), 1–20. https://doi.org/10.1167/JOV.21.13.10

Euler, T., Haverkamp, S., Schubert, T., & Baden, T. (2014). Retinal bipolar cells: Elementary building blocks of vision. *Nature Reviews Neuroscience*, *15*(8), 507–519. https://doi.org/10.1038/nrn3783

Eyman, R. K. (1966). *The effect of subject sophistication on ratio and discrimination scales*. University of Southern California.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology*, *15*(2), 154–160. https://doi.org/10.1016/j.conb.2005.03.010

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

Fahle, M., Poggio, T., & Poggio, T. A. (2002). *Perceptual Learning*. MIT Press.

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

Fernandez, E. (2018). Development of visual Neuroprostheses: trends and challenges.

Bioelectronic Medicine, *4*(12), 1–8.

Fernandez, E., Soto, C., Alfaro, A., Gonzalez, P., Lozano, A., Pena, S., … Normann, R. A. (2019). Development of a Cortical Visual Neuroprosthesis for the Blind: Preliminary Results. *Investigative Ophthalmology & Visual Science*, *60*(4021).

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

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, Ione, & 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).

Fiorentini, A., & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, *287*.

Francis, P. J. (2006). Genetics of inherited retinal disease. *Journal of the Royal Society of Medicine*, *99*(4), 189–191. https://doi.org/10.1258/jrsm.99.4.189

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1204. https://doi.org/10.1038/nn.2889

Fritsche, L. G., Igl, W., Bailey, J. N. C., Grassmann, F., Sengupta, S., Bragg-Gresham, J. L., … Heid, I. M. (2016). A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nature Genetics*, *48*(2), 134–143. https://doi.org/10.1038/ng.3448

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

Fu, Y., Kaneko, M., Tang, Y., Alvarez-Buylla, A., & Stryker, M. P. (2015). A cortical disinhibitory circuit for enhancing adult plasticity. *ELife*, *2015*(4), 1–12. https://doi.org/10.7554/eLife.05558

Fujikado, T., Kamei, M., Sakaguchi, H., Kanda, H., Endo, T., Hirota, M., … Nishida, K. (2016a). One-Year Outcome of 49-Channel Suprachoroidal-Transretinal Stimulation Prosthesis in Patients With Advanced Retinitis Pigmentosa. *Investigative Ophthalmology & Visual Science*, *57*(14), 6147–6157. https://doi.org/10.1167/iovs.16-20367

Fujikado, T., Kamei, M., Sakaguchi, H., Kanda, H., Endo, T., Hirota, M., … Nishida, K. (2016b). 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

Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2004). Learning Strengthens the Response of Primary Visual Cortex to Simple Patterns. *Current Biology*, *14*(7), 573–578. https://doi.org/10.1016/j

Galanter, E. (1962). Contemporary psychophysics. In *New Directions in Psychology*. New York:

Holt.

Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and cartesian gratings in macaque visual cortex. *Science*, *259*(5091), 100–103. https://doi.org/10.1126/science.8418487

Garita-Hernandez, M., Lampič, M., Chaffiol, A., Guibbal, 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

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

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, *31*(5), 681–697. https://doi.org/10.1016/S0896-6273(01)00424-X

Green, C. S., & Bavelier, D. (2010). Training Induced Learning. *Brain*, *23*(4), 692–701. https://doi.org/10.1037/a0014345.Exercising

Green, C. S., & Bavelier, D. (2019). Action-Video-Game Experience Alters the Spatial Resolution of Vision. *Psychological Science*, *18*(1), 88–94. https://doi.org/10.1177/0956797619889044

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Green, S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534–537. https://doi.org/10.1038/nature01647

Green, 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, S., Li, R., & Bavelier, D. (2010). Perceptual learning during action video game playing. *Topics in Cognitive Science*, *2*(2), 202–216. https://doi.org/10.1111/j.1756-8765.2009.01054.x

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10–11), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27*, 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220

Guo, T., Yang, C. Y., Tsai, D., Muralidharan, M., Suaning, G. J., Morley, J. W., … Lovell, N. H. (2018). Closed-loop efficient searching of optimal electrical stimulation parameters for preferential excitation of retinal ganglion cells. *Frontiers in Neuroscience*, *12*(MAR), 1–12. https://doi.org/10.3389/fnins.2018.00168

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

Haim, M. (2002). The epidemiology ofretinitis pigmentosa in Denmark. *Acta Ophthalmologica Scandinavica*, *233*, 1–34.

Hamel, C. (2006). Retinitis pigmentosa. *Orphanet Journal of Rare Diseases*, *1*(1), 1–12. https://doi.org/10.1186/1750-1172-1-40

Hartong, D., Berson, E., & Dryja, T. (2006). Retinitis pigmentosa. *The Lancet*, (368), 1795–1809. https://doi.org/10.1016/0039-6257(88)90085-9

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

Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience*, *27*, 549–579. https://doi.org/10.1146/annurev.neuro.27.070203.144327

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

Ho, A. C., Humayun, M. S., Dorn, J. D., Da Cruz, L., Dagnelie, G., Handa, J., … Greenberg, R. J. (2015). Long-Term Results from an Epiretinal Prosthesis to Restore Sight to the Blind. *Ophthalmology*, *122*(8), 1547–1554. https://doi.org/10.1016/j.ophtha.2015.04.032

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*. Springer Unternational Publishing.

Horton, J. C., & Hocking, D. R. (1998). Monocular core zones and binocular border strips in primate striate cortex revealed by the contrasting effects of enucleation, eyelid suture, and retinal laser lesions on cytochrome oxidase activity. *Journal of Neuroscience*, *18*(14), 5433–5455. https://doi.org/10.1523/jneurosci.18-14-05433.1998

Horton, J. C., & Hoyt, W. F. (1991). The Representation of the Visual Field in Human Striate Cortex: A Revision of the Classic Holmes Map. *Archives of Ophthalmology*, *109*(6), 816–824. https://doi.org/10.1001/archopht.1991.01080060080030

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional

architecture in the cat's visual cortex. *The Journal of Physiology*, *160*, 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Hubel, David H., & Wiesel, T. N. (1959). Receptive Fields of Single Neurones in the Cat's Striate Cortex. *Journal of Physiology*, *148*, 574–591. https://doi.org/10.1109/SOCC.2011.6085109

Humayun, M. S., Dorn, J. D., Cruz, L., Dagnelie, G., Sahel, J., Stanga, P. E., … Greenberg, R. J. (2012). Interim Results from the International Trial of Second Sight's Visual Prosthesis. *Ophthalmology*, *119*(4), 779–788. https://doi.org/10.1016/j.ophtha.2011.09.028.Interim

Humayun, M. S., Weiland, J. D., Fujii, G. Y., Greenberg, R., Williamson, R., Little, J., … De Juan, E. (2003). Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Research*, *43*, 2573–2581. https://doi.org/10.1016/S0042-6989(03)00457-7

Hussain, Z., Astle, A. T., Webb, B. S., & McGraw, P. V. (2014). The challenges of developing a contrast-based video game for treatment of amblyopia. *Frontiers in Psychology*, *5*(OCT), 1–17. https://doi.org/10.3389/fpsyg.2014.01210

Italiano, M. L., Guo, T., Lovell, N. H., & Tsai, D. (2022). Improving the spatial resolution of artificial vision using midget retinal ganglion cell populations modeled at the human fovea. *Journal of Neural Engineering*, *19*(3). https://doi.org/10.1088/1741-2552/ac72c2

Jager, R. D., Mieler, W. F., & Miller, J. W. (2008). Age-Related Macular Degeneration. *New England Journal of Medicine*, *24*(358).

James, W. (1890). The perception of reality. *Principles of Psychology*, *2*, 283–324.

Jeter, P. E., Dosher, B. A., Petrov, A., & Lu, Z. L. (2009). Task precision at transfer determines

specificity of perceptual learning. *Journal of Vision*, *9*(3), 1–13. https://doi.org/10.1167/9.3.1

Jia, K., Zamboni, E., Kemper, V., Rua, C., Reis Goncalves, N., Ka Tsun Ng, A., … Kourtzi, Z. (2020). Recurrent Processing Drives Perceptual Plasticity. *Current Biology*, *30*(21), 4177-4187.e4. https://doi.org/10.1016/j.cub.2020.08.016

Kaas, J. H., Krubitzer, L. A., Chino, Y. M., Langston, A. L., Edward, H., & Blair, N. (1990). Reorganization of Retinotopic Cortical Maps in Adult Mammals After Lesions of the Retina Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2873942 Reorganization Reorganization of Retinotopic of Re. *Science*, *248*(4952), 229–231. Retrieved from http://www.annualreviews.org/doi/10.1146/annurev.neuro.051508.135516

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(11), 4966–4970. https://doi.org/10.1073/pnas.88.11.4966

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, *365*(6443), 250–252. https://doi.org/10.1038/365250a0

Kasowski, J., & Beyeler, M. (2022). Immersive Virtual Reality Simulations of Bionic Vision. In *Augmented Humans 2022* (pp. 82–93).

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

Kim, J. S., Greene, M. J., Zlateski, A., Lee, K., Richardson, M., Turaga, S. C., … Seung, H. S. (2014). Space-time wiring specificity supports direction selectivity in the retina. *Nature*, *509*(7500), 331–336. https://doi.org/10.1038/nature13240

Kingdom, F. A. A., Baldwin, A. S., & Schmidtmann, G. (2015). Modeling probability and additive summation for detection across multiple mechanisms under the assumptions of signal detection theory. *Journal of Vision*, *15*(5), 1–16. https://doi.org/10.1167/15.5.1

Koprinkova-hristova, P., Bocheva, N., Nedelcheva, S., Stefanova, M., Genova, B., Kraleva, R., & Kralev, V. (2020). STDP Plasticity in TRN Within Hierarchical Spike Timing Model of Visual Information Processing. In *FIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 279–290). Springer International Publishing. https://doi.org/10.1007/978-3-030-49161-1

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

Law, C., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory- motor , but not a sensory , cortical area. *Nature Neuroscience*, *11*(4), 505–513. https://doi.org/10.1038/nn2070

Lee, J. H., Wang, J. H., Chen, J., Li, F., Edwards, T. L., Hewitt, A. W., & Liu, G. S. (2019). Gene therapy for visual loss: Opportunities and concerns. *Progress in Retinal and Eye*

Research, *68*(August 2018), 31–53. https://doi.org/10.1016/j.preteyeres.2018.08.003

Levkovitz, Y., Caftori, R., Avital, A., & Richter-Levin, G. (2002). The SSRIs drug Fluoxetine, but not the noradrenergic tricyclic drug Desipramine, improves memory performance during acute major depression. *Brain Research Bulletin*, *58*(4), 345–350. https://doi.org/10.1016/S0361-9230(01)00780-8

Lewis, P. M., Ackland, H. M., Lowery, A. J., & Rosenfeld, J. V. (2015). Restoration of vision in blind individuals using bionic devices: A review with a focus on cortical visual prostheses. *Brain Research*, *1595*, 51–73. https://doi.org/10.1016/j.brainres.2014.11.020

Lewis, T. L., & Maurer, D. (2009). Effects of early pattern deprivation on visual development. *Optometry and Vision Science*, *86*(6), 640–646. https://doi.org/10.1097/OPX.0b013e3181a7296b

Li, N., & DiCarlo, J. J. (2008). Unsupervised Natural Experience Rapidly Alters Invariate Object Representation in Visual Cortex. *Science*, *321*(5895), 1502–1508.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, *12*(5), 549–551. https://doi.org/10.1038/nn.2296

Lim, L. S., Mitchell, P., Seddon, J. M., Holz, F. G., & Wong, T. Y. (2012). Age-related macular degeneration. *The Lancet*, *379*(9827), 1728–1738. https://doi.org/10.1016/S0140-6736(12)60282-7

Linietsky, J. (2014). Godot Engine - First public release!

Liu, Y., Fattah, N., & Degenaar, P. (2020). Newcastle Visual Prosthesis Implantable Control Unit. In *27th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*

(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

Lu, Y., Yan, Y., Chai, X., Ren, Q., Chen, Y., & Li, L. (2013). Electrical stimulation with a penetrating optic nerve electrode array elicits visuotopic cortical responses in cats. *Journal of Neural Engineering*, *10*(3). https://doi.org/10.1088/1741-2560/10/3/036022

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

Marc, R. E., Jones, B. W., Watt, C. B., & Strettoi, E. (2003). Neural remodeling in retinal degeneration. *Progress in Retinal and Eye Research*, *22*(5), 607–655. https://doi.org/10.1016/S1350-9462(03)00039-9

Marmor, M. F., & Wolfensberger, T. J. (1998). The retinal pigment epithelium. *Function and DIsease*, 103–134.

Masland, R. H. (1988). Amacrine cells. *Trends in Neurosciences*, *11*(9), 405–410. https://doi.org/10.1016/0166-2236(88)90078-1

Masland, R. H. (2012). The Neuronal Organization of the Retina. *Neuron*, *76*(2), 266–280. https://doi.org/10.1016/j.neuron.2012.10.002

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

Matthews, N., Liu, Z., Geesaman, B. J., & Qian, N. (1999). Perceptual learning on orientation and direction discrimination. *Vision Research*, *39*(22), 3692–3701. https://doi.org/10.1016/S0042-6989(99)00069-3

Maya Vetencourt, J., Sale, A., Viegi, A., Baroncelli, L., Pasquale, R. De, Leary, O. F. O., … Maffei, L. (2008). The Antidepressant Fluoxetine Restores Plasticity in the Adult Visual Cortex. *Science*, *320*, 385–388.

Mendelson, J. R., & Wells, E. F. (2002). Age-related changes in the visual cortex. *Vision Research*, *42*(6), 695–703.

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

Mokwa, W., Goertz, M., Koch, C., Krisch, I., Trieu, H. K., & Walter, P. (2008). Intraocular epiretinal prosthesis to restore vision in blind humans. *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 5790–5793. https://doi.org/10.1109/iembs.2008.4650530

Mollon, J. D., & Danilova, M. V. (1996). Three remarks on perceptual learning. *Spatial Vision*, *10*(1), 51–58. https://doi.org/10.1163/156856896X00051

Monge, M., Raj, M., Nazari, M. H., Chang, H. C., Zhao, Y., Weiland, J. D., … Emami, A. (2013). A fully intraocular high-density self-calibrating epiretinal prosthesis. *IEEE Transactions on Biomedical Circuits and Systems*, *7*(6), 747–760. https://doi.org/10.1109/TBCAS.2014.2298334

Montes, V. R., Gehlen, J., Lück, S., Mokwa, W., Müller, F., Walter, I. P., & Offenhäusser, A. (2019). Toward a bidirectional communication between retinal cells and a prosthetic device - A proof of concept. *Frontiers in Neuroscience*, *13*(APR), 1–19. https://doi.org/10.3389/fnins.2019.00367

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

Muqit, M., Velikay-Parel, M., Weber, M., Dupeyron, G., Audemard, D., Corcostegui, B., & Le Mer, Y. (2019). Six-month safety and efficacy of the intelligent retinal implant system II device in retinitis pigmentosa. *Ophthalmology*, *126*(4), 637–639.

Murphey, 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

Nagel, G., Szellas, T., Huhn, W., Kateriya, S., Adeishvili, N., Berthold, P., … Bamberg, E. (2003). Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(24), 13940–13945. https://doi.org/10.1073/pnas.1936192100

Najarpour Foroushani, A., Pack, C. C., & Sawan, M. (2018). Cortical visual prostheses: From microstimulation to functional percept. *Journal of Neural Engineering*, *15*(2). https://doi.org/10.1088/1741-2552/aaa904

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

Oei, A. C., & Patterson, M. D. (2015). Enhancing perceptual and attentional skills requires common demands between the action video games and transfer tasks. *Frontiers in Psychology*, *6*(Feb), 1–11. https://doi.org/10.3389/fpsyg.2015.00113

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

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology*, *15*(1), 33–58. https://doi.org/10.1016/0193-3973(94)90005-1

Op De Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience*, *26*(50), 13025–13036. https://doi.org/10.1523/JNEUROSCI.2481-06.2006

Oswalt, A. D., Datta, P., Talbot, N., Mirzadeh, Z., & Greger, B. (2020). A vision prosthesis

based on electrical stimulation of the primary visual cortex using epicortical microelectrodes. *BioRxiv*.

Pagon, R. A. (1988). Retinitis Pigmentosa. *Survey of Ophthalmology*, *33*(3), 137–177.

Palanker, D., Le Mer, Y., Mohand-Said, S., & Sahel, J. A. (2022). Simultaneous perception of prosthetic and natural vision in AMD patients. *Nature Communications*, *13*(1), 1–6. https://doi.org/10.1038/s41467-022-28125-x

Palanker, Daniel, 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

Pan, Z.-H., Lu, Q., Bi, A., Dizhoor, A. M., & Abrams, G. W. (2015). Optogenetic Approaches to Restoring Vision. *Annual Review of Vision Science*, *1*(1), 185–210. https://doi.org/10.1146/annurev-vision-082114-035532

Panetsos, F., Sanchez-Jimenez, A., Cerio, E. D. De, Diaz-Guemes, I., & Sanchez, F. M. (2011). Consistent phosphenes generated by electrical microstimulation of the visual thalamus. An experimental approach for thalamic visual neuroprostheses. *Frontiers in Neuroscience*, *5*(JUL), 1–12. https://doi.org/10.3389/fnins.2011.00084

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

Pellicciari, M. C., Miniussi, C., Rossini, P. M., & Gennaro, L. De. (2009). Increased Cortical Plasticity in the Elderly: Changes in the Somatosensory Cortex After Paired Associative Stimulation. *Neuroscience*, *163*(1), 266–276. https://doi.org/10.1016/j.neuroscience.2009.06.013

Peterhans, E., & von der Heydt, R. (1991). Subjective contours - bridging the gap between psychophysics and physiology. *Trends in Neurosciences*, *14*(3), 112–119. https://doi.org/10.1016/0166-2236(91)90072-3

Petrs-silva, H. (2014). Advances in gene therapy technologies to treat retinitis pigmentosa. *Clinical Ophthalmology*, *8*, 127–136.

Pezaris, J. S., & Reid, C. R. (2007). Demonstration of artificial visual percepts generated through thalamic microstimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7670–7675. https://doi.org/10.1073/pnas.0608563104

Picaud, S., & Sahel, J. A. (2014). Retinal prostheses: Clinical results and future challenges. *Comptes Rendus - Biologies*, *337*(3), 214–222. https://doi.org/10.1016/j.crvi.2014.01.001

Polat, U., Ma-Naim, T., Belkin, M., & Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(17), 6692–6697. https://doi.org/10.1073/pnas.0401200101

Polikov, V. S., Tresco, P. A., & Reichert, W. M. (2005). Response of brain tissue to chronically implanted neural electrodes. *Journal of Neuroscience Methods*, *148*, 1–18. https://doi.org/10.1016/j.jneumeth.2005.08.015

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

Pourtois, G., Rauss, K. S., Vuilleumier, P., & Schwartz, S. (2008). Effects of perceptual learning on primary visual cortex activity in humans. *Vision Research*, *48*(1), 55–62. https://doi.org/10.1016/j.visres.2007.10.027

Prem Senthil, M., Khadka, J., & Pesudovs, K. (2017). Seeing through their eyes: Lived experiences of people with retinitis pigmentosa. *Eye*, *31*(5), 741–748. https://doi.org/10.1038/eye.2016.315

Prosseda, P. P., Tran, M., Kowal, T., Wang, B., & Sun, Y. (2022). Advances in Ophthalmic Optogenetics: Approaches and Applications. *Biomolecules*, *12*(2), 1–13. https://doi.org/10.3390/biom12020269

Provansal, M., Marazova, K., Sahel, J. A., & Picaud, S. (2022). Vision Restoration by Optogenetic Therapy and Developments Toward Sonogenetic Therapy. *Translational Vision Science and Technology*, *11*(1), 1–8. https://doi.org/10.1167/tvst.11.1.18

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

Rodman, H. R., & Moore, T. (1997). Development and Plasticity of Extrastriate Visual Cortex in

Monkeys. In *Extrastriate Cortex in Primates* (pp. 639–672). https://doi.org/10.1007/978-1-4757-9625-4_14

Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., … Vanduffel, W. (2012). Toward a Unified Theory of Visual Area V4. *Neuron*, *74*(1), 12–29. https://doi.org/10.1016/j.neuron.2012.03.011

Rojer, 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.

Rosenzweig, M. R., & Bennett, E. L. (1996). Psychobiology of plasticity: Effects of training and experience on brain and behavior. *Behavioural Brain Research*, *78*(1), 57–65. https://doi.org/10.1016/0166-4328(95)00216-2

Roska, B., & Sahel, J.-A. (2018). Restoring vision. *Nature*, *557*(7705), 359–367. https://doi.org/10.1038/s41586-018-0076-4

Roth, Z. N., & Zohary, E. (2015). Fingerprints of learned object recognition seen in the fMRI activation patterns of lateral occipital complex. *Cerebral Cortex*, *25*(9), 2427–2439. https://doi.org/10.1093/cercor/bhu042

Rousche, P. J., & Normann, R. A. (1998). Chronic recording capability of the utah intracortical electrode array in cat sensory cortex. *Journal of Neuroscience Methods*, *82*, 1–15. https://doi.org/10.1016/S0165-0270(98)00031-4

Sabbah, N., Authié, C. N., Sanda, N., Mohand-Said, S., Sahel, J. A., & Safran, A. B. (2014). Importance of eye position on spatial localization in blind subjects wearing an argus II retinal prosthesis. *Investigative Ophthalmology and Visual Science*, *55*(12), 8259–8266. https://doi.org/10.1167/iovs.14-15392

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

Sanes, J. R., & Masland, R. H. (2015). The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification. *Annual Review of Neuroscience*, *38*, 221–246. https://doi.org/10.1146/annurev-neuro-071714-034120

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2009). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*(1), 53–60. https://doi.org/10.1038/nrn2737

Sato, M., & Stryker, M. P. (2008). Distinctive features of adult ocular dominance plasticity. *Journal of Neuroscience*, *28*(41), 10278–10286. https://doi.org/10.1523/JNEUROSCI.2451-08.2008

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

Schmid, L. M., Rosa, M. G. P., Calford, M. B., & Ambler, J. S. (1996). Visuotopic reorganization in the primary visual cortex of adult cats following monocular and binocular retinal lesions. *Cerebral Cortex*, *6*(3), 388–405. https://doi.org/10.1093/cercor/6.3.388

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

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, *412*(6846), 549–553. Retrieved from www.nature.com

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

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). https://doi.org/10.1201/9781466593534-27

Seitz, A. R. (2017). Perceptual learning. *Current Biology*, *27*(13), R631–R636. https://doi.org/10.1016/j.cub.2017.05.053

Seitz, A. R. (2020). Perceptual Learning : How Does the Visual Circuit Change through Experience? *Current Biology*, *30*(21), R1309–R1311. https://doi.org/10.1016/j.cub.2020.08.097

Seitz, A., & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, *9*(7), 329–334. https://doi.org/10.1016/j.tics.2005.05.010

Shah, N. P., & Chichilnisky, E. J. (2020). Computational challenges and opportunities for a bi-directional artificial retina. *Journal of Neural Engineering*, *17*. https://doi.org/10.1088/1741-2552/aba8b1

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)*, 714–718. https://doi.org/10.1109/NER.2019.8716987

Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics*, *52*(5), 582–588. https://doi.org/10.3758/BF03206720

Sinclair, N. C., Shivdasani, M. N., Perera, T., Gillespie, L. N., McDermott, H. J., Ayton, L. N., & Blamey, P. J. (2016). The appearance of phosphenes elicited using a suprachoroidal retinal prosthesis. *Investigative Ophthalmology and Visual Science*, *57*(11), 4948–4961. https://doi.org/10.1167/iovs.15-18991

Smirnakis, S. M., Brewer, A. A., Schmid, M. C., Tolias, A. S., Schuz, A., Augath, M., … Logothetis, N. K. (2005). Lack of long-term cortical reorganization after macaque retinal lesions. *Nature*, *435*(7040), 300–307. https://doi.org/10.1038/nature03495

Smith, A. T., Greenlee, M. W., Singh, K., Kraemer, F. M., & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging. *The Journal of Neuroscience*, *18*(10), 3816–3830. https://doi.org/10.1016/j.jat.2016.12.002

Smith, W., Assink, J., Klein, R., Mitchell, P., Klaver, C. C. W., Klein, B. E. K., … de Jong, P. T. V. M. (2001). Risk Factors for Age-related Macular Degeneration. *Ophthalmology*, *108*(4), 697–704.

Stingl, K., Bartz-Schmidt, K. U., Besch, D., Braun, A., Bruckmann, A., Gekeler, F., … Zrenner, E. (2013). Artificial vision with wirelessly powered subretinal electronic implant alpha-

IMS. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1757), 1–8. https://doi.org/10.1098/rspb.2013.0077

Stingl, K., Bartz-Schmidt, K. U., Besch, D., Chee, C. K., Cottriall, 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., Cottriall, 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

Striem-amit, E., Cohen, L., Dehaene, S., & Amedi, A. (2012). Reading with Sounds : Sensory Substitution Selectively Activates the Visual Word Form Area in the Blind. *Neuron*, *76*(3), 640–652. https://doi.org/10.1016/j.neuron.2012.08.026

Sunness, J. S., Liu, T., & Yantis, S. (2004). Retinotopic mapping of the visual cortex using functional magnetic resonance imaging in a patient with central scotomas from atrophic macular degeneration. *Ophthalmology*, *111*(8), 1595–1598. https://doi.org/10.1016/j.ophtha.2003.12.050

Szpiro, S. F. A., & Carrasco, M. (2015). Exogenous Attention Enables Perceptual Learning. *Psychological Science*, *26*(12), 1854–1862. https://doi.org/10.1177/0956797615598976

Team, R. C. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

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*, 0–7. https://doi.org/10.3389/fnins.2020.00262

Tootell, R. B. H., Mendola, J. D., Hadjikhani, N., Ledden, P. J., Liu, A. K., Reppas, J. B., … Dale, A. M. (1997). Functional analysis of V3A and related areas in human visual cortex. *The*, *17*(18), 7060–7078. https://doi.org/10.1109/CAOL.2003.1250524

Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The "Catch the Spies" task. *Cognitive Development*, *20*(3), 373–387. https://doi.org/10.1016/j.cogdev.2005.05.009

Troyk, P. R. (2017). *Artificial Vision*. Springer, Cham.

Tsai, D., Chen, S., Protti, D. A., Morley, J. W., Suaning, G. J., & Lovell, N. H. (2012). Responses of Retinal Ganglion Cells to Extracellular Electrical Stimulation, from Single Cell to Population: Model-Based Analysis. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0053357

Tschetter, W. W., Alam, N. M., Yee, C. W., Gorz, M., Douglas, R. M., Sagdullaev, B., & Prusky, G. T. (2013). Experience-enabled enhancement of adult visual cortex function. *Annals of Internal Medicine*, *158*(6), 5362–5366. https://doi.org/10.1523/JNEUROSCI.5229-12.2013

Tsushima, Y., & Watanabe, T. (2009). Roles of attention in perceptual learning from perspectives of psychophysics and animal learning. *Learning & Behavior*, *37*(2), 126–132.

Twyford, P., Cai, C., & Fried, S. (2014). Differential responses to high-frequency electrical stimulation in on and off retinal ganglion cells. *Journal of Neural Engineering*, *11*. https://doi.org/10.1088/1741-2560/11/2/025001

Vaina, L. M., Belliveau, J. W., Des Roziers, E. B., & Zeffiro, T. A. (1998). Neural systems underlying learning and representation of global motion. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(21), 12657–12662. https://doi.org/10.1073/pnas.95.21.12657

Van Gelder, R. N. (2015). Photochemical approaches to vision restoration. *Vision Research*, *111*, 134–141. https://doi.org/10.1016/j.visres.2015.02.001

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

Veraart, C., Raftopoulos, C., Mortimer, J. T., Delbeke, J., Pins, D., Michaux, G., … Wanet-Defalque, M. C. (1998). Visual sensations produced by optic nerve stimulation using an implanted self-sizing spiral cuff electrode. *Brain Research*, *813*(1), 181–186. https://doi.org/10.1016/S0006-8993(98)00977-9

Wandell, B. A. (1991). Computational neuroimaging of human visual cortex. *Annual Review of Neuroscience*, *22*, 145–173. https://doi.org/10.1016/b978-0-12-374620-7.00010-8

Wandell, B. A., & Smirnakis, S. M. (2009). Plasticity and stability of visual field maps in adult primary visual cortex. *Nature Reviews Neuroscience*, *10*(12), 873–884. https://doi.org/10.1038/nrn2741

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

Weitz, A. C., Nanduri, D., Behrend, M. R., Gonzalez-Calle, A., Greenberg, R. J., Humayun, M. S., … Weiland, J. D. (2015). Improving the spatial resolution of epiretinal implants by increasing stimulus pulse duration. *Science Translational Medicine*, *7*(318), 1–12. https://doi.org/10.1126/scitranslmed.aac4877

Wenger, E., Schaefer, S., Noack, H., Kühn, S., Mårtensson, J., Heinze, H. J., … Lövdén, M. (2012). Cortical thickness changes following spatial navigation training in adulthood and aging. *NeuroImage*, *59*(4), 3389–3397. https://doi.org/10.1016/j.neuroimage.2011.11.015

Wiesel, T. N., & Hubel, D. H. (1963). Effects of Visual Deprivation on Morphology and

Physiology of Cells in the Cat's Lateral Geniculate Body. *Journal of Neurophysiology*, *26*(6), 978–993.

Wilke, R., Gabel, V. P., Sachs, H., Schmidt, K. U. B., Gekeler, F., Besch, D., … Zrenner, E. (2011). Spatial resolution and perception of patterns mediated by a subretinal 16-electrode array in patients blinded by hereditary retinal dystrophies. *Investigative Ophthalmology and Visual Science*, *52*(8), 5995–6003. https://doi.org/10.1167/iovs.10-6946

Williams, N. R., & Okun, M. S. (2013). Deep brain stimulation (DBS) at the interface of neurology and psychiatry. *Journal of Clinical Investigation*, *123*(11), 4546–4556. https://doi.org/10.1172/JCI68341

Wong, W. L., Su, X., Li, X., Cheung, C. M. G., Klein, R., Cheng, C. Y., & Wong, T. Y. (2014). Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis. *The Lancet Global Health*, *2*(2), e106–e116. https://doi.org/10.1016/S2214-109X(13)70145-1

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

y Cajal, R. (1972). *The Structure of the Retina*. CC Thomas, Springfield Illinois.

Yamasaki, D. S., & Wurtz, R. H. (1991). Recovery of function after lesions in the superior temporal sulcus in the monkey. *Journal of Neurophysiology*, *66*(3), 651–673. https://doi.org/10.1152/jn.1991.66.3.651

Yang, T., & Maunsell, J. H. R. (2004). The Effect of Perceptual Learning on Neuronal

Responses in Monkey Visual Area V4. *Journal of Neuroscience*, *24*(7), 1617–1626. https://doi.org/10.1523/JNEUROSCI.4442-03.2004

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

Zeng, F. G. (2004). Trends in Cochlear Implants. *Trends in Amplification*, *8*(1), 1–34. https://doi.org/10.1177/108471380400800102

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

Zheng, Y., Jia, S., Yu, Z., Liu, J. K., & Huang, T. (2021). Unraveling neural coding of dynamic natural visual scenes via convolutional recurrent neural networks. *Patterns*, *2*(10), 100350. https://doi.org/10.1016/j.patter.2021.100350

Zrenner, E. (2002). Will retinal implants restore vision? *Science*, *295*(5557), 1022–1025. https://doi.org/10.1126/science.1067996