



Published in final edited form as:

Vision Res. 2021 May ; 182: 58–68. doi:10.1016/j.visres.2021.01.008.

Multisensory Perception in Argus II Retinal Prosthesis Patients: Leveraging Auditory-Visual Mappings to Enhance Prosthesis Outcomes

Noelle R. B. Stiles^{1,2,*}, Vivek R. Patel¹, James D. Weiland^{2,3}

¹Department of Ophthalmology, University of Southern California, 1450 San Pablo Street, Los Angeles, CA, 90033, USA

²Department of Biomedical Engineering, University of Michigan, 2800 Plymouth Road, Ann Arbor, MI, 48109, USA

³Department of Ophthalmology and Visual Sciences, University of Michigan, 1000 Wall Street, Ann Arbor, MI, 48109, USA

Abstract

Crossmodal mappings associate features (such as spatial location) between audition and vision, thereby aiding sensory binding and perceptual accuracy. Previously, it has been unclear whether patients with artificial vision will develop crossmodal mappings despite the low spatial and temporal resolution of their visual perception (particularly in light of the remodeling of the retina and visual cortex that takes place during decades of vision loss). To address this question, we studied crossmodal mappings psychophysically in Retinitis Pigmentosa patients with partial visual restoration by means of Argus II retinal prostheses, which incorporate an electrode array implanted on the retinal surface that stimulates still-viable ganglion cells with a video stream from a head-mounted camera. We found that Argus II patients ($N = 10$) exhibit significant crossmodal mappings between auditory location and visual location, and between auditory pitch and visual elevation, equivalent to those of age-matched sighted controls ($N = 10$). Furthermore, Argus II patients ($N = 6$) were able to use crossmodal mappings to locate a visual target more quickly with auditory cueing than without.

Overall, restored artificial vision was shown to interact with audition via crossmodal mappings, which implies that the reorganization during blindness and the limitations of artificial vision did not prevent the relearning of crossmodal mappings. In particular, cueing based on crossmodal

*Correspondence to: nstiles@usc.edu.

Author Contributions

Conceptualization, N.R.B.S.; Methodology, N.R.B.S, V.R.P., and J.D.W.; Resources, N.R.B.S, V.R.P, and J.D.W.; Writing – Original Draft, N.R.B.S; Writing – Review & Editing, N.R.B.S, V.R.P., and J.D.W.; Funding Acquisition, N.R.B.S, V.R.P., and J.D.W.; Supervision, V.R.P. and J.D.W.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Competing Interests

The authors declare no competing interests.

mappings was shown to improve visual search with a retinal prosthesis. This result represents a key first step toward leveraging crossmodal interactions for improved patient visual functionality.

Keywords

Retinal Prostheses; Auditory-Visual Integration; Crossmodal Mappings; Vision Restoration; Blindness

1. Introduction

Intraocular retinal prostheses are capable of restoring visual function to those who have become fully blind (light perception or less) in adulthood with Retinitis Pigmentosa (RP) (Ayton, Barnes, Dagnelie, Fujikado, Goetz, Hornig, Jones, Muqit, Rathbun & Stingl, 2020, Da Cruz, Dorn, Humayun, Dagnelie, Handa, Barale, Sahel, Stanga, Hafezi & Safran, 2016, Luo & Da Cruz, 2016, Stiles, McIntosh, Nasiatka, Hauer, Weiland, Humayun & Tanguay Jr, 2011, Stingl, Bartz-Schmidt, Besch, Chee, Cottrill, Gekeler, Groppe, Jackson, MacLaren & Koitschev, 2015, Stronks & Dagnelie, 2014, Weiland, Cho & Humayun, 2011, Zhou, Dorn & Greenberg, 2013). For example, the Argus II retinal prosthesis consists of a 60-element microelectrode array that electrically stimulates retinal cells based on visual information captured by a camera mounted on glasses (Fig. 1). Numerous studies have documented the visual experience of Argus II recipients, as well as other visual prostheses such as the Alpha-IMS (Da Cruz et al., 2016, Dagnelie, 2018, Luo & Da Cruz, 2016, Stingl et al., 2015, Stronks & Dagnelie, 2014, Weiland et al., 2011, Zhou et al., 2013). Visual perception with retinal prostheses often begins with basic light perception, followed by the perception of visual flashes (phosphenes). With persistent rehabilitative training and device use, a subset of patients begin to see forms, and in some cases to read letters and demonstrate improved navigation (Dagnelie, 2017, Dagnelie, 2018). These patient outcomes are relatively consistent between the two commercially available devices (the Argus II and Alpha-IMS), and are characterized as ultra-low vision (Dagnelie, 2018).

Nevertheless, Argus II prosthetic vision has been shown to differ from normally sighted vision in its spatial and temporal properties, and has been classified as a new form of perception that relies on a “*lexicon of flashes*” (Erickson-Davis & Korzybska, 2020). In particular, Argus II patients exhibit spatial acuities that can range from basic shape perception down to only a few pixels, phosphenes, or bright flashes (Castaldi, Lunghi & Morrone, 2020, Erickson-Davis & Korzybska, 2020). The phosphenes are often irregularly shaped (due to axonal stimulation) (Beyeler, Nanduri, Weiland, Rokem, Boynton & Fine, 2019), are limited in number of brightness levels, and are variable in size (Castaldi et al., 2020, Luo, Zhong, Clemo & Da Cruz, 2016, Luo & Da Cruz, 2016, Zhou et al., 2013). In addition, Argus II perception has a low temporal resolution accompanied by the fading of stimuli over time (Luo et al., 2016, Luo & Da Cruz, 2016, Zhou et al., 2013). Finally, perception with the Argus II retinal prosthesis is derived from a head-mounted camera and has a narrow field of view, requiring head-scanning of the environment for most visual tasks (Luo et al., 2016, Luo & Da Cruz, 2016, Zhou et al., 2013). All of these constraints in combination with a retina and cortex remodeled during blindness (Castaldi et al., 2020,

Cunningham, Shi, Weiland, Falabella, de Koo, Zacks & Tjan, 2015, Mowad, Willett, Mahmoudian, Lipin, Heinecke, Maguire, Bennett & Ashtari, 2020) have the consequence that Argus II patients require months of training to perform basic visual tasks.

Given all of these differences relative to natural vision, the new and unusual provision of artificial perception by means of retinal prostheses such as the Argus II may not be automatically associated with audition in the same way that natural vision is, and may need to be acquired through post-implantation experience that improves with duration of use. Nevertheless, auditory-visual interactions (if they prove to be feasible with artificial vision) could generate a more holistic perception of the environment, with the potential for improvements in the tasks of daily living through multisensory training and cueing (Cappagli, Finocchietti, Cocchi, Giammari, Zumiani, Cuppone, Baud-Bovy & Gori, 2019, Frassinetti, Bolognini, Bottari, Bonora & Ladavas, 2005, Seitz, Kim & Shams, 2006, Tinelli, Purpura & Cioni, 2015).

In this paper, we take the first step toward evaluating auditory-visual interactions in artificial vision patients by studying crossmodal mappings, which is one type of multisensory crosstalk that is common in sighted individuals. Crossmodal mappings are associations of perceptual features across the senses. These sensory mappings can be based on structural (*e.g.*, temporal or spatial) similarities in sensory signals, on statistical regularities characteristic of the sensory environment, or on semantic or linguistic representations that bridge the senses (Bernstein & Edelman, 1971, Deroy & Spence, 2016, Klapetek, Ngo & Spence, 2012, Mudd, 1963, Parise, Knorre & Ernst, 2014, Pratt, 1930, Roffler & Butler, 1968, Spence, 2011, Stiles & Shimojo, 2015). Crossmodal mappings are often measured in psychophysical experiments by simple stimuli (flashes and beeps), which can be presented with sufficient duration (multiple seconds) that the stimuli can be perceived even with the ultra-low resolution vision characteristic of retinal prostheses.

We evaluated two basic crossmodal mappings in Argus II retinal prosthesis users and sighted controls, the mapping of spatial location between audition and vision (Spence, 2011), and the relative mapping of auditory pitch to visual elevation (Mudd, 1963, Pratt, 1930, Roffler & Butler, 1968). These two mappings were selected for testing in Argus II patients based on their robustness and prevalence in the sighted population (Spence, 2011).

Crossmodal mappings have also been used with low vision patients to cue visual attention to a target during visual search tasks, thereby speeding task performance (Bernstein & Edelman, 1971, Bolognini, Rasi, Coccia & Ladavas, 2005, Frassinetti et al., 2005, Klapetek et al., 2012). Visual search cueing is a potentially useful psychophysical measure for the Argus II population, as cueing could provide a means to ease the cognitive load of interpreting artificial vision and improve the typically slow visual search times for these patients (Liu, Stiles & Meister, 2018, Mante & Weiland, 2018, Parikh, Itti, Humayun & Weiland, 2013). In this paper, we evaluated visual search tasks with auditory cueing to determine whether cueing based on spatial location could improve Argus II visual search speed.

2. Methods

2.1. Participant Details

Ten Argus II patients (5 male and 5 female) participated in the crossmodal matching experiment (Experiment 1 in Fig. 2 and Fig. 3A, demographic information in Table 1). Six of the ten participants (A4 and A6 – A10; 4 male and 2 female) also performed the crossmodal matching task with randomized sound order (Experiment 2), the visual search task (Experiment 3 in Fig. 4A,) and the timed localization task (Experiment 4 in Fig. S6). The Argus II patients have vision loss due to Retinitis Pigmentosa and were an average age of 65.1 years old ($SD = 9.42$ years) at the time of the experiment. The Argus II patients had the Argus II retinal prosthesis device implanted for a minimum of 10 months when they were tested ($M = 27.40$ months, $SD = 10.58$ months) (Fig. 2 and Table 1). The patients had vision loss for a minimum of 16 years when they performed the experiments herein ($M = 26.00$ years, $SD = 11.11$ years) and had natural visual perception of light perception or less. If an Argus II patient reported light perception in either eye, that eye was covered with an eye patch during testing to prevent it from interfering with the visual task (with the exception of one eye in Patient A2; details are provided in Table 1). Additional patient information is provided in Table 1.

Two of the Argus II patients (A4 and A9) that were tested had been implanted previously with the earlier generation of the Argus device (Argus I; 4×4 microstimulator array) in one eye; the Argus II was then implanted in the other eye. As the Argus II implanted eye was not previously implanted with the Argus I, which in any case is significantly lower in resolution, this period of Argus I perception was included within the period of blindness in Table 1.

The crossmodal matching experiments (Experiment 1, and Experiment 2) with Argus II patients were performed either at the University of Southern California or in the case of Experiment 1 also at the University of Michigan. The same researchers set up and supervised the task at both locations with the same computer and black felt setup. The visual search experiment (Experiment 3), and timed localization experiment (Experiment 4) were performed only at the University of Southern California.

Ten sighted control individuals (3 male and 7 female) participated in the crossmodal matching experiments (Experiment 1, and Experiment 2) at the University of Southern California (Table 2). The sighted participants had an average age of 63.5 years ($SD = 4.70$ years). Sighted participants performed the same tasks as the Argus II patients, with the experimenter providing oral instructions and entering the participant responses on the keyboard.

All experiments were approved by either the University of Southern California or the University of Michigan Institutional Review Boards, as appropriate, and all participants gave informed written consent. This research adhered to the Declaration of Helsinki.

Table 1 details the demographic information and visual perception capabilities of the Argus II patients. The visual disease that caused blindness in each of the patients is Retinitis Pigmentosa (RP). The remaining natural visual perception for each eye (as reported by the

patient) is presented in Columns 6 and 7, with the indications of “LP” for light perception and “No LP” for no light perception. The use of an eye patch during the experiment for each of the patients’ eyes is indicated in Columns 8 and 9, with “Yes” for the use of an eye patch and “No” for the absence of an eye patch. All participants who reported light perception used an eye patch (with the exception of the right eye for A2, due to a limited supply of available eye patches). See also Table S1.

Table 2 details the demographic information and perceptual capabilities (self-reported) of the sighted control participants. The fourth column details whether the participants reported wearing glasses or contacts for either reading or distance vision. The final two columns detail the self-reported hearing loss of each participant as: No (None), mild (Mild), moderate (Mod), or severe (Sev) hearing loss.

2.2. Experimental Setup

All of the experimental tasks were performed on a 27-inch iMac computer (2560 by 1440 pixel resolution), placed on a black felt covered table in front of a black felt covered wall and window (shown in Fig. 3A). The experimental room was dimly lit during the task, with the interior lights turned off but an adjacent window partially uncovered (all of the tasks were performed during the day). Participants sat in an office swivel chair about 20 inches from the computer monitor on which visual stimuli were displayed. The experimenter sat on the left side of the participant, providing oral instructions and entering all participant responses using a keyboard. Participants in all of the experimental tasks used Sony MDR-ZX110NC Headphones (placed over the Argus II retinal prosthesis device in the patients tested). The sound volume was adjusted to be comfortable for the participant performing the task (sounds were played in MATLAB by using the sound function at the default setting (*i.e.*, 8192 sample rate)).

If the Argus II patients had light perception, they wore an opaque mask that prevented any natural visual perception (Table 1). The Argus II participants were encouraged to feel the edges of the computer monitor with their hands in order to determine the position of the visual stimuli relative to the screen’s edges. While the University of Southern California and the University of Michigan experimental setups were in different locations, both rooms were designed to include all of the features detailed above.

2.3. Experiment 1

2.3.1. Crossmodal Mappings Tested—Two crossmodal mappings were tested: spatial location matching between auditory and visual stimuli, and the matching between auditory stimuli pitch and visual stimuli elevation. (Note: Our aim for the second crossmodal mapping was to test for crossmodal mapping of *relative* auditory pitch to visual elevation). Location matching can be considered a structural crossmodal mapping, where the common structure of the brain’s spatial sensory mapping for audition and vision enables the two sensory spatial maps to be crossmodally associated (Spence, 2011). The crossmodal mapping of pitch and elevation is a semantically mediated correspondence, in which the language for representing the stimuli is common between the two modalities (*i.e.* high elevation, and high pitch) (Spence, 2011).

Correct crossmodal stimulus pairings include: (1) For the auditory-visual spatial location matching task, the left located sound paired with the left located flash, and two other location matched pairs; (2) For the auditory pitch to visual elevation matching task, the high pitch sound paired with the horizontal bar at the top of the screen, and two other similarly matched pairs. The auditory-to-visual (crossmodal) matching fraction correct in each case was first determined for each participant in each task, excluding all trials in which the visual stimulus was incorrectly localized, and then averaged over the participants.

Experiment 1 (Argus II patients and sighted control participants) was divided into two blocks; the first block tested for auditory-visual location matching, and the second block tested for auditory pitch to visual elevation matching (Fig S1).

2.3.2. Experiment 1 Block 1 Auditory and Visual Stimuli—For the spatial location matching task (Block 1), the visual stimuli were large *vertical* bars presented on a computer monitor on the left, in the center, or on the right. The vertical bar stimulus was 2.75 inches wide by 13.13 inches high or approximately 8 degrees by 36 degrees visual angle on the left edge, center, or right edge of the screen. The location of the visual stimulus was randomized among the three possible locations, and the auditory stimuli were presented in left to right order. Randomization was performed with the randperm function (default settings) in MATLAB.

The auditory stimuli were beeps presented on the left, in the center, or on the right. The auditory stimulus for Block 1 was a 200 Hz tone, with left-right amplitude differences indicating its location. The left sound was generated by 100% amplitude in the left ear and 0% amplitude in the right ear, the center sound by 50% amplitude in the left and right ears, and the right sound by 0% amplitude in the left ear and 100% amplitude in the right ear. For both Block 1 and Block 2 of Experiment 1, the auditory stimuli were 0.5 seconds long. Participants performed 18 trials in Experiment 1 Block 1, which included 6 trials for each visual stimulus location (6 trials \times 3 locations = 18 trials).

2.3.3. Experiment 1 Block 2 Auditory and Visual Stimuli—For the auditory pitch to visual elevation matching task, the visual stimuli were large *horizontal* bars presented on a computer monitor at the top, in the middle, or at the bottom of the screen. The horizontal bars were 23.25 inches wide by 2.75 inches high or approximately 60 degrees by 8 degrees visual angle (full width of the computer screen). The horizontal bars could be located at the top edge of the screen, in the vertical center of the screen, or at the bottom edge of the screen. The location of the visual stimulus was randomized among the three possible locations, and the auditory stimuli were presented in order from high pitch to low pitch. Randomization was performed with the randperm function (default settings) in MATLAB.

The auditory stimuli were beeps with high, middle, or low pitch. The auditory stimuli for Task 2 were centrally located tones (50% amplitude in the left and right ears), with the following potential frequencies: 800 Hz (high frequency tone), 300 Hz (mid-frequency tone), and 50 Hz (low frequency tone). The auditory stimuli were generated in MATLAB using the sine function. Participants performed 18 trials in Experiment 1 Block 2, which included 6 trials for each visual stimulus location (6 trials \times 3 locations = 18 trials).

2.3.4. Experiment 1 Procedure—For both crossmodal mappings tested, patients were asked to view a visual stimulus and indicate the location of the visual stimulus, then three sounds were played and participants selected the sound that seemed to best match the visual stimulus (Fig S1). Participants were not provided with any additional instructions on matching the stimuli. The auditory stimuli were not described beyond their order (First, Second, and Third sound), in order to prevent semantic indicators of crossmodal matching.

Both Experiment 1 blocks began when the visual stimulus started flashing on the computer monitor. Each visual stimulus flashed on for 0.5 seconds then off for 0.5 seconds ten times in succession. While the visual stimulus was flashing, the participant was asked to locate the flashing visual stimulus on the screen. Following these ten flashes, the visual stimulus remained stationary on the screen until the experimenter pressed the 1 key on the keyboard. The participant then verbally reported left, center, or right for Experiment 1 Block 1 (auditory-visual location matching) or top, middle, or bottom for Experiment 1 Block 2 (auditory pitch to visual elevation matching). The experimenter then pressed the 1 key to proceed, at which time three sounds played. For Block 1 (auditory-visual location matching), the sounds were Left, Center, and then Right. For Block 2 (auditory pitch to visual elevation matching), the sounds were High, Middle, and then Low pitch. The participant could then either report the sound they thought best matched the visual stimulus by choosing the first, second, or third sound, or ask to have the sounds replayed. If the sounds were requested to be replayed, the experimenter pressed the 1 key, and the three sounds played again in the participant's headphones. When the participant reported their answer, the experimenter pressed the 2 key and then entered both the auditory sound and the visual location reported by the participant.

2.4. Experiment 2

While in Experiment 1 the order of the visual stimuli is randomized, the order of the auditory stimuli presented to the participants is not randomized by design. In contrast, the order of *both* the auditory and visual stimuli presented to participants is randomized in Experiment 2 in order to further fortify the results garnered in Experiment 1. In Experiment 2 (like in Experiment 1) the matching of auditory-visual spatial location and the matching of auditory pitch to visual elevation are both evaluated. The experimental blocks, stimuli, and procedures for Experiment 2 were the same as for Experiment 1 in both the Argus II patients and the sighted participants. Experiment 2 was performed after Experiment 1, and the same instructions were repeated for Experiment 2 as for Experiment 1.

2.5. Experiment 3

2.5.1. Experiment 3 Blocks—In Experiment 3 a visual search task was performed by Argus II patients with a cluttered visual stimulus (including a visual search target and visual search distractors). Experiment 3 was divided into three blocks (Fig S2). In the first block Argus II patients were trained to localize the visual search target among distractors (with feedback from the experimenter). (Note: Training was required for this visual search task as patients have difficulty with perceiving shapes and therefore needed to be directly trained on distinguishing the small disk distractors from the large rectangular target. The visual task was designed to be challenging to push patients to use the auditory information as a

supplement to the visual information.) The second Experiment 3 block presented co-located auditory and visual stimuli (visual target and distractors). The third block had the visual stimuli presented alone (baseline visual search). The goal for all three blocks was for the patient to locate the visual target as quickly and accurately as possible. Participants performed 18 trials in each Block of Experiment 3, which included 6 trials for each visual stimulus location (6 trials \times 3 locations = 18 trials per block, with 3 blocks for a total of 54 trials).

2.5.2. Experiment 3 Auditory and Visual Stimuli—The target of the visual search was a large white vertical bar while the visual search distractors were two vertical columns of three aligned circles (shown in Fig. 4A). The large vertical bar was 2.75 inches wide by 13.13 inches high or approximately 8 degrees by 36 degrees visual angle, and was located either on the far left, center or the far right of the screen. The vertical bars on the far left and far right of the screen were shifted 1.19 inches inward from the monitor edge. For each location of the vertical bar, two vertical columns of three disks each (1.38 inches (approximately 4 degrees) in diameter with a 3.69 inch vertical center-to-center distance) were located in the two other vertical bar locations. When located on the far left or far right of the screen, the disks were 2.56 inches from each disk center to the adjacent vertical edge of the screen. The disks in the center of the screen were then located 9.13 inches center-to-center distance from circles on the far left or far right of the screen (horizontal distance). For example, if the vertical bar was on the left edge of the screen, a column of disk distractors would be present in the center of the screen and on the far right edge of the screen. The location of the vertical bar was randomized among the three possible locations using the randperm function (default settings) in MATLAB. Fig. 4A shows a typical image of the screen during the experiment.

The auditory stimulus was a tone of 2,731 Hz lasting 0.07 seconds, with left-right amplitude differences indicating its location. The left sound was generated by 100% amplitude in the left ear and 0% amplitude in the right ear, the center sound by 50% amplitude in the left and right ears, and the right sound by 0% amplitude in the left ear and 100% amplitude in the right ear. The auditory beeps were synchronized to start with the onset of each visual flash. The location of the auditory beep was matched to the location of the vertical bar in Block 2 of Experiment 3 (Blocks 1 and 3 had no auditory stimuli).

The auditory cueing for the visual search was designed to provide a directional cue, (*i.e.* left, right, or center), without the exact location of the visual target. Participants were told to use the auditory cues to aid with visual search, but also that they were required to locate the visual target with their artificial vision before responding. Precise auditory cues with Head-Related Transfer Functions (HRTFs) were not used in order to prevent the sole use of auditory information for localization (*i.e.* the participant would not use their vision to locate the target but simply point by auditory information alone to the target locations). This is of particular concern given that the targets must be large for the participant to locate them with the Argus II device. In addition, the visual processing with Argus II is slow and requires significant visual attention, and therefore may not be used if auditory information is sufficient to complete the task. Therefore, to ensure that the participant did in fact use the Argus II to locate the target, auditory stimuli were provided that were directional cues rather

than the precise locations of the visual targets. The head-movements of the participant were also monitored to confirm that patients looked for the target by actively scanning their camera across the computer monitor, and that their head was pointed in the direction of the visual stimulus when they located it.

2.5.3. Experiment 3 Procedure—Each trial within the task began when the visual stimulus began flashing (on for 0.5 seconds and then off for 0.5 seconds) (Fig S2). In four of the Argus II patients tested, the experimenter alerted the patient when the visual stimulus started flashing on the screen to indicate that they should start their visual search (the remaining two patients, A4 and A7, performed visual search continuously rather than being cued to begin searching on each trial). During all trials, the patients searched for the target until they located it, and then they were asked to point to the target on the screen with one finger. Once the target was located, the experimenter pressed 1 on the keyboard, which made the visual stimulus disappear. The experimenter then entered the visual flash location that the participant reported via the keyboard. The computer recorded the reported visual flash location as well as the time between the flash onset and the visual target localization. Feedback during training (Block 1) was verbal (such as “the target is farther to the left or right”).

2.6. Experiment 4

2.6.1. Experiment 4 Design—Six Argus II patients also performed Experiment 4, a timed localization experiment similar to Experiment 3. During Experiment 4 patients performed a localization task with both congruent and incongruent auditory-visual stimuli, and with no visual clutter (visual target presented alone) (Fig S3).

2.6.2. Experiment 4 Auditory and Visual Stimuli—The visual stimulus in Experiment 4 was the randomized vertical bar stimulus presented in Experiment 3 (with no disk distractor stimuli presented). (Note: The absence of distractors in comparison to Experiment 3 made the Experiment 4 visual task significantly easier. Experiment 3 used visual distractors to purposely make the visual task difficult. The visual task difficulty in Experiment 3 pushed patients to use the auditory information to ease their visual processing load. Experiment 4 was a simple task (no visual distractors) and therefore patients had less of an imperative to use auditory cueing. Consequently, the use of auditory information in Experiment 4 was substantially diminished relative to Experiment 3.)

The auditory stimuli in Experiment 4 were the same as Experiment 3. The location of the auditory beep was matched to the location of vertical bar in Block 1 of Experiment 4 (congruent condition), but was in a different location in Block 2 (incongruent condition) (Block 3 had visual stimuli alone). Participants performed 18 trials in each block of Experiment 4, which included six trials for each visual stimulus location for Blocks 1 and 3 (6 trials × 3 locations = 18 trials in Blocks 1 and 3). The auditory-visual stimuli (same location, Block 1) or just the visual stimuli alone (Block 3) were randomized in order. For Block 2, the sounds were presented in a location that was different from the flash location. Therefore, of three possible sound locations (left, center, or right), the sound could be in one of the two locations that were different than that of the visual flash location. There were

three flash locations and two possible sound locations different than the flash location, which generated six stimulus combinations with three repeats of each combination (3 trials \times 6 stimuli combinations = 18 trials in Block 2). The six auditory-visual stimulus combinations for Block 2 were randomized in order.

2.6.3. Experiment 4 Procedure—The timing of the task and stimuli in Experiment 4 were the same as Experiment 3 (Fig S3). In Blocks 1 and 2 of Experiment 4, an auditory tone was co-presented with the visual flash. Participants were told that there may or may not be an auditory stimulus present, and that the auditory stimulus may or may not be helpful in locating the visual flash. During each trial, the experimenter told the patient when the visual stimulus had begun to indicate that they should start their visual search.

2.7. Argus II Retinal Prosthesis System

The Argus II retinal prosthesis system is manufactured and sold by Second Sight Medical Products, located in Sylmar, California, USA. It was approved by the FDA in 2013 and received a CE mark in Europe in 2011 for implantation in patients with severe Retinitis Pigmentosa (bare light perception or less). The device consists of a microelectrode array with 60 electrodes in a 6×10 array that is proximity coupled to the retinal surface for direct stimulation of retinal ganglion cells via electrical pulses. The electrodes are each 200 μm in diameter, and the entire array subtends about 20 degrees of visual angle diagonally. The microelectrode array is connected via wire to a microelectronics package that is attached to the sclera of the eye by a scleral buckle. The microelectronics package receives both power and information from an external transmitter coil mounted on a pair of glasses through a receiver coil that is incorporated in the scleral buckle adjacent to the microelectronics package. The pair of glasses also incorporates a small camera that transmits a live video feed via a cable to a Visual Processing Unit (VPU). The small computing device, or VPU, processes the video and translates the visual information into stimulation parameters, and then sends this information back through the cable to the glasses-mounted coil for transmission to the microelectrode array. For additional device design details see Zhou *et al.* and Luo *et al.* (Luo, Zhong, Merlini, Anafloos, Arsiero, Stanga & Da Cruz, 2014, Zhou *et al.*, 2013).

Second Sight estimates that over 350 Argus II devices have been implanted worldwide (2019). Implantation surgeries occur at select medical centers and hospitals (20 locations are listed on the Second Sight Website) with specially trained ophthalmological surgeons and clinical coordinator staff. Following the surgical implantation of the device, four weeks are typically set aside for ocular healing and recovery (2019). At that time, the patient returns to the clinic for device setup, where the thresholds for each of the electrodes are determined. After the electrode thresholds are estimated for safe retinal stimulation, the patient turns on the Argus II device for the first time, and has their first visual experience in the clinic. The patient is then permitted to take the Argus II device home for use in their daily life. A subset of patients also return for a second visit to the clinic to perform basic functional tasks for Second Sight such as motion detection and object localization on a touch screen; these clinical trial results were published in Humayun *et al.* (Humayun, Dorn, Da Cruz, Dagnelie, Sahel, Stanga, Cideciyan, Duncan, Elliott & Filley, 2012).

Following device implantation and activation, Argus II patients can choose whether they would like to participate in the visual rehabilitation training provided by Second Sight. This visual rehabilitation training is structured to aid the patient with learning to use the prosthesis in the tasks of daily living, and is often performed in their own home (McDonald, 2019). It starts with basic visual tasks (such as seeing light and dark, or identifying white stripes on a black fabric), and progresses to more complex tasks (such as navigation in natural environments). On average this rehabilitation training lasts up to 3 to 4 hours in duration for each of 3 or more visits by a visual rehabilitation specialist. For the patients with the Argus II device that we tested, 9 out of 10 had performed this rehabilitation training with Second Sight. In addition, these patients used their device 3 days per week on average after rehabilitation training. Detailed training information for each Argus II patient is listed in Table S1 in the Supplemental Information.

2.8. Quantification and Statistical Analyses

One-sample two-tailed student t-tests were used to assess the significance of results relative to chance (chance is 1/3 for all of the 3AFC tasks) (MATLAB function t-test, default settings). Paired sample two-tailed t-tests were used to calculate the significance of differences between distinct measures made within the same participant group (paired t-test, MATLAB function t-test, default settings). If a paired sample t-test was used, this is specified in the Results section within the set of statistical metrics. Two-sample (unpaired) two-tailed t-tests were used to calculate the significance of differences between participant groups (unpaired t-test, MATLAB function t-test2, default settings). A measurement was deemed significant if the given t-test generated a p -value of less than 0.05. All correlation analyses were performed with the corr function in MATLAB.

3. Results

3.1. Crossmodal Mappings in Argus II Patients (N = 10) (Experiment 1)

The fraction correct for crossmodal matching in ten Argus II patients was significantly above chance for both auditory-visual spatial location matching, and auditory pitch to visual elevation matching (Auditory-visual spatial location matching task, $M = 0.97$, $SD = 0.09$, $t(9) = 22.61$, $p = 3.07 \times 10^{-9}$; Auditory pitch to visual elevation matching task, $M = 0.86$, $SD = 0.25$, $t(9) = 6.71$, $p = 8.78 \times 10^{-5}$) (Fig. 3). In addition, the fraction of trials in which the visual stimulus was correctly localized was significantly above chance for both crossmodal matching tasks (Auditory-visual spatial location matching task, *i.e.*, vertical bar localization, $M = 0.97$, $SD = 0.05$, $t(9) = 42.61$, $p = 1.08 \times 10^{-11}$; Auditory pitch to visual elevation matching task, *i.e.*, horizontal bar localization, $M = 0.73$, $SD = 0.23$, $t(9) = 5.39$, $p = 4.40 \times 10^{-4}$).

The crossmodal matching fraction correct was not significantly correlated with participant age, duration of Argus II device use, or duration of blindness in either auditory-to-visual matching task.

Argus II patients correctly matched auditory and visual stimuli for two different crossmodal mappings significantly above chance. This result confirms that crossmodal mappings can be reinstated after decades of blindness.

3.2. Crossmodal Mappings in Sighted Controls (N = 10) Compared to Argus II Patients (N = 10) (Experiment 1)

Age-matched sighted controls also performed Experiment 1 (the auditory-visual matching experiment) using natural vision and audition (Sighted participant age: $M = 63.5$ years, $SD = 4.70$ years; Argus II patient age: $M = 65.1$ years, $SD = 9.42$ years). The fraction correct in crossmodal matching for sighted participants (with the incorrect visual localization trials removed, as described in Experiment 1) was significantly above chance for both auditory-visual spatial location matching and auditory pitch to visual elevation matching (Auditory-visual spatial location matching task, $M = 0.98$, $SD = 0.04$, $t(9) = 54.45$, $p = 1.19 \times 10^{-12}$; Auditory pitch to visual elevation matching task, $M = 0.96$, $SD = 0.07$, $t(9) = 27.07$, $p = 6.19 \times 10^{-10}$). In addition, the visual localization fraction correct was significantly above chance for the sighted participants, as it was for the Argus II patients (Auditory-visual spatial location matching task, *i.e.*, vertical bar localization, $M = 1$, $SD = 0.00$, $t(9) = \infty$, $p = 0.00$; Auditory pitch to visual elevation matching task, *i.e.*, horizontal bar localization, $M = 1$, $SD = 0.00$, $t(9) = \infty$, $p = 0.00$).

The sighted participants did not match the auditory stimuli to visual stimuli significantly more accurately than the Argus II patients (Auditory-visual spatial location matching task, $t(18) = -0.56$, $p = 0.58$; Auditory pitch to visual elevation matching task, $t(18) = -1.22$, $p = 0.24$) (Fig. 3B). Therefore, the crossmodal mappings measured in Argus II patients were similar in strength to age-matched sighted participants.

3.3. Crossmodal Mappings in Argus II and Sighted Individuals with Randomized Sound Order (Experiment 2)

Experiment 2 (Argus II patients, $N = 6$; sighted participants, $N = 10$) mirrored Experiment 1 except for the addition of further randomization of the multisensory stimuli. Experiment 2 had very similar results to Experiment 1, including significant multisensory stimuli matching in Argus II patients and sighted controls, and no significant difference between the groups. Experiment 2 is reported in detail in the Supplemental Information (Fig. S5).

3.4. Argus II Visual Search with Auditory Cueing (N = 6) (Experiment 3)

Argus II localization of the visual target was significantly above chance for both the auditory-visual task and the vision alone task (Auditory-visual task, $M = 0.92$, $SD = 0.10$, $t(5) = 14.61$, $p = 2.71 \times 10^{-5}$; Vision alone task, $M = 0.90$, $SD = 0.12$, $t(5) = 11.16$, $p = 1.01 \times 10^{-4}$). While the fraction correct was higher for the auditory-cued visual search task as compared to the fraction correct for the vision alone visual search task, they were not significantly different (Paired t-test, $t(5) = 0.59$, $p = 0.58$). However, the Argus II patients ($N = 6$) located the visual target significantly faster when auditory cues were present than when the visual stimulus was presented alone (Paired t-test, $t(5) = -3.03$, $p = 0.03$) (Fig. 4B). Therefore, crossmodal mappings for location can be used to increase visual search speed in

cluttered environments. Potential applications of this type of multimodal cueing for patients with artificial vision will be explored in the Discussion section.

Correlations were performed for the visual search performance relative to patient age, duration of prosthesis use, and the duration of blindness, and they were not found to be significantly correlated.

3.5. Argus II Timed Localization with Congruent and Incongruent Auditory Cueing (N = 6) (Experiment 4)

An additional localization experiment with both congruent and incongruent auditory-visual stimuli, and with no visual clutter (visual target presented alone) was performed by six Argus II patients (Experiment 4). The results of Experiment 4 showed on average a faster visual localization with congruent auditory-visual stimuli than incongruent stimuli, or visual stimuli alone (Fig. S6). Additional information on Experiment 4 is in the discussion section “The Role of Attention in Auditory-Cued Visual Search”.

4. Discussion

4.1. Overview

We have shown that Argus II patients can match features consistently and correctly across the senses by using two different crossmodal mappings between audition and vision. Furthermore, we verified that the crossmodal mappings in Argus II patients were similar to that observed in age-matched controls without vision loss by using a comparison to sighted participants performing the same tasks. Finally, we showed by means of a visual search task with co-presented auditory location cueing that these crossmodal mappings can possibly be used to increase visual search speed with ultra-low resolution retinal prostheses.

4.2. The Role of Attention in Auditory-Cued Visual Search

Attention and arousal can cause an increase in visual search speed when visual search with auditory cueing (congruent stimuli) is compared to visual search alone (as in Experiment 3). To address this concern with Experiment 3, Experiment 4 was performed in which both congruent and incongruent auditory cueing was tested in Argus II patients. Unfortunately, on average the Argus II patients were not significantly faster at the congruent auditory-visual condition relative to the incongruent auditory-visual condition. The smaller and more variable effect in Experiment 4 (congruent and incongruent timed localization) *vs.* Experiment 3 (congruent visual search), is likely due to an easier task (Experiment 4 did not have visual distractors whereas Experiment 3 did) and differences in patient search strategy (Experiment 4 did not instruct patients to use the auditory cues whereas Experiment 3 did). Nonetheless, the congruent auditory-visual localization was performed faster than the incongruent auditory-visual localization in Experiment 4. The individual Argus II patient results show that 4 out of 6 patients had a faster localization for the congruent relative to incongruent stimuli, and 2 out of 6 patients had a significantly faster localization for the congruent relative to incongruent stimuli. Therefore, two of the Argus II patients had faster localization with auditory cueing due to auditory-visual mappings rather than attentional improvements. High variability across the Argus II population is not unusual, in fact Garcia

et al showed integration of non-visual and visual self-motion cues in only 2 out of 4 Argus II patients tested (Garcia, Petrini, Rubin, Da Cruz & Nardini, 2015). Argus II patient variability in auditory-cued localization could be caused by differences in strategy, rehabilitation training, or crossmodal reorganization during blindness.

4.3. Visual Rehabilitation with Retinal Prostheses

With several visual restoration therapies in development, including stem cell therapies (Kashani, Lebkowski, Rahhal, Avery, Salehi-Had, Dang, Lin, Mitra, Zhu & Thomas, 2018, Roska & Sahel, 2018), gene therapies (Apte, 2018, Lam, Davis, Gregori, MacLaren, Girach, Verriotto, Rodriguez, Rosa, Zhang & Feuer, 2019), and artificial prosthetic devices, there is significant neuroscientific and medical interest in measuring the degree of crossmodal interactions that can be rehabilitated after a prolonged period of vision loss. While retinal prostheses are still in an early stage of development, the number of psychophysical assessments of subject testing and use with the device is rapidly increasing (Fernandes, Diniz, Ribeiro & Humayun, 2012, Humayun et al., 2012, Theogarajan, 2012, Zrenner, 2013). Argus II patients have been able to perform basic reach and grasp tasks, detect motion, distinguish common objects, and read letters and words (Castaldi, Cicchini, Cinelli, Biagi, Rizzo & Morrone, 2016, Da Cruz, Coley, Dorn, Merlini, Filley, Christopher, Chen, Wuyyuru, Sahel & Stanga, 2013, Dorn, Ahuja, Caspi, Da Cruz, Dagnelie, Sahel, Greenberg, McMahon & Group, 2013, Kotecha, Zhong, Stewart & Da Cruz, 2014, Luo et al., 2014). These initial tasks have shown both a large variability in patient functionality and slower task performance than predicted (Beyeler et al., 2019, Fornos, Sommerhalder, Da Cruz, Sahel, Mohand-Said, Hafezi & Pelizzone, 2012, Luo et al., 2016, Yanai, Weiland, Mahadevappa, Greenberg, Fine & Humayun, 2007). Limitations in task performance in all of these emerging visual restoration therapies could be aided by the use of crossmodal interactions to enhance learning with multimodal training (Seitz et al., 2006), and to improve visual search via crossmodal cueing.

4.4. Artificial Sense Integration with Natural Perception

The restoration of crossmodal interactions between artificial hearing and natural vision have been explored in cochlear implant research. It has been shown that cochlear implant users can integrate artificial audition with visual face stimuli in the McGurk Illusion (Rouger, Fraysse, Deguine & Barone, 2008, Schorr, Fox, van Wassenhove & Knudsen, 2005). In addition, implant users show normal crossmodal motion adaptation between artificial audition and natural vision (Fengler, Müller & Röder, 2018). It can be argued that the cochlear implant is in part successful due to neuroplasticity, which has allowed the integration of artificial sensation with the other remaining senses.

Like retinal prostheses, sensory substitution is a form of artificial vision-like sensation. However, in sensory substitution, visual information is translated into an auditory signal, and blind individuals are trained to interpret it. It has been shown that this auditory-visual “artificial” sensation can be integrated with natural tactile sensation by naive and trained individuals (Stiles & Shimojo, 2015). In addition, it was found that natural vision could be integrated with auditory-visual sensory substitution via a sensory-motor learning task (Levy-Tzedek, Novick, Arbel, Abboud, Maidenbaum, Vaadia & Amedi, 2012). Overall, the brain

has proven to be a plastic system, which seems to integrate new artificial information sources with existing senses to improve task performance.

Based in part on these results in artificial hearing and sensory substitution users, we evaluated crossmodal mappings in artificial vision users. These Argus II patients demonstrated that crossmodal interactions can also be rehabilitated with prosthetic vision. The crossmodal mappings in Argus II patients were not significantly different from sighted controls, although Argus II patients had a higher variability in crossmodal interaction strength. This result is consistent with the finding that crossmodal plasticity generated a high functional variability in the cochlear implant population (Buckley & Tobey, 2011, Lee, Lee, Oh, Kim, Kim, Chung, Lee & Kim, 2001, Sandmann, Dillier, Eichele, Meyer, Kegel, Pascual-Marqui, Marcar, Jäncke & Debener, 2012).

4.5. Retinal Prostheses and Multisensory Cueing

We have demonstrated that a visual target stimulus can be located in a cluttered visual display more quickly when a co-located auditory cue is simultaneously presented. This crossmodal cueing could be used to improve visual search with low-resolution artificial vision, for example, by using an automated object identification and auditory cueing algorithm. The object location information used in this type of cueing could be automatically estimated using several methods, including the automated visual processing of the video stream by machine learning, or the localization of pre-placed bar codes or beacons on objects or locations of interest (Liu et al., 2018). Previous research has explored the development and testing of automated navigational and object-identification devices and algorithms (Mante & Weiland, 2018, Parikh et al., 2013), in which the cued information is integrated in low vision and blind individuals with existing natural low vision or natural tactile perception (*e.g.*, cane navigation). Our results show that a cueing algorithm could provide benefits in visual search to ultra-low artificial vision patients, who have unique visual limitations and rehabilitation potential. In particular, the relearning of crossmodal interactions in these patients allows for the improvement of artificial vision by its integration with natural auditory perception.

4.6. Perceptual versus Cognitive Processing of Crossmodal Correspondences

The multisensory research community has extensively investigated crossmodal correspondences in the sighted, and has shown that correspondences are an integral part of how the senses interact, and cannot be reduced to purely linguistic correlations (Spence, 2011). The second crossmodal mapping that we tested, the mapping of auditory pitch to spatial elevation, has been shown through speeded classification tasks in numerous papers to be perceptually mediated (Ben-Artzi & Marks, 1995, Bernstein & Edelstein, 1971, Evans & Treisman, 2010, Melara & O'Brien, 1987, Miller, 1991, Patching & Quinlan, 2002). In a crossmodal mappings review paper, Charles Spence stated that “the crossmodal correspondence between auditory pitch and visual elevation constitutes one of the more robust associations to have been reported to date” (Spence, 2011). In addition, Miller *et al.* refined the testing of elevation and pitch mapping to even rule out decisional/response selection as a mechanism, further supporting the perceptual processing of this mapping (Miller, 1991). In this paper we show that Argus II patients had a significant mapping

between auditory pitch and visual elevation. Given these previous studies, it is likely that the crossmodal mapping for elevation and pitch operates at the same perceptual level in Argus II users.

We also tested the spatial mapping of visual and auditory stimuli in Argus II patients, and showed that these patients on average exhibited this spatial mapping as well. In addition, the evaluation of this crossmodal mapping was supported by a second experiment, which used a timed visual search (one type of test for perceptual processing), and found that on average patients had a significantly speeded cued response to the auditory-visual stimuli. In an even more rigorous test, which had both congruent and incongruent stimuli, 2 out of 6 patients still had a significantly speeded response to congruent auditory-visual location stimuli relative to incongruent stimuli. (Note: In this congruent and incongruent task, participants were told that sound cues may or may not be useful for locating the visual flash; therefore, they were not encouraged to use a cognitive strategy). Therefore, it was shown that in at least a subset of patients, spatial stimuli were likely processed at a non-cognitive level.

4.7. Early Visual Restoration in the Early Blind

Retinal prostheses restore vision in adulthood in the late blind, however vision loss can also be treated during the critical period in patients with blindness from birth from cataracts (Chen, Wu, Chen, Zhu, Li, Thorn, Ostrovsky & Qu, 2016, Gandhi, Kalia, Ganesh & Sinha, 2015, Guerreiro, Putzar & Röder, 2015, Held, Ostrovsky, de Gelder, Gandhi, Ganesh, Mathur & Sinha, 2011, Putzar, Goerendt, Lange, Rösler & Röder, 2007). The plasticity of the critical period in childhood permits these early blind patients to have the potential for a more complete visual recovery than if the same patients were restored vision at an older age. In comparison to Argus II patients, cataract patients can have quite different rehabilitation methods, challenges, and outcomes. In particular, cataract patients have vision loss from birth, they receive visual restoration by cataract surgery thereby restoring full-resolution natural visual perception, and their visual restoration is accomplished during the visual critical period. In contrast, Argus II patients are late blind with normal vision often into the patient's mid-twenties before the onset of vision loss. Argus II patients also have longer periods of vision loss (often spanning many decades) than cataract patients, and Argus II patient visual restoration is provided at low resolution with a biomedically implanted device (Humayun et al., 2012). Crossmodal interactions have been shown to be partially recovered in cataract patients. In particular, recovered congenital cataract patients have been found to be able to match basic tactile and visual shapes (Held et al., 2011), but could not match high frequency sounds with angular shapes (Sourav, Kekunnaya, Shareef, Banerjee, Bottari & Röder, 2019). Like the cataract patients, the Argus II patients were shown in this paper to recover some basic auditory-visual mappings. Additional research which uses neuroimaging could highlight additional similarities and differences between the visual and crossmodal neural processing of these two patient groups.

4.8. Crossmodal Interactions and Low Vision

Auditory-cued visual search has been previously studied in low vision patients to determine if multimodality could improve situational awareness, and the efficiency of visual search. In particular, patients with partial vision loss due to cortical damage (hemianopia) have been

shown to improve visual search and visual detection with auditory-visual stimuli (Cappagli et al., 2019, Frassinetti et al., 2005, Tinelli et al., 2015). Furthermore, training with auditory-cued visual search generated improvements in visual oculomotor exploration strategies, that transferred to daily life activities (Bolognini et al., 2005). In this paper, we showed that the visually restored (Argus II patients) have the ability to associate spatially coherent auditory and visual stimuli, and the ability to use auditory cueing to improve visual search. Low vision research in hemianopia highlights the benefits of longer term auditory-visual search training, which could further improve Argus II patient's visual perception. Therefore, extended training with auditory-visual search tasks is a promising possibility for future research on rehabilitation protocol development. (Note: A key difference between the low vision patients (with cortical or retinal vision loss) and Argus II patients, is that low vision patients have never had a period of complete blindness (or light perception alone). Therefore, low vision patients maintain multimodal interactions between vision and the other senses, despite a degradation in visual field of view and visual resolution. In contrast, Argus II patients (in addition to the very low resolution of their vision) are re-learning to see after decades of vision loss. Therefore, while low-vision research can be quite informative for rehabilitation training techniques and as a comparison to Argus II patients, in many ways Argus II patients are an entirely unique group, with many more challenges for the crossmodal interactions than patients with low vision.)

5. Conclusion

Our results suggest that restoration of even limited visual perception in the late blind can interact with audition through crossmodal mappings. This implies that remodeling of visual regions during blindness and the challenges of artificial vision do not prohibit the rehabilitation of these types of auditory-visual interactions. Finally, we showed that visual search can be accelerated by the simultaneous presentation of auditory location cues in the visually restored, which supports the potential use of crossmodal cueing as a method to improve prosthetic visual functionality.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful for financial support from the USC Roski Eye Institute, the National Institutes of Health (Grant Number U01 EY025864); the Philanthropic Educational Organization Scholar Award Program; and the Arnold O. Beckman Postdoctoral Scholars Fellowship Program. We are thankful to Professor Mark S. Humayun for providing laboratory space in the Ginsburg Institute for Biomedical Therapeutics to perform these experiments.

References

- (2019). Second Sight, Our Therapy, FAQ. 2019 (Second Sight Medical Products Website: <https://www.secondsight.com/faq/>).
- Apte RS (2018). Gene therapy for retinal degeneration. *Cell*, 173 (1), 5. [PubMed: 29570997]
- Ayton LN, Barnes N, Dagnelie G, Fujikado T, Goetz G, Hornig R, Jones BW, Muqit MM, Rathbun DL, & Stingl K (2020). An update on retinal prostheses. *Clinical Neurophysiology*, 131 (6), 1383–1398. [PubMed: 31866339]

- Ben-Artzi E, & Marks LE (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics*, 57 (8), 1151–1162. [PubMed: 8539090]
- Bernstein IH, & Edelman BA (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, 87 (2), 241. [PubMed: 5542226]
- Beyeler M, Nanduri D, Weiland JD, Rokem A, Boynton GM, & Fine I (2019). A model of ganglion axon pathways accounts for percepts elicited by retinal implants. *Scientific Reports*, 9 (1), 9199. [PubMed: 31235711]
- Bolognini N, Rasi F, Coccia M, & Ladavas E (2005). Visual search improvement in hemianopic patients after audio-visual stimulation. *Brain*, 128 (12), 2830–2842. [PubMed: 16219672]
- Buckley KA, & Tobey EA (2011). Cross-modal plasticity and speech perception in pre- and postlingually deaf cochlear implant users. *Ear and Hearing*, 32 (1), 2–15. [PubMed: 20829699]
- Cappagli G, Finocchietti S, Cocchi E, Giammari G, Zumiani R, Cuppone AV, Baud-Bovy G, & Gori M (2019). Audio motor training improves mobility and spatial cognition in visually impaired children. *Scientific Reports*, 9 (1), 1–9. [PubMed: 30626917]
- Castaldi E, Cicchini GM, Cinelli L, Biagi L, Rizzo S, & Morrone MC (2016). Visual BOLD response in late blind subjects with Argus II retinal prosthesis. *PLoS Biology*, 14 (10).
- Castaldi E, Lunghi C, & Morrone MC (2020). Neuroplasticity in adult human visual cortex. *Neuroscience & Biobehavioral Reviews*, 112, 542–552. [PubMed: 32092315]
- Chen J, Wu E-D, Chen X, Zhu L-H, Li X, Thorn F, Ostrovsky Y, & Qu J (2016). Rapid integration of tactile and visual information by a newly sighted child. *Current Biology*, 26 (8), 1069–1074. [PubMed: 27040777]
- Cunningham SI, Shi Y, Weiland JD, Falabella P, de Koo LCO, Zacks DN, & Tjan BS (2015). Feasibility of structural and functional MRI acquisition with unpowered implants in Argus II retinal prosthesis patients: A case study. *Translational Vision Science & Technology*, 4 (6), 6–6.
- Da Cruz L, Coley BF, Dorn J, Merlini F, Filley E, Christopher P, Chen FK, Wuyyuru V, Sahel J, & Stanga P (2013). The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss. *British Journal of Ophthalmology*, 97 (5), 632–636.
- Da Cruz L, Dorn JD, Humayun MS, Dagnelie G, Handa J, Barale P-O, Sahel J-A, Stanga PE, Hafezi F, & Safran AB (2016). Five-year safety and performance results from the Argus II retinal prosthesis system clinical trial. *Ophthalmology*, 123 (10), 2248–2254. [PubMed: 27453256]
- Dagnelie G (2017). Patient-Reported Outcomes (PRO) for Prosthetic Vision. In: *Artificial Vision* (pp. 21–27): Springer.
- Dagnelie G (2018). Retinal Prostheses: Functional Outcomes and Visual Rehabilitation. In: *Retinal Prosthesis* (pp. 91–104): Springer.
- Deroy O, & Spence C (2016). Crossmodal correspondences: Four challenges. *Multisensory Research*, 29 (1–3), 29–48. [PubMed: 27311290]
- Dorn JD, Ahuja AK, Caspi A, Da Cruz L, Dagnelie G, Sahel J-A, Greenberg RJ, McMahon MJ, & Group AIS (2013). The detection of motion by blind subjects with the epiretinal 60-electrode (Argus II) retinal prosthesis. *JAMA Ophthalmology*, 131 (2), 183–189. [PubMed: 23544203]
- Erickson-Davis C, & Korzybska H (2020). What do blind people “see” with retinal prostheses? Observations and qualitative reports of epiretinal implant users. *BioRxiv*.
- Evans K, & Treisman A (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10 (1).
- Fengler I, Müller J, & Röder B (2018). Auditory and visual motion adaptation in congenitally deaf cochlear implant users: Evidence for typical crossmodal aftereffects. *Journal of Hearing Science*, 8 (2).
- Fernandes RAB, Diniz B, Ribeiro R, & Humayun M (2012). Artificial vision through neuronal stimulation. *Neuroscience Letters*, 519 (2), 122–128. [PubMed: 22342306]
- Fornos AP, Sommerhalder J, Da Cruz L, Sahel JA, Mohand-Said S, Hafezi F, & Pelizzone M (2012). Temporal properties of visual perception on electrical stimulation of the retina. *Investigative Ophthalmology & Visual Science*, 53 (6), 2720–2731. [PubMed: 22447863]

- Frassinetti F, Bolognini N, Bottari D, Bonora A, & Làdavas E (2005). Audiovisual integration in patients with visual deficit. *Journal of Cognitive Neuroscience*, 17 (9), 1442–1452. [PubMed: 16197697]
- Gandhi T, Kalia A, Ganesh S, & Sinha P (2015). Immediate susceptibility to visual illusions after sight onset. *Current Biology*, 25 (9), R358–R359. [PubMed: 25942545]
- Garcia S, Petrini K, Rubin GS, Da Cruz L, & Nardini M (2015). Visual and non-visual navigation in blind patients with a retinal prosthesis. *PLoS One*, 10 (7).
- Guerreiro MJ, Putzar L, & Röder B (2015). The effect of early visual deprivation on the neural bases of multisensory processing. *Brain*, 138 (6), 1499–1504. [PubMed: 25808371]
- Held R, Ostrovsky Y, de Gelder B, Gandhi T, Ganesh S, Mathur U, & Sinha P (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, 14 (5), 551. [PubMed: 21478887]
- Humayun MS, Dorn JD, Da Cruz L, Dagnelie G, Sahel J-A, Stanga PE, Cideciyan AV, Duncan JL, Elliott D, & Filley E (2012). Interim results from the international trial of Second Sight's visual prosthesis. *Ophthalmology*, 119 (4), 779–788. [PubMed: 22244176]
- Kashani AH, Lebkowski JS, Rahhal FM, Avery RL, Salehi-Had H, Dang W, Lin C-M, Mitra D, Zhu D, & Thomas BB (2018). A bioengineered retinal pigment epithelial monolayer for advanced, dry age-related macular degeneration. *Science Translational Medicine*, 10 (435), eaao4097. [PubMed: 29618560]
- Klapetek A, Ngo MK, & Spence C (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, 74 (6), 1154–1167.
- Kotecha A, Zhong J, Stewart D, & Da Cruz L (2014). The Argus II prosthesis facilitates reaching and grasping tasks: a case series. *BMC Ophthalmology*, 14 (1), 71. [PubMed: 24885164]
- Lam BL, Davis JL, Gregori NZ, MacLaren RE, Girach A, Verriotto JD, Rodriguez B, Rosa PR, Zhang X, & Feuer WJ (2019). Choroideremia gene therapy phase 2 clinical trial: 24-month results. *American Journal of Ophthalmology*, 197, 65–73. [PubMed: 30240725]
- Lee DS, Lee JS, Oh SH, Kim S-K, Kim J-W, Chung J-K, Lee MC, & Kim CS (2001). Deafness: cross-modal plasticity and cochlear implants. *Nature*, 409 (6817), 149.
- Levy-Tzedek S, Novick I, Arbel R, Abboud S, Maidenbaum S, Vaadia E, & Amedi A (2012). Cross-sensory transfer of sensory-motor information: Visuomotor learning affects performance on an audiomotor task, using sensory-substitution. *Scientific Reports*, 2, 949. [PubMed: 23230514]
- Liu Y, Stiles NR, & Meister M (2018). Augmented reality powers a cognitive assistant for the blind. *eLife*, 7, e37841. [PubMed: 30479270]
- Luo YH, Zhong JJ, Clemo M, & Da Cruz L (2016). Long-term repeatability and reproducibility of phosphene characteristics in chronically implanted Argus II retinal prosthesis subjects. *American Journal of Ophthalmology*, 170, 100–109. [PubMed: 27491695]
- Luo YH-L, & Da Cruz L (2016). The Argus® II retinal prosthesis system. *Progress in Retinal and Eye Research*, 50, 89–107. [PubMed: 26404104]
- Luo YH-L, Zhong J, Merlini F, Anafloos F, Arsiero M, Stanga PE, & Da Cruz L (2014). The use of Argus® II retinal prosthesis to identify common objects in blind subjects with outer retinal dystrophies. *Investigative Ophthalmology & Visual Science*, 55 (13), 1834–1834.
- Mante N, & Weiland JD (2018). Visually impaired users can locate and grasp objects under the guidance of computer vision and non-visual feedback. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1–4): IEEE.
- McDonald RB (2019). My Brain Implant for Bionic Vision: The First Trial of Artificial Sight for the Blind. (p. 129): Independently published.
- Melara RD, & O'Brien TP (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, 116 (4), 323.
- Miller J (1991). Channel interaction and the redundant-targets effect in bimodal divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, 17 (1), 160. [PubMed: 1826309]
- Mowad TG, Willett AE, Mahmoudian M, Lipin M, Heinecke A, Maguire AM, Bennett J, & Ashtari M (2020). Compensatory Cross-Modal Plasticity Persists After Sight Restoration. *Frontiers in Neuroscience*, 14, 291. [PubMed: 32477041]

- Mudd S (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology*, 66 (4), 347. [PubMed: 14051851]
- Parikh N, Itti L, Humayun M, & Weiland J (2013). Performance of visually guided tasks using simulated prosthetic vision and saliency-based cues. *Journal of Neural Engineering*, 10 (2), 026017. [PubMed: 23449023]
- Parise CV, Knorre K, & Ernst MO (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, 111 (16), 6104–6108.
- Patching GR, & Quinlan PT (2002). Garner and congruence effects in the speeded classification of bimodal signals. *Journal of Experimental Psychology: Human Perception and Performance*, 28 (4), 755. [PubMed: 12190249]
- Pratt CC (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, 13 (3), 278.
- Putzar L, Goerendt I, Lange K, Rösler F, & Röder B (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience*, 10 (10), 1243–1245. [PubMed: 17873871]
- Roffler SK, & Butler RA (1968). Factors that influence the localization of sound in the vertical plane. *The Journal of the Acoustical Society of America*, 43 (6), 1255–1259. [PubMed: 5659493]
- Roska B, & Sahel J-A (2018). Restoring vision. *Nature*, 557 (7705), 359. [PubMed: 29769667]
- Rouger J, Fraysse B, Deguine O, & Barone P (2008). McGurk effects in cochlear-implanted deaf subjects. *Brain Research*, 1188, 87–99. [PubMed: 18062941]
- Sandmann P, Dillier N, Eichele T, Meyer M, Kegel A, Pascual-Marqui RD, Marcar VL, Jäncke L, & Debener S (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain*, 135 (2), 555–568. [PubMed: 22232592]
- Schorr EA, Fox NA, van Wassenhove V, & Knudsen EI (2005). Auditory-visual fusion in speech perception in children with cochlear implants. *Proceedings of the National Academy of Sciences*, 102 (51), 18748–18750.
- Seitz AR, Kim R, & Shams L (2006). Sound facilitates visual learning. *Current Biology*, 16 (14), 1422–1427. [PubMed: 16860741]
- Sourav S, Kekunnaya R, Shareef I, Banerjee S, Bottari D, & Röder B (2019). A protracted sensitive period regulates the development of cross-modal sound–shape associations in humans. *Psychological Science*, 30 (10), 1473–1482. [PubMed: 31483197]
- Spence C (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73 (4), 971–995.
- Stiles NR, & Shimojo S (2015). Auditory sensory substitution is intuitive and automatic with texture stimuli. *Scientific Reports*, 5, 15628. [PubMed: 26490260]
- Stiles NRB, McIntosh BP, Nasiatka PJ, Hauer MC, Weiland JD, Humayun MS, & Tanguay AR Jr (2011). An intraocular camera for retinal prostheses: Restoring sight to the blind. In: Serpenguzel A, & Poon AW (Eds.), *Optical Processes in Microparticles and Nanostructures: A Festschrift Dedicated to Richard Kounai Chang on His Retirement from Yale University* (p. 385): World Scientific Publishing Co.
- Stingl K, Bartz-Schmidt KU, Besch D, Chee CK, Cottrill CL, Gekeler F, Groppe M, Jackson TL, MacLaren RE, & Koitschev A (2015). Subretinal visual implant alpha IMS–clinical trial interim report. *Vision Research*, 111, 149–160. [PubMed: 25812924]
- Stronks HC, & Dagnelie G (2014). The functional performance of the Argus II retinal prosthesis. *Expert Review of Medical Devices*, 11 (1), 23–30. [PubMed: 24308734]
- Theogarajan L (2012). Strategies for restoring vision to the blind: Current and emerging technologies. *Neuroscience Letters*, 519 (2), 129–133. [PubMed: 22414860]
- Tinelli F, Purpura G, & Cioni G (2015). Audio-visual stimulation improves visual search abilities in hemianopia due to childhood acquired brain lesions. *Multisensory Research*, 28 (1–2), 153–171. [PubMed: 26152056]
- Weiland JD, Cho AK, & Humayun MS (2011). Retinal prostheses: current clinical results and future needs. *Ophthalmology*, 118 (11), 2227–2237. [PubMed: 22047893]

- Yanai D, Weiland JD, Mahadevappa M, Greenberg RJ, Fine I, & Humayun MS (2007). Visual performance using a retinal prosthesis in three subjects with retinitis pigmentosa. *American Journal of Ophthalmology*, 143 (5), 820–827. e822. [PubMed: 17362868]
- Zhou DD, Dorn JD, & Greenberg RJ (2013). The Argus® II retinal prosthesis system: An overview. *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on* (pp. 1–6): IEEE.
- Zrenner E (2013). Fighting blindness with microelectronics. *Science Translational Medicine*, 5 (210), 210ps216–210ps216.

Highlights

- Argus II retinal prostheses restore low resolution vision to the blind
- Argus II patients exhibit crossmodal mappings between vision and audition
- Auditory cueing of a visual stimulus speeds up Argus II patient visual search
- Auditory-visual interactions could be leveraged to improve patient functionality



Fig. 1. Argus II Retinal Prosthesis System Components

The external components of the Argus II Retinal Prosthesis System. The visual information is recorded by the glasses mounted camera and then fed to a Visual Processing Unit (VPU), which translates visual information into electrical stimulation parameters. The visual signal is then sent via wire to the coil on the glasses, which wirelessly communicates through an implanted coil to a microelectrode stimulator array that is proximity coupled to the retina. Additional details on the Argus II device are detailed in the Methods section.

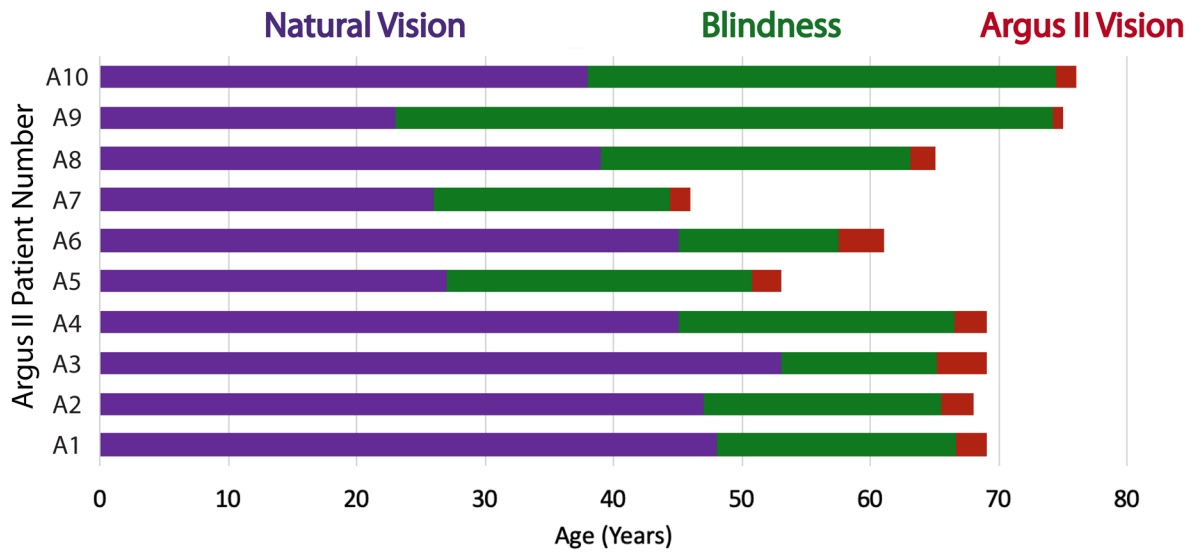


Fig. 2. Argus II Patient Timelines

Timeline for natural visual perception, vision loss, and vision restoration with the Argus II retinal prosthesis in each of the patients tested. See also Table 1.

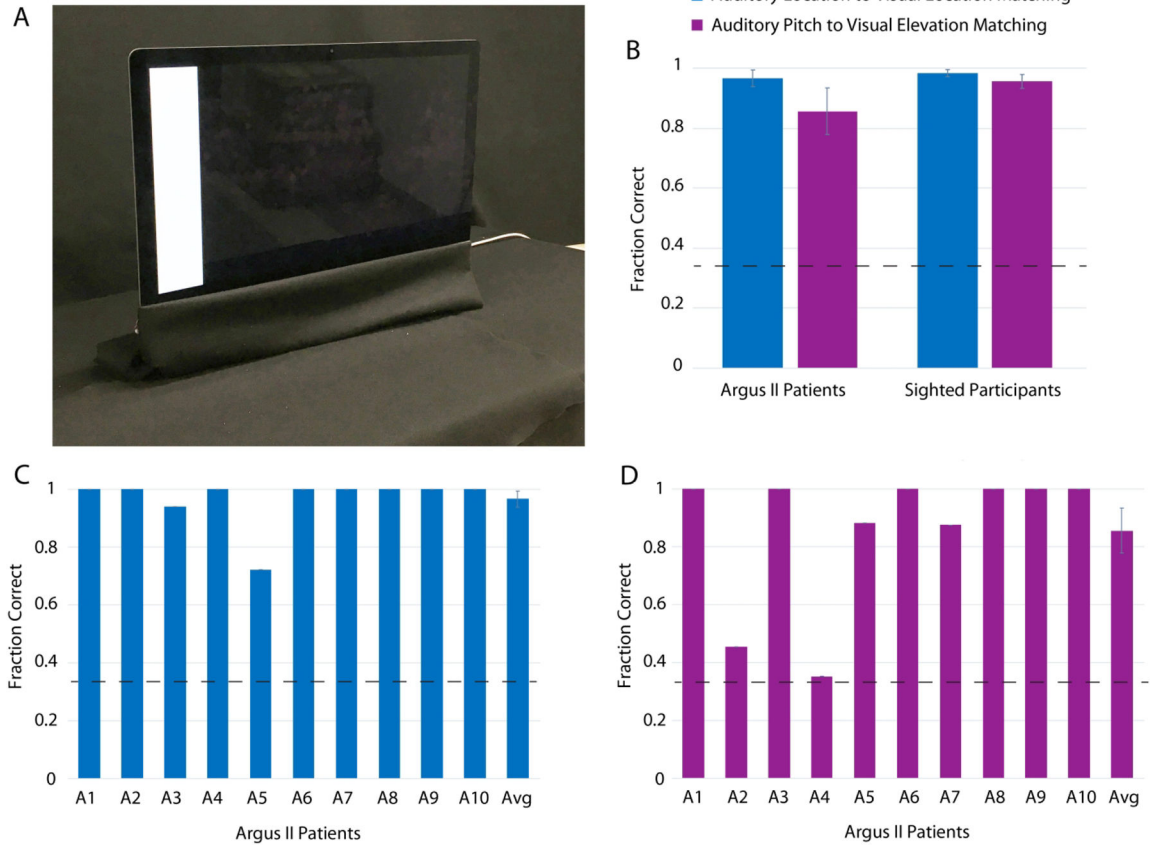


Fig. 3. Crossmodal Correspondences Results in Argus II and Sighted Subjects (Experiment 1)

The experimental setup at the University of Southern California for both crossmodal correspondence experiments is shown in Panel A. The fraction correct for matching auditory to visual stimuli is shown in Panel B for both Argus II patients ($N=10$) and age-matched sighted controls ($N=10$). The fraction correct for matching auditory to visual stimuli in Argus II patients ($N=10$) is shown in Panels C and D (Panel C for the mapping of auditory location to visual location, and Panel D for the mapping of auditory pitch to visual elevation). The individual patient fraction correct results for each Argus II patient in Panels C and D do not have error bars as they represent one data point for each patient; however, the averages in Panels C and D do have error bars as they sum across patients. The error bars represent \pm standard error of the mean across participants. The dashed lines represent chance. See also Figs. S4 and S5, and Table S2.

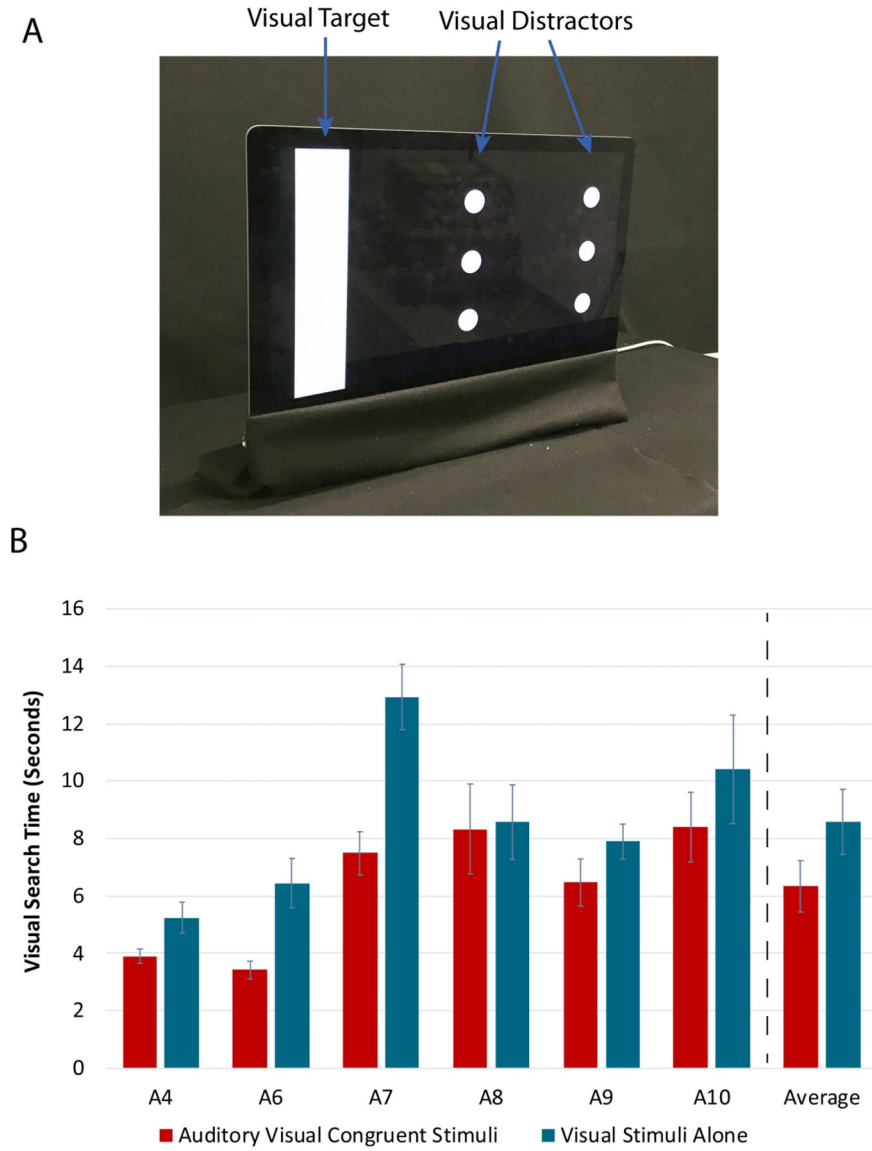


Fig. 4. Visual Search Task Setup and Results with Argus II Patients (Experiment 3)

A representative image of the visual stimuli in the visual search task is shown in Panel A with the visual target and visual distractors labeled. The time to detection results for the visual search task in Argus II patients ($N=6$) is shown in Panel B. The full length of the error bars represent \pm standard error of the mean across trials for individual results, and across participants for average results. See also Fig. S6.

Table 1.

Argus II Patient Information

Subject ID	Age	Gender	Duration Blind (Years)	Duration with Argus II (Months)	Vision Left	Vision Right	Eye Patch Left	Eye Patch Right
A1	69	M	21	28	LP	LP	Yes	Yes
A2	68	F	21	29.5	LP	LP	Yes	No
A3	69	F	16	45.5	No LP	LP	No	Yes
A4	69	M	24	30	LP	LP	Yes	Yes
A5	53	F	26	27.5	LP	No LP	Yes	Yes
A6	61	F	16	42	LP	No LP	Yes	No
A7	46	M	20	19.5	LP	LP	Yes	Yes
A8	65	M	26	23	LP	No LP	Yes	Yes
A9	75	F	52	10	No LP	No LP	No	No
A10	76	M	38	19	No LP	No LP	Yes	Yes

M = Male, F = Female, LP = Light Perception, and RP = Retinitis Pigmentosa

Table 2.

Sighted Control Participant Information

Subject ID	Age	Gender	Glasses or Contacts?	Hearing Loss Left Ear	Hearing Loss Right Ear
S1	61	M	Yes	None	None
S2	66	F	Yes	None	None
S3	68	F	Yes	None	None
S4	69	F	Yes	None	None
S5	64	F	No	Mod	Mod
S6	58	F	Yes	None	None
S7	63	F	Yes	None	None
S8	69	F	Yes	None	None
S9	55	M	Yes	None	None
S10	62	M	No	None	None

M = Male, F = Female, None = No Hearing Loss, Mild = Mild Hearing Loss, Mod = Moderate Hearing Loss, and Sev = Severe Hearing Loss

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript