Taylor & Francis
Taylor & Francis Group

Disability and Rehabilitation: Assistive Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iidt20

Preparing participants for the use of the tongue visual sensory substitution device

Zahide Pamir, Jae-Hyun Jung & Eli Peli

To cite this article: Zahide Pamir, Jae-Hyun Jung & Eli Peli (2022) Preparing participants for the use of the tongue visual sensory substitution device, Disability and Rehabilitation: Assistive Technology, 17:8, 888-896, DOI: 10.1080/17483107.2020.1821102

To link to this article: https://doi.org/10.1080/17483107.2020.1821102

Published online: 30 Sep 2020.

DISABILITY AND REHABILITATION: ASSISTIVE TECHNOLOGY
2022, VOL. 17, NO. 8, 888–896
https://doi.org/10.1080/17483107.2020.1821102

ORIGINAL RESEARCH

Preparing participants for the use of the tongue visual sensory substitution device

Zahide Pamir, Jae-Hyun Jung and Eli Peli

Department of Ophthalmology, Schepens Eye Research Institute of Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

ABSTRACT

Purpose: Visual sensory substitution devices (SSDs) convey visual information to a blind person through another sensory modality. Using a visual SSD in various daily activities requires training prior to use the device independently. Yet, there is limited literature about procedures and outcomes of the training conducted for preparing users for practical use of SSDs in daily activities.

Methods: We trained 29 blind adults (9 with congenital and 20 with acquired blindness) in the use of a commercially available electro-tactile SSD, BrainPort. We describe a structured training protocol adapted from the previous studies, responses of participants, and we present retrospective qualitative data on the progress of participants during the training.

Results: The length of the training was not a critical factor in reaching an advanced stage. Though performance in the first two sessions seems to be a good indicator of participants' ability to progress in the training protocol, there are large individual differences in how far and how fast each participant can progress in the training protocol. There are differences between congenital blind users and those blinded later in life.

Conclusions: The information on the training progression would be of interest to researchers preparing studies, and to eye care professionals, who may advise patients to use SSDs.

ARTICLE HISTORY
Received 13 July 2020
Accepted 4 September 2020

KEYWORDS
BrainPort; training; sensory substitution device; visual-to-tactile; tongue

➤ **IMPLICATIONS FOR REHABILITATION**

- There are large individual differences in how far and how fast each participant can learn to use a visual-to-tactile sensory substitution device for a variety of tasks.
- Recognition is mainly achieved through top-down processing with prior knowledge about the possible responses. Therefore, the generalizability is still questionable.
- Users develop different strategies in order to succeed in training tasks.

Introduction

Visual sensory substitution devices (SSDs) provide a non-invasive assistive technology for the blind/visually impaired. General-purpose SSDs aim at improving blind people's functional performance in a variety of everyday tasks such as light detection, object recognition and mobility, and consequently enhancing independence of blind individuals [1,2].

SSDs convey visual information through other sensory modalities such as touch or audition. Understanding the information normally processed by one modality when presented in another may not be an intuitive and automatic task [3], but rather a learned skill like reading or language that requires training and practice [4,5] (but see Stiles and Shimojo [6] for results indicating that perception is intuitive, to some extent, with an auditory sensory substitution).

One example of a general-purpose SSD that requires extensive training prior to practical use of the device for daily activities is the BrainPort (Wicab Inc., Madison, WI), a commercially available SSD that conveys video camera-based visual information through electro-tactile stimulation on the tongue. The Wicab's website [7] (and the user manual) states that a typical 10 hours (h) of supervised training is necessary before using the BrainPort independently. Studies conducted to evaluate the functional performance of the BrainPort also pointed out the necessity of training by a professional and following the training home self-practice by users in activities of daily living; and they described a structured training protocol which is carried out over a few 3-h sessions and usually lasts for a total of 10–15 h [e.g. 3,5,8]. The training protocol starts with familiarising participants with the device, its components, purpose, and limitations, how it converts visual information to tactile stimulation on the tongue, and how to interpret the stimulation. The protocol involves several tasks with increasing levels of complexity, such as understanding the spatial relations between the stimulation on the tongue and surrounding visual environment, recognition of high-contrast basic shapes, letter recognition, and recognition of common wall signs such as a restroom sign.

Although participants were trained by following the structured training protocol in the aforementioned studies [3,5,8], details related to their individual performance during the training have not been reported. For example, Grant et al. [5] reported that 57 participants who completed their study were trained by an experienced BrainPort trainer prior to functional testing with the device. However, they did not report whether all participants progress

CONTACT Zahide Pamir zahide_pamir@meei.harvard.edu Department of Ophthalmology, Schepens Eye Research Institute of Massachusetts Eye and Ear, Harvard Medical School, Boston, MA, USA

Table 1. Demographic characteristics of participants.

Aetiology	Congenital Number of participants	Acquired Number of participants	Acquired Onset before training (years)
Retinopathy of Prematurity	4	1	37
Diabetic Retinopathy		4	10, 13, 23, 39
Retinitis Pigmentosa	1	3	20, 34, 48
Leber's Optic Neuropathy		2	11, 12
Glaucoma		2	33, 43
Trauma		2	2, 13
Leber's Congenital Amaurosis	1		–
Microphthalmia	1		–
Uveitis		1	45
Detached Retina and Band Keratopathy		1	53
Cone Dystrophy		1	15
Immune Deficiency		1	4
Retinal Blastoma		1	57
Optic Nerve Damage		1	6
Not Known	2		–
Total	**9**	**20**	

similarly during the training or performed at an expected success level in all the steps included in the training protocol. Similarly, 100 participants were trained by Nau et al. [3], but functional testing was carried out only with 18 participants. They did not explain why more than 80% of participants that they had trained were not included in the experiments that followed. As a result, the following questions usually remained unanswered in these papers [3,5,8]: Are participants expected to reach a certain expertise level in order to be considered to have completed the training's objective of preparing users for daily activities, or is the completion of the training defined just by time (e.g. 10 h)? How does performance vary among participants? Does it take a similar time for each participant to complete a particular task in the training? Are there any participants who were not able to complete all training tasks? Are there any predictive participant's characteristics for better or poorer training outcome? Such details about the performance and the progress during the structured training applied for rehabilitation purposes might be useful for evaluating the potential of the adoption of SSDs by users.

Some studies trained participants only on a single targeted task in order to evaluate the performance in that particular task, e.g. basic shape discrimination [9] or motion direction discrimination [10]. In these studies, training is usually tailored for the requirements of the study, and the SSD is used not as a visual rehabilitation aid but a research tool to answer specific questions such as the spatial or temporal resolution of the tactile system [11,12]. In such studies, the progression in performance is usually reported during the training, and performance is usually expected to asymptote before the final evaluation is conducted. This type of training is outside the scope of the current paper.

Over 2 years, we trained 29 blind participants (20 acquired and 9 congenital blind) following a structured training protocol. We adapted the protocol described by Nau et al. [3], with modifications, as described in the method section. Participants who completed the training participated in our studies where they were trained again especially for the targeted task tested in the study [e.g. 13]. Our aim was not to systematically investigating the performance in the training protocol [3,5,8], but merely to prepare our participants for the use of the device in various studies. However, during the training, we observed some strategies intuitively developed by participants, and that some participants experienced difficulties in some training tasks. To the best of our knowledge, in addition to the unanswered questions mentioned above, such observations on the training procedure have not been reported in the literature before. In order to fill this gap, we retrospectively analysed our training records and present here a qualitative summary of participants' performance and reactions. We also provide a detailed description of the training protocol. We followed a structured training protocol with pre-defined stages. Yet, since the training was not the aim of the study, the exposure of each participant to certain tasks differed based on the progress and interest of the participant, and the acceptable level of performance at each task was decided by the trainer subjectively, matching what might happen in clinical practice. Therefore, no statistical comparisons were performed, and this paper should be considered as a retrospective observational report. Information in this paper should be of value to researchers planning future studies, to eye care professionals who may advise patients regarding the use of this device, and to clinicians who may train people in the use of the device for daily activities.

Method

Participants

Thirty functionally blind adults (10 with congenital and 20 with acquired blindness) with the visual ability of light perception, hand motion or worse bilaterally, from a variety of aetiologies (see Table 1) were recruited. One congenitally blind participant was excluded without any training because of damage to the cortical lobe and inability to differentiate his right from his left, which is important to be able to use the BrainPort effectively. Exclusion criteria included cortical blindness from any cause, oral lesions or piercings, perceptual abnormalities in the tongue or skin based on personal report of participants, and low scores (below 23 out of 27) on the Mini Mental State Examination [14]. Participants (21 men) ranged in age from 31 to 86 ($M = 60$, SD $=$ 12) years.

Demographic information, cause of blindness and onset, and any abnormalities in the tactile system or any cortical damage were recorded based on personal reports. Training protocol adhered to the tenets of the Declaration of Helsinki, and the protocol and informed consent were approved by the Massachusetts Eye and Ear Institutional Review Board.

Apparatus and stimuli

The BrainPort

The BrainPort is an electro-tactile tongue SSD. It delivers the visual information acquired through a head-mounted camera to a

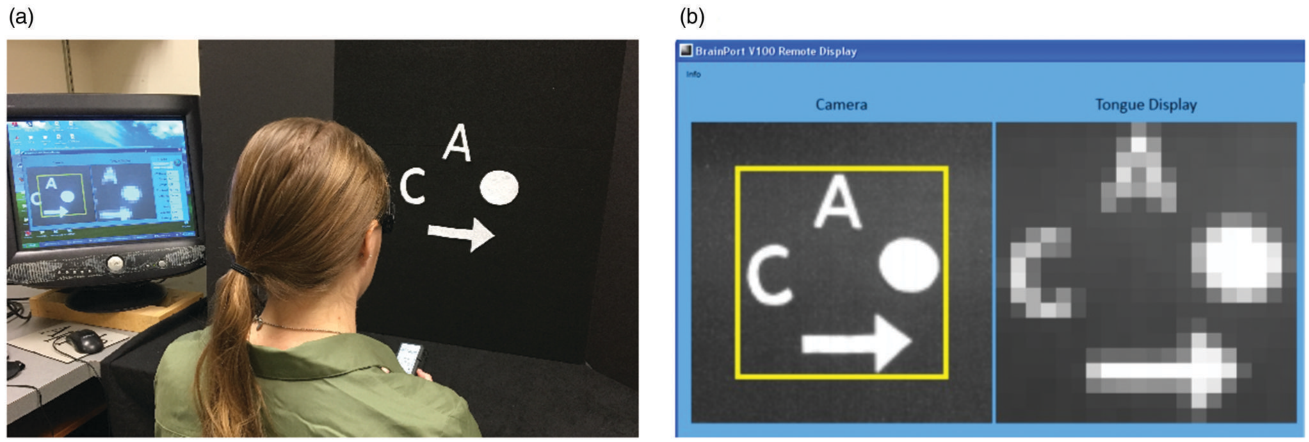

Figure 1. BrainPort testing environment. (a) Black trifold cardboard covered centrally in black felt standing on a black-clothed table. Sample stimuli cut out of white felt (letters A and C, a circle, and an arrow) are shown on the black surface to be viewed with BrainPort. The monitoring screen is seen to the left. (b) The BrainPort monitoring screen. The left window shows the camera field of view together with the yellow cropping box indicating the visual field of the tongue display (default 24° × 24°, controlled by the zoom). The right window illustrates the electrical stimulation sent to the IOD at the resolution of 20 × 20 pixels. Note that the default field of view presented on the IOD is much smaller than the felt-covered trifold (about 10% of the area).

20 × 20 grid of electrodes on a 25.8 × 25.8 mm intra-oral device (IOD) placed on the tongue (for more information on the device see Nau et al. [3]). Three versions of the device, V100, V200, and Vision Pro were used during the training. The dimension and resolution of IODs are similar in all three versions. In BrainPort V100, the camera is located on sunglasses and there is a separate handheld controller unit, whereas in the other two versions both the camera and controller are placed on a headset. For the purpose of the training, these two devices are identical. The controller mainly enables the user to turn on and off the device, adjust stimulation intensity and field of view (zoom), and select different viewing modes (e.g. inverting light and dark options). In the default configuration, visual stimuli are captured by the camera and then sent to the IOD, which can be monitored by sighted observers on a computer display through a Wi-Fi interface.

Training environment

Stimuli were presented on the centre board (91 × 60 cm) of a trifold cardboard and was covered in a black felt cloth. The trifold board was placed vertically open on a table that was covered in black, non-reflective fabric. This setup allowed for an all-black, non-cluttered background area for object viewing. Stimuli were cut from a white felt cloth that could be flexibly attached to the felt-covered trifold (Figure 1).

Procedure

We adapted the first four levels of the training protocol described in Nau et al. [3] and created a nine-stage training protocol by adding new tasks (see Stages 2, 3, 4, and 6 below).

The pace through each training session was adjusted by the trainer based on individual participant's progress and interest. Some participants wished to spend more time on some tasks although they were already performing satisfactorily at that stage. Others got bored and wanted to move on to other tasks although they were not yet successful at given stage. In the latter case, the trainer returned to the prior incomplete task after spending some time on the new task.

Each session lasted 2–3 h depending on each participant's fatigue and interest. Trainer subjectively decided on the acceptable level of performance at each stage required to move forward. The purpose of the training was to advance participants' usage of the device. The trainer prioritised keeping participants motivated and if the trainer noted that the participant was getting bored with one task, the trainer moved to the next stage after successfully completing fewer trials than required and completed by other participants. Although some participants completed all the stages in the training protocol, they were asked to participate in a few more sessions to determine if the performance will improve further, such as recognising letters or objects in a shorter time with more practice on already completed tasks.

The nine stages in the training protocol are described below.

Device familiarization and mapping of the tongue stimulation to the visual/spatial world

The participants were introduced to the operation of the device and were taught to use the controller; turning the device on and off and adjusting the intensity level. A white circle or square target (approximately 8° × 8°) was placed on the board centrally so that it would likely appear within the centre of IOD field of view (24° × 24°). Participants were asked to put the IOD onto their tongue and adjust the intensity to a comfortable level.

Next, while viewing the monitor, the trainer presented a handheld white Styrofoam bar (about 2.5 × 2.5 × 30 cm, approximately 2° × 2°× 24° but varies slightly depending on where the trainer presented the bar) to different edges of the field of view of the IOD (Yellow frames in Figure 1) stimulating one edge at a time. The trainer explained which side of the IOD/tongue stimulation corresponds to what part of the visual field (i.e. right to right, left to left, an object at the upper visual field stimulates the back of the tongue, and an object on the lower visual field stimulates the tip of the tongue). When participants could correctly locate the stimulation on the tongue, they were asked to hold the Styrofoam bar in their hand and practice stimulation of the four edges of the IOD visual field on their own. The main purpose of this task was to introduce participants to the IOD "field of view" and its relation to the real-world directions.

Static bar orientation

The trainer attached a white bar (approximately 2° × 30°) to the black vertical board at one of four different orientations (horizontal, vertical, left- or right-tilted) and participants were required to report the orientation of the bar. Tracing strategies with the head or tip of the tongue helped participants having difficulty in

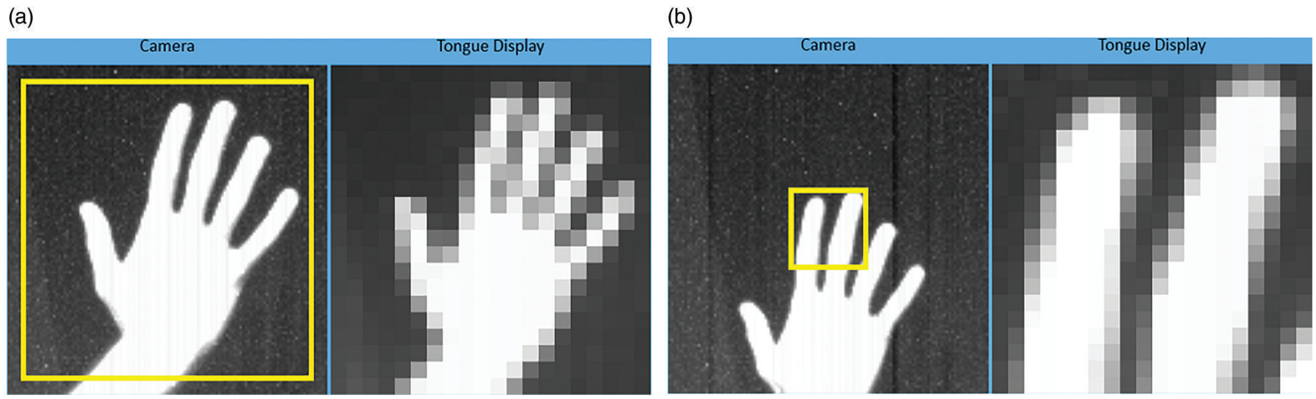

Figure 2. The use of the zoom function. On the left the camera view, with the yellow frame marking the portion of the image sent to the IOD. On the right is a simulation of the signal sent to the electrode array. Note that the sensation on the tongue may not have the same resolution and certainly does not have the same number of grey levels. (a) Trainer's hand seen within the default 24° field of view. The gaps between fingers may not be easily recognisable with the tongue. Participants report that this stimulation feels like a single large blob. (b) Zoomed-in onto two fingers makes it possible to distinguish the two fingers and the gaps between the fingers.

recognising orientation. In head tracing, they moved the head-mounted camera and noted the direction of head movement necessary to keep the stimulation generated by the bar stimulus essentially in the same position on the IOD during the movement. In tip tracing, they traced the stimulation on the IOD with the tip of their tongue while keeping the head static. Some of the participants used the tracing strategies intuitively. For others having difficulty in this task, the trainer introduced to both strategies. All participants reported benefiting from using head or tip tracing.

Localisation

The trainer attached a circle (about 8° in diameter) at a random location on the board (about 87°×57°) and instructed the participants to locate it within the 24°×24° default field of the device using head scanning movements. The purpose of this task was to familiarise participants with the relationship between head movements and the external world (e.g. when they move their head towards left in relation to their body, they would detect a stimulus on their left). No specific scanning strategy was suggested. Participants typically used random movements in searching for the stimuli.

Counting targets

The trainer placed multiple circles (2, 3, or 4) of different sizes at different locations on the board and asked participants to count the number of objects on the board. The purpose of this task was to strengthen the localisation skill of Stage 3. Keeping track of each found object's location improved their ability to integrate head movement/spatial location relationship; without keeping track of each object's spatial location, they might count the same object more than once. After the participant was able to correctly report the number of objects on the board, the trainer brought the objects closer together and encouraged the participant to differentiate the objects even when there was only a small gap between them. To make the task more interesting, sometimes the trainer asked participants to locate the biggest stimulus on the board.

Using zoom-in and -out

The trainer introduced participants to the zoom control of the system and explained to the participant that zooming-out is an efficient method for detection and localisation (as it increases the field of view), while zooming-in is helpful for recognising details of the visual stimulus (as it increases the apparent resolution). The trainer asked participants to locate a stimulus on the board by first zooming-out, head scanning to locate the stimulus, and then zoom-in to try to recognise the details of the stimulus. Participants were asked to locate and examine their own hand or the trainer's hand using the zoom functions. In the default zoom setting, the whole hand (at arm length) is visible within the field of view, but the gaps between fingers cannot be appreciated on the IOD (see Figure 2 for illustration). Participants were taught to zoom-in and explore the fingers and spaces between them at a higher resolution. The hand exploration task in the default zoom setting also helped participants to comprehend the field of view of the camera by moving the hand through the field and exploring the extent of stimulation on the IOD caused by their own hand from the arm-length distance.

Discriminating the direction of Tumbling E and Landolt C stimuli

To become more familiar with the zooming function, participants were trained to identify the direction of the opening in a tumbling E and Landolt C targets placed on the board in one of the four different directions (up, down, left, or right).

Recognising basic shapes

Simple geometrical shapes cut out of white felt cloth such as triangle, square, circle, hexagon, diamond, arrow, and heart (spanning approximately 10°) were placed on the black board. Participants were asked to identify the shape by utilising both zooming-in and tracing of the electrode array with the tip of the tongue or by tracing with the head, and name it. During early pilot testing of the device and training of the first few participants, both "filled" and "outlined" shapes were used [15]. Our participants reported that recognising outlined shapes was easier because the outline was easier to trace with the head or the tip of the tongue. Therefore, only the outlined shapes were used in most of the training sessions for most participants. Note that creating contours with image processing may also convert filled shapes to outlined shapes [16]. The BrainPort provides two edge enhancement modes: In the first mode, image is converted to just edges. In the second mode, edges are calculated and presented on top of the image. Neither of these was used in our training.

Recognising letters and numbers

The participants (specifically congenitally blind participants) were first asked if they are familiar with letters and numbers. If they

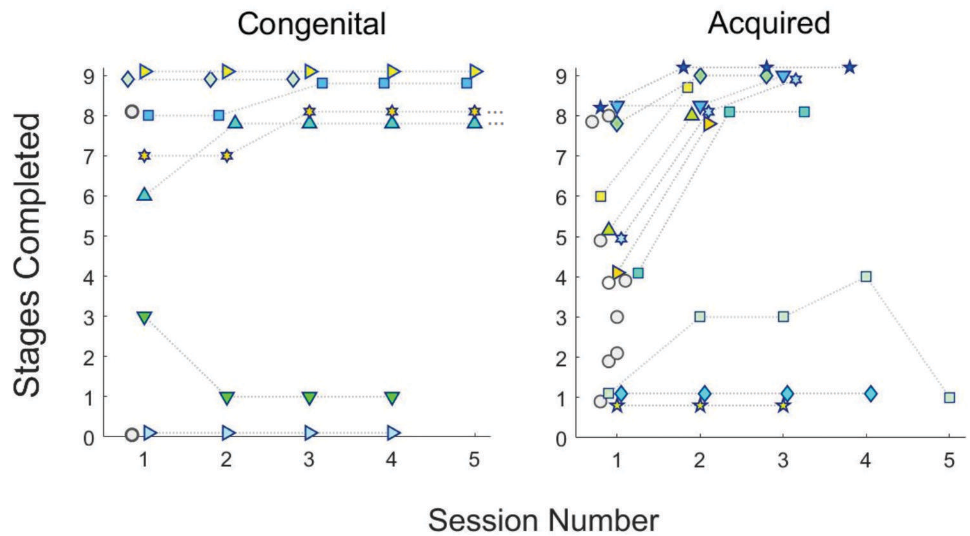

Figure 3. The training stages completed in the first five training sessions for all 29 participants. The data shows that if participants can complete Stage 4 in the first session, they can complete all the stages (or reach the last Stage) in the training protocol within the first five sessions. Each symbol represents one participant. Participants who attended only one training session were indicated by grey circles. Participants who attended more than five sessions are indicated by three dots at the end of the fifth session.

were not confident, (capital) letters and numbers cut from cloth were put on the table one by one and they were asked to identify them by tracing with their fingers. All the participants were able to identify most of the letters and numbers. Only letters and numbers that participants could identify by touch were used for further training. The participants were asked to name letters and numbers (approximately 10° height) by utilising zooming-in and tracing the IOD with the tip of the tongue or by using head tracing. For some participants, only a single letter at a time was presented on the board. For others who could easily recognise a single letter on the board, three letters, which would form a word such as "DOG", were presented. For a few highly performing participants, ten letters were placed randomly on the board, and participants were asked to locate a target letter among the others.

Recognising real objects on a table

An actual object (i.e. a mug with a handle, a paper coffee cup, a funnel, a hand sanitiser bottle) was placed on the table in front of the participant. The participants were asked to name the object using the BrainPort without prior knowledge of objects (in their first exposure to each object). After they explored each object using the BrainPort, they could touch the object. Next, multiple objects (any combination of the four objects mentioned above) were placed on the table, and participants were asked to name the objects and verbally report the location of each object relative to others. The participants were also introduced to occlusion and perspective in this stage, as they were encouraged to look at the objects from different angles (multiple viewpoints) and different physical distances (perspective) while they were sitting on a chair or standing up. Seeing objects from multiple viewpoints did not provide much additional information for circular symmetric objects such as a paper coffee cup, but it did help to detect distinct properties of asymmetrical objects such as the handle of a mug. Also, in order to determine the size of the objects, we suggested participants to use their thumb or index finger at arm's length as a reference of size, but keeping the finger and an object within the same IOD field of view was found to be difficult because of the occlusion by the finger.

Results and discussion

Each participant attended different number of sessions (minimum 1, maximum 11 sessions; See Figure 3 for each participant's progress in the first five sessions. Out of nine participants with congenital blindness, six (66%) reached the last stage in the training protocol (i.e. successfully recognised letters and numbers); 4 at the end of the 1st session, 1 at the end of the 2nd session and 1 at the end of the 3rd session. Three out of the six were successful in recognising the real objects satisfactorily. Some participants were tested on all letters that they are familiar with and some with just a few of these letters depending on time restrictions in the session and participant's interest. Some of them were able to name the letter in the second or third attempt, and those were considered as a correct response. Two congenitally blind participants could not progress further than Stage 1 (one reached Stage 3 in the 1st session but the performance was worse in the following sessions) even after four training sessions (approximately 12 h of training and they were not trained further). One was only able to tell if stimulation is on or off at the end of the 1st session (he was not trained further). Two out of nine participants attended only one training session lasted approximately 3 h (one could only tell if the stimulus is on/off and the other completed Stage 8).

Out of 20 participants with acquired blindness, 10 (50%) reached the last stage in the training protocol (5 at the end of the 1st session, 5 at the end of the 2nd session). Five out of the 10 (who reached the last stage) were successful in recognising the real objects satisfactorily. Four out of 20 could not progress further than Stage 1 (i.e. spatial mapping; 1 at the end of the 4th, 1 at the end of the 3rd, and 1 at the end of the 1st session. One reached Stage 4 at the 4th session but performance dropped at the 5th session). Five could progress only up to Stage 2 ($N = 2$), Stage 3 ($N = 1$), or Stage 4 ($N = 2$) at the end of the 1st session (they were not trained further). Nine out of 20 participants attended only one training session (see the grey circles in Figure 3, right plot).

Out of 29 blind participants, 12 either dropped out or did not want to continue to participate in the training. Two dropped out of the study because they were not able to keep the device in the mouth for long durations without gagging. Two showed no progress beyond Stage 3 even after at least four sessions, and two more showed no progress beyond Stage 1 even after at least three sessions. Therefore, these four participants were not called back for further training. Six reported that they were not interested in participating in the training anymore because of disinterest in the device (4 at Stage 8, 1 at Stage 4 and 1 at Stage 2).

The performance in the first two sessions seems to be a good indicator of participants' ability to progress in the training protocol. Participants who could complete at least the fourth stage within the 1st session were at least able to recognise letters and numbers in the following sessions. Participants who were stuck at one of the first four stages in the 1st session did not show any further progress, for some of them, even after four training sessions. Therefore, the stage completed in the first session seems to be a good indicator of future progress. As can be seen in the Figure 3, most (15 out of 16 participants) who can recognise at least letters and numbers reached this performance level within the first two sessions. Also, although some participants completed all stages, they were asked to attend more training sessions to see if further practice affects the performance at least in terms of recognition time. We haven't noticed such an improvement.

Observations on the training stages

Device familiarization and spatial mapping of the tongue in relation to the visual world

During the spatial mapping task, most of the participants were not able to detect stimulation at the back of the tongue with the adjusted comfortable intensity level. They reported feeling the stimulation when stimulation is moved towards the tip of the tongue (lower visual field) which is consistent with reports in the literature about lower sensitivity at the back of the tongue [11,13,17,18]. Two participants were not able to feel stimulation on the right side of the tongue. Another participant was not able to feel stimulation on any part of the tongue except the tip. Yet, all three participants were able to advance in the training using tip tracing and completed the Stage 8.

It was difficult to explain the lower and upper visual field especially to participants with congenital blindness. Therefore, we tried to give them real life examples to clarify the spatial configurations in a visual environment and its representation on the IOD by the device. For instance, we told participants that if they are looking at a person, the head is in the upper visual field, and the feet are in the lower visual field. Hence, the head should be presented on the back of the tongue (on the upper rows of the IOD), and the feet should be presented on the tip of the tongue (on the lower rows of the IOD closer to the cable). Few participants stated that the spatial configuration of the device was not intuitive for them as they thought that it would make more sense for them to present the upper visual field on the tip of the tongue. This observation is consistent with previous studies which reported individual differences in natural spatial perspective for spatial patterns presented on the body [19–21]. Yet, there was not such reversal in the left/right relationship.

Static bar orientation

Most of the participants were not able to perceive the orientation of the bar when they were first exposed to the stimulation. They developed different strategies to succeed in this task. For example, many participants developed a strategy to move their head horizontally and vertically. If the bar is vertical, the stimulation does not change when they slightly move their head up and down, and thus they concluded that the bar is vertical. Similarly, when the bar is horizontal, stimulation does not change when they move their head from left to right. This strategy suggests that they did not actually perceive the orientation but used a cognitive approach to discriminate the change in sensation with the head movements [22,23]. Another strategy they used was tracing the stimulus with the head or tip of the tongue (See "2. Static Bar Orientation" under the method section for more details). Note that head tracing is different than head scanning strategy. In the head scanning strategy, they did not trace the stimulus; they moved their head along different directions and noted the changes in stimulation. They usually cognitively assembled the percept using the smaller parts they could locate. For instance, instead of perceiving a smooth diagonal bar, they reported detecting stimulation "on the top right and bottom left of the IOD," and then interpreted that "it should be a diagonal/left-tilted bar." Thus they were able to perform in the multiple-choice task without actually perceiving bar or line orientation at any instant. Similar cognitive strategies have been recently reported to be used by blind people implanted retinal prosthesis in orientation and motion discrimination tasks [23].

Localisation & target counting

Participants were successful in localising objects by head scanning and bringing the detected stimulation roughly to the centre of the IOD (trainer tracked the position in the BrainPort monitoring view which can be seen in Figure 1). However, participants with congenital blindness had difficulty in relating the spatial location of the object to their body and on the blackboard. For instance, even if they move their head far to the right, if the stimulation falls onto the left side of the IOD, they tended to ignore the position of their head and reported that the object is on their left side relative to their body. A similar problem was reported for patients with tunnel vision using a head-mounted display system to search for objects outside their residual central field [24]. In that situation, the patients responded with eye and head movement based on the position of the image of the target relative to the fovea without regards to the head position [25]. Therefore, an extra time was required to explain the relationship between head position and observed visual field to congenitally blind participants. In addition, some participants assumed that the camera was located on the controller held in the hand in the BrainPort V100 version of the device. Therefore, with this version of the device, it took even more time for participants to learn the head position/spatial location relationship. The later versions of the device appeared to solve this problem by moving the controller to the headset. However, with that heavier headset, some participants reported that the new design was not comfortable causing headache or skin irritation.

Introducing zoom & discriminating Tumbling E and Landlot C direction

The concept of zooming-in and -out (field of view change) was hard to grasp for most of the congenitally blind participants. They usually mistook it for a function showing distance between the stimulus and the camera or varying intensity. Participants with acquired blindness grasped the concept of zooming easily.

Participants mainly used one of two strategies to detect the direction of the gap in Tumbling E or Landlot C. Some of them zoomed-in all the way so that only a small portion of the test target was presented on the IOD and traced the stimulus with their head. Participants were allowed to use similar head tracing movements (or head-like movements *via* a computer mouse) in

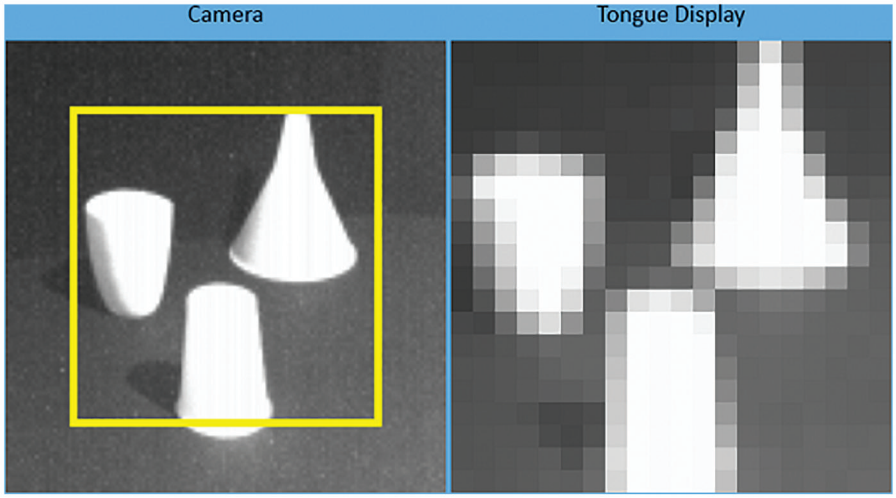

Figure 4. A screenshot taken from BrainPort monitoring screen with real objects on the table. Left panel shows the camera view and the yellow square indicates the IOD's field of view. Right panel shows the IOD simulated view. Although limited, there are some depth cues in the camera view such as shadows of the objects and the faint boundary between the table and the wall enabling interpretation of the depth with normal sight. Although shadows may be seen in the IOD simulated view, at least at a minified view, they do not seem to be perceivable on the IOD, perhaps because of the limited dynamic range on the tongue [17]. Stimulation on the tongue seems mostly binary [13].

previous acuity studies with Tumbling E or T [26,27]. Others zoomed-in only until all parts of the object were presented on the IOD and then traced the pattern with the tip of the tongue. Some participants had difficulty using gentle head movements. Therefore, they frequently lost the stimulus when zoomed-in. For those participants, tip tracing was more effective. Participants who used the zoom effectively progressed better in the subsequent stages.

Recognising shapes & letters and numbers

Although Vincent et al. [15] showed that recognition performance was not affected by the shape being filled or outlined while using the BrainPort, our participants preferred the outlined shapes because they were easier to trace with the head or tip of the tongue. As in the bar orientation task, participants first reported the components of shapes, letters, and numbers using the zoom-in function with either head or tip tracing. They then deduced the response from the components they identified. For instance, one participant stated that "There are two legs, and a round top. Therefore, it should be R." It seems that top-down processing is frequently involved in these forced-choice tasks. Participants combine the information that they could detect from the stimulation with the possible responses based on the assigned task and infer the response. Therefore, these tasks required a lot of focus and cognitive processing, which is fatiguing [28]. One participant who was successful at recognising letters and numbers stated that "It takes a lot of focus and time so I cannot imagine wanting to use it regularly although I enjoy the in-lab tasks."

Eight blind and eight normally-sighted participants who trained under this protocol participated in a study where we tested different presentation modes with a letter recognition task [13]. All of them completed at least Stage 8, letter and number recognition, and an additional targeted training specific to the task tested. Although the task in the experiment was similar to the training task at Stage 8 (except time limitation), percent correct performance varied from 39% to 69%. This indicates that an acceptable (above chance) performance in the training does not necessarily lead to a nearly perfect performance even in a relatively similar task. It should be noted that normally-sighted participants' performance was not included in this report because the training protocol was not strictly followed when training them, and most of them received a shorter training for an hour. Despite the differences in training between blind and normally-sighted participants, both groups performed similarly in the experiment (See Pamir et al. [13] for further details).

Recognising real objects

When moved from the board to the tabletop, due to two-dimensional to three-dimensional expansion of the environment, more variables were inevitably added to the task including distance, perspective, and occlusion, all of which created some confusion for the participants. For instance, many participants perceived objects located at different distances from the participant as if they were located at different heights, rather than different distances. This interpretation makes sense because when a three-dimensional scene is converted into a static two-dimensional low-resolution image, most depth cues are unavailable without reference of size and distance, and the outcome image is consistent with participants' reports (See Figure 4 for an example).

Another major challenge for recognising real objects in space with the BrainPort is the size and distance relationship. When both are unknown to the user, it is almost impossible to decide if the target is a closer small object or a large object located far away, which makes bottom-up recognition difficult. Most of our participants reported the size of objects relative to the other objects in the scene. For instance, when a mug, a sanitiser bottle, and a funnel are placed side by side on the table, participants reported which one is the tallest and which one is the widest based on the image position on the IOD. We had great difficulty in explaining this relationship to congenitally blind participants. Because this is not an intuitive concept for them, even if they comprehend the logic behind it, they frequently tended to forget about it while concentrating on recognition. As reported above, attempts to scale object size with subject hand was not successful due to interference between the hand image and the other object.

General observations

Our experience with BrainPort training suggests that the length of the training is not a critical factor in reaching an advanced stage with our current training protocol. There are large individual differences in how far and how fast each participant can progress in the training protocol [29].

Object recognition with the BrainPort was mainly achieved through top-down processing using prior knowledge of the training stimuli and environment. Participants usually recognise the parts of the stimulus and assemble it in their mind based on the possible responses. Therefore, even though they are successful in the training, the generalizability to everyday tasks is questionable.

Usually participants could successfully progress in the protocol to the last stage if they were able to use head or tip tracing and zooming function effectively. The limited spatial resolution of the tongue might be the reason why these strategies are essential to be successful in recognising stimuli with SSDs [13].

Acknowledgments

We would like to thank Kassandra Lee, Cyril Nyankerh, and M. Umut Canoluk for their valuable contributions in training participants.

Author contributions statement

All authors conceived and formulated the study. Z.P drafted the manuscript. All authors edited the manuscript.

Disclosure statement

Drs. Peli and Jung have a patent and a patent application on image processing for visual prosthesis, both assigned to the Schepens Eye Research Institute. Dr. Z.P. declare no competing interests.

Funding

This work was supported by the US Department of Defense under Grant W81XWH-16-1-0033; NIH core grant under Grant P30-EY003790.

ORCID

Zahide Pamir https://orcid.org/0000-0003-1813-400X
Jae-Hyun Jung http://orcid.org/0000-0002-5709-6045
Eli Peli http://orcid.org/0000-0002-1340-9257

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article.

References

[1] Loomis JM, Klatzky RL, Giudice NA. Sensory substitution of vision: importance of perceptual and cognitive processing. In: Manduchi R, Kurniawan S, editors. Assistive technology for blindness and low vision. Boca Raton (FL): CRC Press; 2012. p. 162–191.

[2] Stronks HC, Mitchell EB, Nau AC, et al. Visual task performance in the blind with the BrainPort V100 vision aid. Expert Rev Med Devices. 2016;13(10):919–931.

[3] Nau AC, Pintar C, Arnoldussen A, et al. Acquisition of visual perception in blind adults Using the BrainPort artificial vision device. Am J Occup Ther. 2014;69(1): 6901290010p1–6901290010p8.

[4] Deroy O, Auvray M. Reading the world through the skin and ears: a new perspective on sensory substitution. Front Psychol. 2012;3:457.

[5] Grant P, Spencer L, Arnoldussen A, et al. The functional performance of the Brainport V100 device in persons who are profoundly blind. J Vis Impair Blind. 2016;110(2):77–88.

[6] Stiles NRB, Shimojo S. Auditory sensory substitution is Intuitive and automatic with texture stimuli. Sci Rep. 2015; 5:15628.

[7] Wicab Inc. Madison (WI); [cited 2020 Jun 30]. Available from: https://www.wicab.com/brainport-vision-pro

[8] Nau A, Bach M, Fisher C. Clinical tests of ultra-low vision used to evaluate rudimentary visual perceptions enabled by the BrainPort vision device. Transl Vis Sci Technol. 2013; 2(3):1.

[9] Ptito M, Matteau I, Wang AZ, et al. Crossmodal recruitment of the ventral visual stream in congenital blindness. Neural Plast. 2012;2012:304045.

[10] Matteau I, Kupers R, Ricciardi E, et al. Beyond visual, aural and haptic movement perception: hMT+ is activated by electrotactile motion stimulation of the tongue in sighted and in congenitally blind individuals. Brain Res Bull. 2010; 82(5–6):264–270.

[11] Moritz JJ, Turk P, Williams JD, et al. Perceived intensity and discrimination ability for lingual electrotactile simulation depends on location and orientation of electrodes. Front Neurosci. 2017;11:186.

[12] Lee VK, Nau AC, Laymon C, et al. Successful tactile based visual sensory substitution use functions independently of visual pathway integrity. Front Hum Neurosci. 2014;8:291.

[13] Pamir Z, Canoluk MU, Jung JH, et al. Poor resolution at the back of the tongue is the bottleneck for spatial pattern recognition. Sci Rep. 2020;10(1):2435.

[14] Kurlowicz L, Wallace M. The mini-mental state examination (MMSE). J Gerontol Nurs. 1999;25(5):8–9.

[15] Vincent M, Tang H, Zhu Z, et al. Shape discrimination using the tongue: implications for a visual-to-tactile sensory substitution device. Multisens Res. 2016;29(8):773–798.

[16] Jung J-H, Aloni D, Yitzhaky Y, et al. Active confocal imaging for visual prosthesis. Vision Res. 2015;111(Pt B):182–196.

[17] Lozano CA, Kaczmarek KA, Santello M. Electrotactile stimulation on the tongue: intensity perception, discrimination, and cross-modality estimation. Somatosensory Motor Res. 2009;26(2–3):50–63.

[18] Wilson JA, Walton LM, Tyler ME, et al. Lingual electrotactile stimulation as an alternative sensory feedback pathway for brain-computer interface applications. J Neural Eng. 2012;9: 045007.

[19] Arnold G, Spence C, Auvray M. Taking someone else's spatial perspective: natural stance or effortful decentring? Cognition. 2016;148:27–33.

[20] Arnold G, Auvray M. Tactile recognition of visual stimuli: specificity versus generalization of perceptual learning. Vision Res. 2018;152:40–50.

[21] Richardson ML, Lloyd-Esenkaya T, Petrini K, et al. Reading with the tongue: Individual differences affect the perception of ambiguous stimuli with the BrainPort. Proceedings of the 2020 CHI Conference on Human Factors in

Computing Systems; Honolulu, HI: Association for Computing Machinery; 2020. p. 1–10.
[22] Peli E. Testing vision is not testing for vision. ARVO; Vancouver, Canada. J Vision. 2019;19(8):51.
[23] Hallum LE, Dakin SC. Retinal electrode array implantation to treat retinitis pigmentosa: a systematic review. University of Auckland. p. 1–17.
[24] Luo G, Woods RL, Peli E. Collision judgment when using an augmented-vision head-mounted display device. Invest Ophthalmol Vis Sci. 2009;50(9):4509–4515.
[25] Luo G, Peli E. Use of an augmented-vision device for visual search by patients with tunnel vision. Invest Ophthalmol Vis Sci. 2006;47(9):4152–4159.
[26] Sampaio E, Maris S, Bach-y-Rita P. Brain plasticity: 'Visual' acuity of blind persons via the tongue. Brain Res. 2001; 908(2):204–207.
[27] Ptito M, Moesgaard SM, Gjedde A, et al. Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. Brain. 2005;128(Pt 3):606–614.
[28] Erickson-Davis C, Korzybska H. What do blind people "see" with retinal prostheses? Observations and qualitative reports of epiretinal implant users. BioRxiv. 2020. DOI:10.1101/2020.02.03.932905
[29] Arnold G, Pesnot-Lerousseau J, Auvray M. Individual differences in sensory substitution. Multisens Res. 2017;30(6): 579–600.