Citation for published version:
Richardson, M, Lloyd-Esenkaya, T, Petrini, K & Proulx, M 2020, Reading with the Tongue: Individual Differences Affect the Perception of Ambiguous Stimuli with the BrainPort. in *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, pp. 1-10. https://doi.org/10.1145/3313831.3376184

DOI:
10.1145/3313831.3376184

Publication date:
2020

Document Version
Peer reviewed version

Link to publication

Reading with the Tongue: Individual Differences Affect the Perception of Ambiguous Stimuli with the BrainPort

Mike L Richardson
Department of Psychology, University of Bath
m.richardson@bath.ac.uk

Tayfun Esenkaya
Department of Computer Science, University of Bath
t.esenkaya@bath.ac.uk

Karin Petrini
Department of Psychology, University of Bath
k.petrini@bath.ac.uk

Michael J Proulx
Department of Psychology, University of Bath
m.j.proulx@bath.ac.uk

ABSTRACT

There is an increasing interest in non-visual interfaces for HCI to take advantage of the information processing capability of the other sensory modalities. The BrainPort is a vision-to-tactile sensory substitution device that conveys information through electro-stimulation on the tongue. As the tongue is a horizontal surface, it makes for an interesting platform to study the brain's representation of space. But which way is up on the tongue? We provided participants with perceptually ambiguous stimuli and measured how often different perspectives were adopted; furthermore, whether camera orientation and gender had an effect. Additionally, we examined whether personality (trait extraversion and openness) could predict the perspective taken. We found that self-centered perspectives were predominantly adopted, and that trait openness may predict perspective. This research demonstrates how individual differences can affect the usability of sensory substitution devices, and highlights the need for flexible and customisable interfaces.

Author Keywords

Sensory substitution; tactile interfaces; individual differences in computing; user preferences.

CSS Concepts

• **Human-centered computing** → *Laboratory experiments; empirical studies in HCI.*

INTRODUCTION

Human-computer interaction (HCI) includes a great number of interface methods, ranging from the conventional monitor, keyboard and mouse; the now widespread touch screen; and the increasingly popular voice interfaces of Siri and Alexa [21]. The 'Brainport' is a device that lies towards the obscure end of the interface spectrum, providing tactile feedback to the user through electrical stimulation on the tongue [7, 14]. It was developed as a vision-into-tactile sensory substitution device (SSD), a machine that converts the information available in one sensory modality into another [31]. However, it also has untapped potential for use as a novel way to provide information through tactile means that could be generalized to other parts of the body. This, in turn, can help further reveal the brain's representation of, and interaction with, space [1, 3].

Previous research has suggested that the tongue is an 'ideal' surface for sensory displays, often citing reasons such as sensitivity, moistness (therefore a better conductor of electrical stimulation, requiring less power consumption than other tactile methods such as vibrating motors), and leaving the hands, ears, and any residual vision free for other purposes [6, 10, 13, 27]. Using the tongue as a display surface provides some unique issues that must be considered. As the camera is designed to be head-mounted, the video feed is as one would view on a traditional screen (see Figure 1 for how the Brainport converts a visual signal). However, because the tongue is a horizontal surface, the video feed must undergo some transformation. As there are not many situations that can quickly be brought to mind in which humans naturally convert vertical space into horizontal, the need for a device to be intuitively mapped may be consequential to enjoyment and uptake, or dropout from device learning [9]. Furthermore, individual differences such as being introverted and extroverted may affect how a user converts this information (e.g., introverted users may take a more self-centered perspective than extroverted users when transferring the vertical information to horizontal). In fact, personal factors like gender [32], and personality [30], can influence how the brain relates to space.

The sensations provided by the BrainPort are novel, and consequently, their processing is cognitively complex. Previous work has shown the benefits of dedicated practice to improve SSD comprehension and also BrainPort specific comprehension [31]. However, as there are a number of ways in which the picture from the BrainPort could be flipped to the flat surface of the tongue, we investigated the influence of camera orientation, gender and personality on the perception of directionally ambiguous stimuli; this will provide insight about the impact of seemingly fundamental individual differences in the brain's integration of space.

CHI '20, April 25–30, 2020, Honolulu, HI, USA

ACM ISBN 978-1-4503-6708-0/20/04…$15.00
https://doi.org/10.1145/3313831.3376184

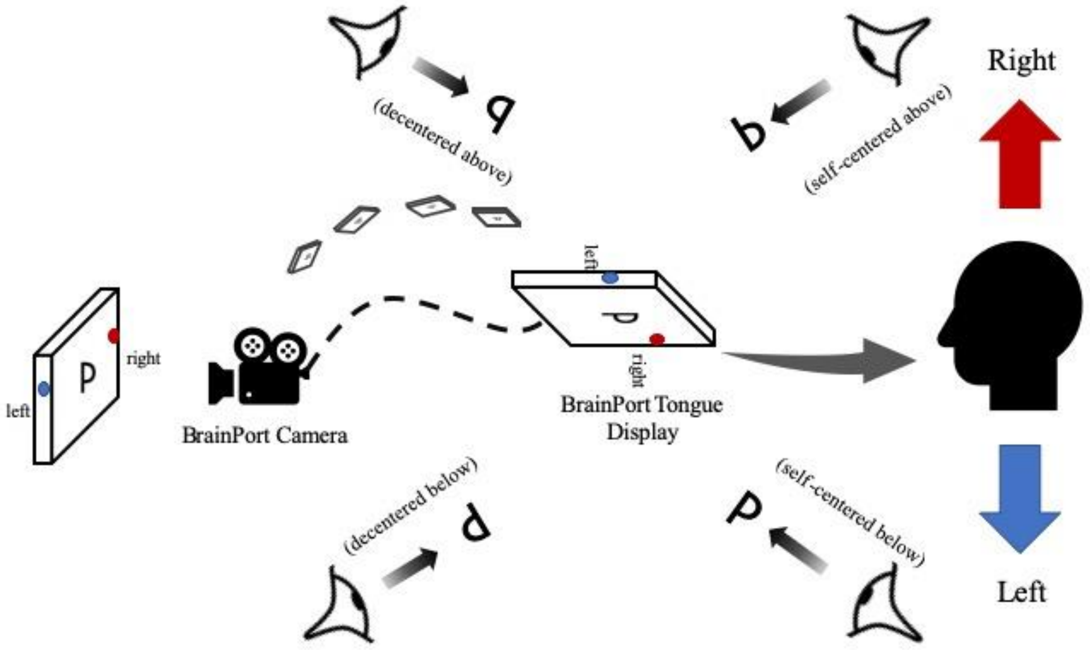


Figure 1. Possible interpretations of the ambiguous letter 'p' when viewed through the BrainPort's tongue display. The top of the camera's field of view appears on the back of the tongue. Whichever perspective the user adopts, will change the percept of the stimuli.

To better understand the process of transforming vertical spatial information into the plane of a horizontal tongue display, and how the BrainPort achieves this, it might be helpful to draw the letter 'p' on a note of paper (mentally or physically). Then, flip the paper over and trace over the 'p' on the opposite side to create a single shape through the plane of the paper (i.e., if the reader were to cut around the shape, it would be the same on either side). To imagine how the BrainPort converts an image, hold the note in front with both hands, like an open book (the note should read 'p'). Now, bending the elbows, bring the top of the 'p' toward the mouth until the paper is horizontal (looking down at the note should read 'b'), this is the shape that would be electro-tactilely vibrated into the tongue (as if one were to drop the paper 'p' onto the tongue).

The fields of sensory substitution and augmentation have several widespread applications which all would benefit more in-depth literature from which to draw on. The most discussed application for this technology is arguably, to recover some perceptual losses from a sensory impairment, with many individuals already using SSDs in their everyday life [33]. There are also specialist uses being considered where further knowledge about how the brain interprets the SSD are unquestionably vital. One example of this is tactile feedback for firefighters; when smoke levels are too high for clear visual search, an ultrasonic rangefinder paired to haptic gloves can offer a new search perspective [12]. Another important employment of tactile feedback research is being used in Human-Drone interfaces [2, 19]. As drones become more available and capable of more complex tasks, providing environmental or navigational information through tactile means could allow the pilot to concentrate visual attention on flying.

As the display needs to map to the external world and we cannot assume that a display of information to any given sense or (in the case of tactile feedback) any location on the body, would be the same as others, so introducing the TDU for non-blind, non-vestibular, other HCI uses, requires this knowledge produced here (e.g. which way is 'up' on the tongue?). While camera orientation is not a unique consideration to tongue displays, we posit that the uniqueness of translating information from an external vertical plane to an internal horizontal one is, so we assessed that here as well.

RELATED WORK

Brown, Macpherson, and Ward [11] conducted an experiment with the vOICe, a vision-to-sound SSD, where they examined the effect of different camera conditions on ease of object location and discrimination. They found that holding the camera with the hand, led to more easily identified objects compared to head mounting, possibly suggesting that individuals can readily shift their perspective to that of the camera. This could potentially mean that when identifying objects through an SSD hand-held camera, users take a camera-embodied perspective (and move their locus of attention with the camera lens). Brown [11] chose to examine the difference between hand-held and head-mounted camera angles, rather than between different positions for the hand-held camera or head-mounted camera. This left a gap for future research to examine the usability of different hand-held camera orientations (for example, in front, or above).

Recent research into low-resolution SSDs (only 128 pixels) found that participants could still make remarkably accurate spatial judgments [26]. While low-resolution devices are useful for gross tasks, such as movement and navigation, they may not provide enough information to form complex perceptions. Studies using the BrainPort have previously demonstrated that participants can quite easily identify rotating letters, reducing in size down to only a few millimeters on the tongue [13, 24, 27]. These experiments typically used the Snellen Tumbling E test (a rotating E that gradually decreases in size), which is useful to measure acuity, as by an optician, but not perspective on a tongue display, as the E is symmetrical along the horizontal axis. As such, the E would appear the same when viewed from above and below. To date, no one has yet tested the BrainPort with truly ambiguous letters that would change meaning when viewed from alternate perspectives, including a combination of decentered, self-centered, above, and below.

A study from Arnold, Spence, and Auvray [5] used vibrotactile motors to stimulate the letters of b, p, q, and d on the torso of participants to observe which perspective was taken. They reported that three different perspectives could likely be taken: 1) head-centered (as if one was looking from the head down at the letter), 2) trunk-centered (perceiving directly forward from the torso), and 3) decentered (perceiving as if looking at the torso from a second-person perspective), which were adopted by 30%, 50%, and 20%, of participants, respectively. Work previous to that of Arnold and colleagues used similar methods to examine adopted perspectives [17, 25, 29], however, in Arnold's and colleagues' study a tactile matrix was used rather than having an experimenter draw the letters onto participants. Removing the experimenter seemed to decrease the likelihood of adopting a decentered perspective in comparison to these past works, although not completely. Arnold, Spence, and Auvray [5] suggest that to some individuals, the decentered perspective 'may be their default' (p. 31), but for most, the presence of an experimenter creating the symbols, sways the perspective to that of the experimenter (decentered). In addition, a study found that individuals with good social skills can more freely adopt a decentered perspective [30]; by taking the other's perspective spatially, they can further grasp the other's perspective empathetically [28]. High trait extraversion and openness, and gender may also serve as markers for the flexible social skills that are required to step into the others' perspective [15, 22, 31]. For example, females have been shown to perform more successfully than males on activities designed to test spatial perspective taking, when the task is dependent on social factors [32].

A follow up review conducted by Arnold, Spence, and Auvray [4] indicated that spatial, personal, and interpersonal factors could influence the perspective adopted when perceiving tactile letters on the body; as part of this review, they included a meta-analysis of studies presenting tactile letters to the head (on the forehead). When discussing the possible perspectives that could be adopted in response to a tactile letter on the forehead, the most common distinctions were found between decentered (looking directly at the forehead from a second-person perspective), or self-centered (perceiving directly forward from the forehead). Furthermore, Arnold, Spence, and Auvray [4] showed that most studies reported the self-centered perspective as most often adopted. However, these experiments do not consider that a perspective could be taken from the eyes, looking up at the letter, much in the way that some participants took a head-centered perspective in the torso experiments (looking down at the letter). The current rhetoric seemingly classifies that head-centric perspectives are generally adopted because it is the head in which vision resides [4, 8], but this may depend on where the stimulus is located (e.g., the head rather than the torso).

DESIGN

We aimed to test some of the spatial and personal factors that could contribute to what perspective is taken when using the BrainPort. To examine spatial factors, we drew upon the work of Brown et al. (2011) and hypothesized that camera position would affect the perspective taken, as they found that holding a camera allowed for more successful object identification, possibly due to adopting the camera's point of view (PoV). We next examined the effect of certain personal factors (extroversion and openness) on decentered perspective. These factors were chosen as past research suggests that they may indicate social skills, and an ability to relate to others, spatially and empathetically [16, 23, 30].

Methods

Participants

Thirty-six individuals volunteered for the experiment (18 female, mean age = 20 ± 1 years), and were reimbursed £5 for their time. All the participants gave written informed consent but were unaware of the study's purpose. Ethics permission was granted by the Department of Psychology Research ethics committee, University of Bath [reference no. 0125-18-14]. After each participant's data collection was completed, they were debriefed, revealing all aspects of the study. The participants all reported no sensory impairments, and the majority were righthanded (N = 32). Sighted individuals were chosen (rather than visually impaired) as the specifics of the present experiment hoped to convey individual variation with anyone using a tongue display, rather than information pertaining to blindness or visually impairments only.

Materials and measures

In addition to the ambiguous letters, for greater generalisability, further stimuli were used in the present study giving four distinct stimulus groupings:

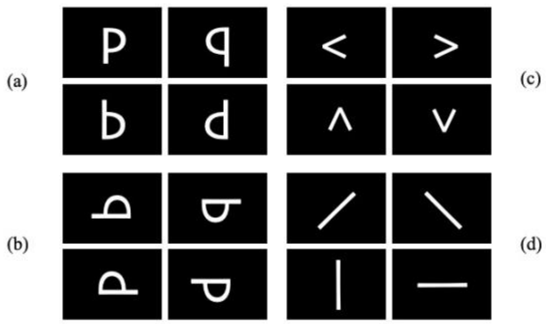


Figure 2. The stimuli that were presented to the BrainPort camera, the letters (a), rotated letters (b), arrowheads (c), and lines (d).

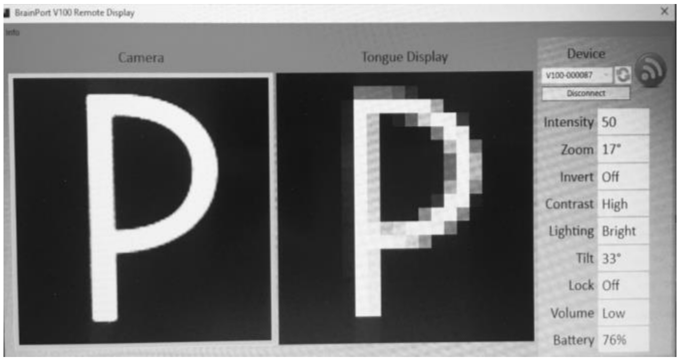
1) four letters; 2) two ambiguous letters ‘d’ and ‘q’ rotated by 90° in either direction; 3) arrowheads; and 4) lines of ascending, descending, vertical, and horizontal orientations (see Figure 2). The rotated letters served as a functional control since the participants would not be able to interpret the letter (see Figure 2b). The ambiguous letters (Figure 2a) could be interpreted in one of four ways depending on perspective, in that the letter ‘p’ could appear to be either a ‘p’, ‘b’, ‘q’, or ‘d’ from varying positions. However, the arrowhead (Figure 2c) and diagonal line (Figure 2d) stimuli could only realistically differ in perspective between opposite pairings (left vs. right, up vs. down, ascending vs. descending). The lines were included as additional measures of accuracy, but not of perspective-taking, as their appearance remains stable independent of perspective in the case of the horizontal and vertical orientation. That is, the diagonal lines could be interpreted as ascending or descending depending on whether they were viewed from above or below, from either the decentered or self-centered perspective, hence no measure of perspective is possible with these stimuli.

The BrainPort V100 (Wicab, USA) was used for the experiment. This is an older version of the device, which has since been updated into the form of the BrainPort Vision Pro. The device consists of a headset and a controller. The headset is formed of a camera mounted to sunglasses, and the tongue display (an array of 400 electrodes, arranged 20 × 20, spaced at 1.32 mm apart); the total size of the tongue display (29.5 mm × 33.8 mm × 7 mm) allows it to sit on the tongue comfortably and inside the mouth [20]. The controller houses the lithium-polymer battery pack, that provides the BrainPort with up to 2 hours of use, and also handles the image processing, along with buttons to control the output (for example, zoom, intensity, contrast). The vRemote (also developed by Wicab, USA) software allows a laptop to wirelessly view the configured settings, the camera input, and the tongue display output. Figure 3 shows how the BrainPort renders the video image to the tongue display, as viewed through the vRemote program. The initial settings were standardized (intensity = 50; zoom = 17°; invert = off; contrast = high); however, the intensity setting was manipulated to provide optimum comfort for the participants, while maintaining a clear projection of the stimulus, based on individual preferences.

Procedure

Prior to conducting the main BrainPort experiment, participants completed the Big Five Inventory (BFI-44; a questionnaire-based measure that aims to assume individuals’ personality traits, succinctly; [22]). The participants’ background information (age, gender, dominant hand) was collected, and they were blindfolded before being guided into the experimental room (to prevent visual information from influencing the user’s performance), then sat in front of the BrainPort. Before any data collection commenced, each participant was encouraged to explore the tongue display to familiarise themselves with it, while the experimenter adjusted the intensity to achieve the participant’s optimal comfort. A short training protocol was used to give the participants some practice with the stimuli and to make sure they understood the task. The training consisted of five trials, identifying a given number of dots, and explaining their appearance (e.g., two dots on the horizontal axis). In this phase, the researcher would give verbal feedback once a response had been made as to whether it was correct or not.

The main task consisted of three different conditions, each with 18 trials (two of the trials presented in the data collection were for other experiments, 16 of the trials were analyzed for the present experiment, with one trial per stimulus, see Figure 2). Participants were allowed up to 10 seconds to respond to each stimulus with a verbal answer. Participants were informed of the stimulus group and, therefore, knew whether to respond with a letter, arrow direction, or line orientation depending on the trial. The conditions consisted of no camera (NoCAM), vertical camera (VertCAM), and horizontal camera (HozCAM). In the no camera condition, participants were told that the stimuli were pre-recorded. In the vertical and horizontal conditions, the participants were given a fake camera to hold in a vertical and horizontal position (see Figure 4), respectively.



V100-000087

Intensity	50
Zoom	17°
Invert	Off
Contrast	High
Lighting	Bright
Tilt	33°
Lock	Off
Volume	Low
Battery	76%

Figure 3. A screenshot of the BrainPort’s input camera signal, and its rendered tactile output on the tongue display.

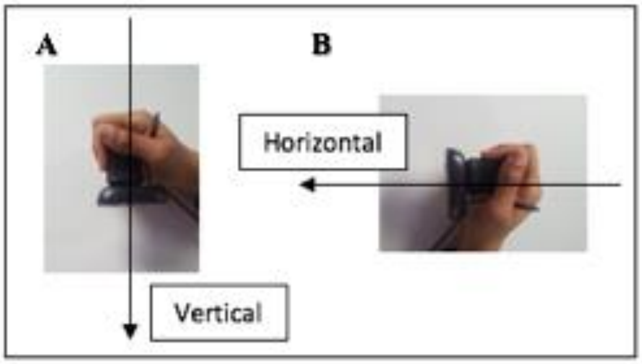


Figure 4. A demonstration of camera position in the vertical (A) and horizontal (B) camera conditions.

The researchers would occasionally make comments about steadying the camera to perpetuate the deception. The reason for using a fake camera was to facilitate identification of the stimuli and to allow for valid comparison to the no camera condition. The stimuli were identical in every condition and presented in a random order. Participants were given a break between conditions; in total, the experiment took roughly one hour to complete.

Data analysis

Due to the different number of possible interpretations depending on stimuli type, the letter-based stimuli were coded slightly differently to the arrows and lines. Additionally, the vertical and horizontal lines were not factored into perspective-taking, only for calculating response accuracy.

For the letters, responses were coded with a number from 1 to 4 depending on the given answer (for the letter ‘p’: 1 = ‘b’ = self-centered from above; 2 = ‘p’ = self-centered from below; 3 = ‘q’ = decentered from above; 4 = ‘d’ = decentered from below, refer again to Figure 1). For the arrowheads, responses were coded as only either self-centered or decentered, as the direction would not change from higher or lower perspectives (for the arrow ‘<’: 1 = ‘left’ = self-centered; 4 = ‘right’ = decentered). Coding responses in this manner was arbitrary and aimed to force a clearer separation between self-centered and decentered during the analysis. The letters could be used to tease apart ‘decentered’ and ‘self-centered’, including the further perspectives of ‘above’ and below’. This was done by examining the most frequently adopted perspective, so that if a participant responded to the letters with perspectives ‘1, 1, 3, 2 (or self-centered above, self-centered above, decentered above, self-centered below’, they would be considered as predominantly self-centered above. We used a frequency driven perspective classification, as explained in the previous example, rather than using the average, as calculating the average across the four participant’s responses would have returned imprecise results. That is, taking the example above the average of 1, 1, 3, 2 responses would have been around 2, suggesting that that participant had a self-centered below perspective, despite only responding with perspective 2 on one occasion. If a participant reported different perspectives an equal amount of time, for example ‘2, 2, 3, 3’, or ‘2, 1, 4, 3’, then that participant was considered as having a mixed perspective. Accuracy for all stimuli was also measured by recording whether the answer was correct or not according to the BrainPort (e.g., if a ‘P’ was shown to the BrainPort camera, and the response ‘P’ was given), and is reported as a proportion across all trials, split between the camera conditions.

RESULTS

As the data were predominately categorical (with the exception of: proportion of correct responses, and extraversion and openness personality questionnaire scores), a chi-square test revealed that there was no association between camera orientation and perspective taken for the letters ($X_2(8, N = 108) = 10.04$, $p = .262$) or arrows ($X_2(4, N = 108) = 4.39$, $p = .356$). There was also no association between gender and perspective adopted for letters or arrows ($X_2(4, N = 108) = 3.19$, $p = .538$; ($X_2(2, N = 108) = 4.29$, $p = .117$, respectively).

Figure 5 shows the percentages of adopted perspectives for the letters and arrowheads within the participants and, it also shows the variation within the arrowhead stimuli between left/right and up/down arrows. As described above, participants were described as having a ‘mixed perspective’ in the case that they had equal self-centered and decentered responses, or no clear mode of response (e.g. for the letters, answering with each of the possible perspectives).

Camera Condition	Total Correct		Letters		Rotated Letters		Arrows		Lines	
	Median	*IQR*	*Median*	*IQR*	*Median*	*IQR*	*Median*	*IQR*	*Median*	*IQR*
No Camera	.75	.27	.75	.75	.25	.25	.75	.50	1	.25
Vertical Camera	.75	.50	.50	.75	.25	.50	.75	.50	1	.25
Horizontal Camera	.79	.32	.75	.75	.25	.25	.75	.50	1	.50

Table 1. Medians and interquartile range for proportion of correct answers given to the different stimuli when observing them through the BrainPort (e.g. if a ‘P’ is shown to the camera, the participant responses with ‘P’).

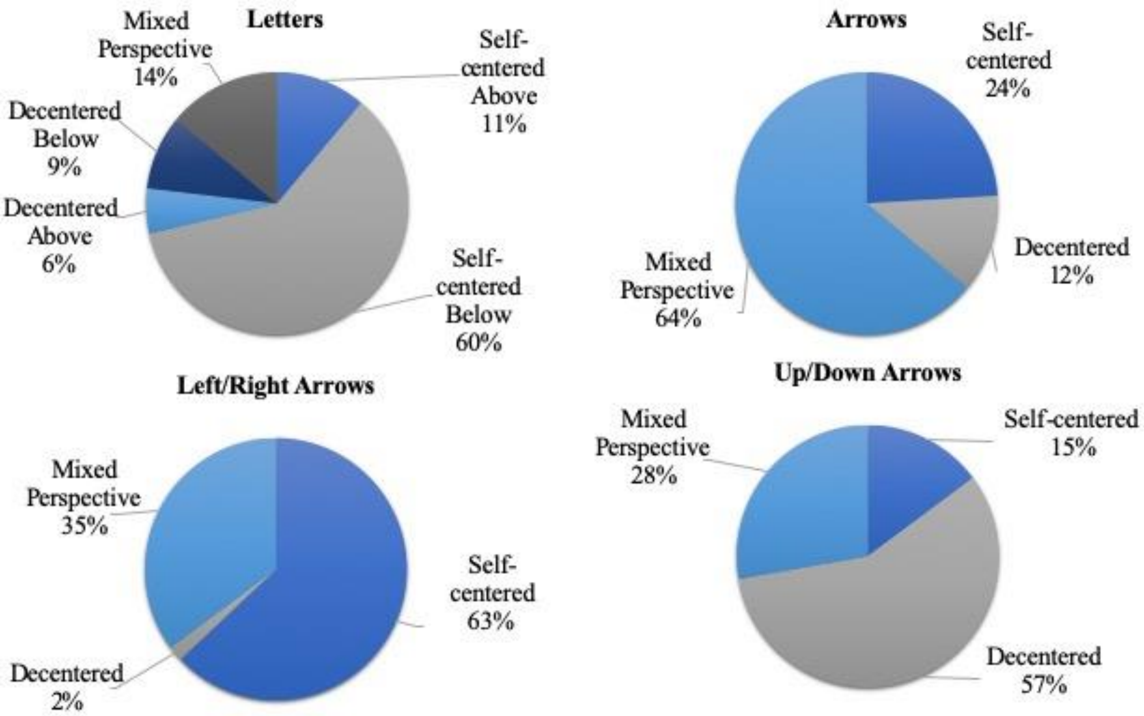


Figure 5: The pie charts at the top display the percentages for each adopted perspective when participants observed the letters and arrows with the BrainPort. The pie charts at the bottom display the different percentages adopted for the arrowhead stimuli.

Next, we examined the level of accuracy for the different types of stimuli (proportion of correct responses according to the BrainPort; e.g. if a ‘p’ is shown to the camera, participant responsed with a ‘p’). Median proportions can be found in Table 1. Wilcoxon Signed Rank tests (with a Bonferroni correction, giving an accepted *P* value of .008) showed that the accuracy in interpreting the letters was significantly higher than that of rotated letters ($Z = 6.28$, $p < .001$); letters less than arrows ($Z = 3.09$, $p = .002$); letters less than lines ($Z = 6.91$, $p < .001$); arrows more than rotated letters ($Z = 7.51$, $p < .001$); arrows less than lines ($Z = 5.61$, $p < .001$); and, lines more than rotated letters ($Z = 8.91$, $p < .001$). Also within the arrowhead category the left/right arrows were correctly identified significantly more than up/down arrows (75% and 66%, respectively) across all trials ($Z = 2.34$, $p = .02$).

The level of extroversion and openness was used as a predictors in a multiple linear regression analysis to examine the effect of these personality traits on participants’ proportion of correct responses (according to the BrainPort), which serves as a proxy for perspective taking (if a person were to answer correctly 100% of the time, they would likely be defaulting to a self-centred perspective). Proportion of correct answers was chosen to be the criterion as it is measured as continuous, rather than categorical, such as adopted perspective. We predicted that extroverted and open individuals would be more likely to adopt a decentered position, and would therefore offer more incorrect answers. Collinearity was tested on the predictors, and openness and extraversion proved to be within accepted values (tolerance = .96, VIF = 1.04; tolerance = .96, VIF = 1.04 respectively; [14, p. 325]). Visual examination of a P-P plot and a scatterplot of the standardised vs. predicted residuals showed no cause for concern. Additionally, autocorrelation was deemed at an acceptable level (Durbin-Watson =1.83). The results of the multiple regression analysis showed no effect of the combination of openness and extraversion on perspective-taking, $F(2,40) = 2.74$, $p = .077$, with an R^2 of .13 and an $R^2_{Adjusted} = .08$. However, examining the predictors individually showed that extroversion did not predict perspective-taking, but openness may do ($ß = -.36$, $t(40) = -2.32$, $p = .026$).

DISCUSSION

The presented study aimed to identify the dominantly adopted perspective when using the tongue via the BrainPort to interpret ambiguous stimuli. Additionally, it sought to examine whether camera orientation (a spatial factor), trait extroversion, and openness (personal factors) had any effect on the adopted perspective. The results indicate that self-centered was the most adopted perspective, and that camera orientation did not have any effect on the adopted perspective. Specifically, for the letter stimuli, slightly more than half the participants (60%) generally took a self-centered from below perspective, as if one were looking up at the tactile letter on tongue, from inside their mouth; and just over a tenth (11%) took a self-centered from above perspective, as if looking down at the tactile letter on the tongue from their eyes. Openness (but not extraversion) may slightly predict the adoption of a decentered perspective. Although, the multiple regression equation was marginally non-significant when factoring in both openness and extraversion.

The results do align reasonably well with Arnold et al. (2016), in that the majority of individuals adopt a self-centered perspective when perceiving the ambiguous letters. However, in later work by the same authors, they commented on the potential for perspective-taking to be predominated by a vision-centric point of view when perceiving tactile stimuli [4]. This does not appear to be the case with the BrainPort, with 60% of participants taking a view from below, as if from inside the body, rather than from the eyes per se.

The observed effect of openness on likeliness to adopted a decentered perspective aligns reasonably well with the work of Shelton and colleagues [30], as they found those with good social skills, more freely adopt a decentered perspective. However, finding that gender did not offer a tangible association with perspective-taking is somewhat surprising. While both males and females can be considered to predominantly adopt a self-centered perspective, the females did so more consistently. One study examining the gender effect on perspective-taking found that females perform better at spatial tasks with a social component [32]. While our task did not include a social component, we expected females to more readily adopt a decentered perspective as a reflection of their social relatability. Future replications of this research with tongue displays may wish to try running the experiment both blindfolded (as presented here), and unblindfolded with the addition of a social agent. It may be that females predominantly change their perspective only in the apparent presence of another. It was also surprising to find that camera orientation did not exhibit an association with perspective. The results of Brown and others [11] showed that the camera position on SSDs could have a dramatic sway over task performance. One possibility for our result is that using a fake camera (to control image presentation for each participant) did not offer the same proprioceptive feedback that an actual camera would.

As the BrainPort does not allow for a secondary camera to be connected (being designed as a standalone unit), it was not feasible to use a live camera in the present study. Perhaps in the future, a programmed accelerometer could be connected to the fake camera to wobble the stimuli on a screen, to more effectively emulate an actual handheld camera. The lack of ‘hackability’ in the BrainPort is surely a limitation set by the device for generalizability in research settings, but does not limit the knowledge generated about tongue displays (i.e. research with the BrainPort is useful for furthering tongue display development, despite customization issues). One possible way to overcome this could be found in the ‘Tongueduino’, a fully-programmable, lower resolution tongue interface [15], that offers a lower cost alternative to the BrainPort. Although, in the present case, the Tongueduino’s resolution would be incapable for presenting complex shapes, such as letters; hopefully, further BrainPort research can improve the functionality of lower cost and hackable tongue displays.

In previous work, there have been results that suggest that sighted people more naturally tend toward a head-centered perspective. The unique interface of the tongue display allows us to tease apart ‘head-centered’; indeed, the BrainPort allows for two head-centered perspectives (from the eyes down, and the tongue up). Arnold and colleagues (2017) suggest that one possible reason for this is that sighted individuals have a high-reliance on vision, due to the wide bandwidth of information conveyance, and therefore other types of perception also gravitate toward the eyes. They also found that when participants were forced to adopt a different perspective, their tactile perception accuracy significantly reduced. Our findings suggest that this link may exist, but in a fine form; the majority of our participants took a tongue-centered perspective (which is incredibly novel without prior BrainPort experience) and responded with generally high accuracy toward stimuli (around 75% correct responses). Arguably we cannot conclusively remark whether the BrainPort forces unnatural perspectives, as camera orientation did not display any effect on perspective adoption, but grounds for speculation surely exist. Additionally, as the experiment was short (to measure intuition rather than learning), there was little chance of brain plasticity changes, to adapt to the specific interpretation of the BrainPort; there could be a scenario where all the perspectives where ‘unnatural’.

One of the more intriguing findings is the difference in perspective between the left/right arrows and the up/down arrows. For the left/right arrows, the majority of participants (63%) took a self-centered perspective and on average were more often correctly identified than in the up/down arrow trials (75% compared to 66%). There was also a reverse to the majority adopting a decentered position in the up/down arrows (57%). This potentially could support the idea, that forcing unnatural perspectives reduces the accuracy of stimulus interpretation, as suggested by Arnold and colleagues (2017). The BrainPort could unnaturally flip up and down arrows in terms of perspective taken, while left and right arrows remain the same when translated onto the tongue (left still points left, but up points to behind the person).

The fact the perspective-taking is not uniform, even within a group that was given a small amount of training, combined with evidence from Arnold et al. [4, 5], that adopting an unnatural perspective detracts from tactile symbol recognition, would strongly suggest that making SSDs as customisable as possible, would be an advantageous boon. Additionally, Wicab, the company that designs and builds

BrainPorts, state on their website that training is required for the BrainPort, which is undertaken at dedicated training centers. We propose to Wicab that prior to their training program, they collect perspective-taking data from the user, and use that to calculate which orientation is naturally preferred; setting up the device in this way may decrease learning time. The ambiguous letters are ideal for such an exercise; revealing the dominant natural perspective and calibrating the BrainPort to match.

It was somewhat surprising that such a high percentage of participants favored the 'self-centered from below' perspective. While the BrainPort training was short, it is perhaps likely that the participants learned quickly the orientation to which the BrainPort flipped images. This is supported by the high percentage of correct answers, independent of perspective adoption (demonstrated by the arrowheads: left/right stimuli were correctly answered 75% of the time and were predominantly self-centered; up/down were correctly answered on 66% of trials, but the majority of responses were decentered). This would suggest that the demonstrated differences in perspective-taking are conservative and that with no training, a more considerable variation should be expected. However, for research purposes, basic training is challenging to remove from a protocol; without any training, responses would likely be too inaccurate to draw any conclusions from the data. Training people with visual impairments who intend to use the device in daily life would not suffer this issue. Additionally, while the present paper examined the specifics of perspective taking, there are many other avenues that should be explored to help inform potential BrainPort daily users, including, hygiene, acceptance, and wearability.

Impact

The impact of the presented study has the potential to dramatically improve the time it takes to gain familiarity with novel displays (like the BrainPort). If interface designers consider from the first stage, the possible individual differences between users, they could increase the percentage of people that find the device initially intuitive. For example, if the designers of the BrainPort (Wicab) could update it, to allow for a greater degree of flexibility as stated here, many potential BrainPort users would likely enjoy gains of the device earlier on in the learning process. Adding further adaptability to any computer interface has untold benefits, and improving the BrainPort will allow for deeper exploration into the cognitive sides of these devices. Using sighted participants in the present experiment has provided a suitable baseline from which to launch a similar protocol with blind persons.

Additionally, to the broader HCI community, our research highlights the general importance of customizable interfaces. Differences in interaction preferences can occur in even the most specific and novel technologies, like the BrainPort. As technologies become more specific, and as tactile methods are introduced into the wider computing industry (for example, feedback in gaming or engineering), it is vital for designers to allow for individual differences by incorporating flexibility in the use of the device.

CONCLUSION

The research of sensory substitution has much to offer the field of HCI, in the form of maximizing information transfer through non-visual displays. Our research into perspective-taking, using the tongue as an interface receptor, shows that when considering tactile displays, it is crucial to strive for the most customizable displays as possible. Factors that contribute to making a device as intuitive as possible can range through personal, interpersonal, and spatial; we tested gender, trait openness and extraversion, and camera orientation. We saw that openness may have played a small role in influencing the adopted perspective, but not to a sufficient enough degree to explain the observed variation within the sample. Making devices highly customizable would allow for individual differences within a user population, regardless of influencing factors. Specifically, regarding the BrainPort, a simple software update could improve the accessibility for users, particularly in the initial stages of acquiring the device.

ACKNOWLEDGMENTS

Authors Richardson, Lloyd-Esenkaya, Petrini, and Proulx would like to thank Cicely Hayes, Meg Carter, Rhea Dcosta, Georgie Waterworth, Lydia Maurel, and Daniella Gregory, for their effort and hard work in collecting the data.z`

REFERENCES

[1] Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S. and Amedi, A. 2014. EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. *Restorative Neurology and Neuroscience*. 32, 2 (Jan. 2014), 247–257. DOI:https://doi.org/10.3233/RNN-130338.

[2] Abdullah, M., Kim, M., Hassan, W., Kuroda, Y. and Jeon, S. 2017. HapticDrone: An Encountered-Type Kinesthetic Haptic Interface with Controllable Force Feedback: Initial Example for 1D Haptic Feedbac. *Adjunct Publication of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17* (New York, New York, USA, 2017),115–117. DOI:https://doi.org/10.1145/3131785.3131821

[3] Amedi, A., Stern, W.M., Camprodon, J.A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P. and Pascual-Leone, A. 2007. Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*. 10, 6 (Jun. 2007), 687–689. DOI:https://doi.org/10.1038/nn1912.

[4] Arnold, G., Spence, C. and Auvray, M. 2017. A unity of the self or a multiplicity of locations? How the graphesthesia task sheds light on the role of spatial perspectives in bodily self-consciousness.

Consciousness and Cognition. 56, (Nov. 2017), 100–114. DOI:https://doi.org/10.1016/J.CONCOG.2017.06.012.

[5] Arnold, G., Spence, C. and Auvray, M. 2016. Taking someone else's spatial perspective: Natural stance or effortful decentring? *Cognition*. 148, (Mar. 2016), 27–33. DOI:https://doi.org/10.1016/J.COGNITION.2015.12.006.

[6] Bach-y-rita, P. 2004. Tactile Sensory Substitution Studies. *Annals of the New York Academy of Sciences*. 1013, 1 (2004), 83–91. DOI:https://doi.org/10.1196/annals.1305.006.

[7] Bach-y-Rita, P. and Kercel, S. 2003. Sensory substitution and the human–machine interface. *Trends in Cognitive Sciences*. 7, 12 (Dec. 2003), 541–546. DOI:https://doi.org/10.1016/J.TICS.2003.10.013.

[8] Bertossa, F., Besa, M., Ferrari, R. and Ferri, F. 2008. Point Zero: A Phenomenological Inquiry into the Seat of Consciousness. *Perceptual and Motor Skills*. 107, 2 (Oct. 2008), 323–335. DOI:https://doi.org/10.2466/pms.107.2.323-335.

[9] Birk, M. and Mandryk, R.L. 2013. Control your game-self. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (New York, New York, USA, 2013), 685. DOI:https://doi.org/10.1145/2470654.2470752

[10] Van Boven, R.W. and Johnson, K.O. 1994. The limit of tactile spatial resolution in humans: grating orientation discrimination at the lip, tongue, and finger. *Neurology*. 44, 12 (Dec. 1994), 2361–6. DOI:https://doi.org/10.1212/wnl.44.12.2361.

[11] Brown, D., Macpherson, T. and Ward, J. 2011. Seeing with Sound? Exploring Different Characteristics of a Visual-to-Auditory Sensory Substitution Device. *Perception*. 40, 9 (Sep. 2011), 1120–1135. DOI:https://doi.org/10.1068/p6952.

[12] Carton, A. and Dunne, L.E. 2013. Tactile distance feedback for firefighters. *Proceedings of the 4th Augmented Human International Conference on - AH '13* (New York, New York, USA, 2013), 58–64. DOI:http://dx.doi.org/10.1145/2459236.2459247

[13] Chebat, D.-R., Rainville, C., Kupers, R. and Ptito, M. 2007. Tactile–'visual' acuity of the tongue in early blind individuals. *NeuroReport*. 18, 18 (Dec. 2007), 1901–1904. DOI:https://doi.org/10.1097/WNR.0b013e3282f2a63.

[14] Danilov, Y. and Tyler, M. 2005. BrainPort: An Alternative Input to the Brain. *Journal of Integrative Neuroscience*. 04, 04 (Dec. 2005), 537–550. DOI:https://doi.org/10.1142/S0219635205000914.

[15] Dublon, G. and Paradiso, J.A. 2012. Tongueduino. (2012), 1453. DOI: https://doi.org/10.1145/2212776.2212482

[16] Eysenck, H.J. 1961. Personality and Social Attitudes. *The Journal of Social Psychology*. 53, 2 (Apr. 1961), 243–248. DOI:https://doi.org/10.1080/00224545.1961.9922122.

[17] Ferrè, E.R., Lopez, C. and Haggard, P. 2014. Anchoring the Self to the Body. *Psychological Science*. 25, 11 (Nov. 2014), 2106–2108. DOI:https://doi.org/10.1177/0956797614547917.

[18] Field, A. 2013. *Discovering Statistics using IBM SPSS Statistics*. Sage.

[19] Funk, M. 2018. Human-drone interaction. *Interactions*. 25, 3 (Apr. 2018), 78–81. DOI:https://doi.org/10.1145/3194317.

[20] Grant, P., Spencer, L., Arnoldussen, A., Hogle, R., Nau, A., Szlyk, J., Nussdorf, J., Fletcher, D.C., Gordon, K. and Seiple, W. 2016. The Functional Performance of the BrainPort V100 Device in Persons who Are Profoundly Blind. *Journal of Visual Impairment & Blindness*. 110, 2 (Mar. 2016), 77–88. DOI:https://doi.org/10.1177/0145482X1611000202.

[21] Hoy, M.B. 2018. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*. 37, 1 (Jan. 2018), 81–88. DOI:https://doi.org/10.1080/02763869.2018.1404391.

[22] John, O.P., Donahue, E.M. and Kentle, R.L. 1991. The big five inventory—versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality.

[23] McCrae, R.R. 1996. Social consequences of experiential openness. *Psychological Bulletin*. 120, 3 (1996), 323–337. DOI:https://doi.org/10.1037/0033-2909.120.3.323.

[24] Nau, A., Bach, M. and Fisher, C. 2013. Clinical Tests of Ultra-Low Vision Used to Evaluate Rudimentary Visual Perceptions Enabled by the BrainPort Vision Device. *Translational Vision Science & Technology*. 2, 3 (Jun. 2013), 1. DOI:https://doi.org/10.1167/tvst.2.3.1.

[25] Parsons, L.M. and Shimojo, S. 1987. Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *Journal of Experimental Psychology: Human Perception and Performance*. 13, 3 (1987), 488–504. DOI:https://doi.org/10.1037/0096-1523.13.3.488.

[26] Richardson, M., Thar, J., Borchers, J., Ward, J. and Hamilton-Fletcher, G. 2019. How much spatial information is lost in the sensory substitution process? Comparing visual, tactile, and auditory approaches. *Perception*. (2019). DOI: https://doi.org/10.1177/0301006619873194

[27] Sampaio, E., Maris, S. and Bach-y-Rita, P. 2001.

Brain plasticity: ‘visual’ acuity of blind persons via the tongue. *Brain Research*. 908, 2 (Jul. 2001), 204–207. DOI:https://doi.org/10.1016/S0006-8993(01)02667-1.

[28] Schober, M.F. 1993. Spatial perspective-taking in conversation. *Cognition*. 47, 1 (Jan. 1993), 1–24. DOI:https://doi.org/10.1016/0010-0277(93)90060-9.

[29] Sekiyama, K. 1991. Importance of head axes in perception of cutaneous patterns drawn on vertical body surfaces. *Perception & Psychophysics*. 49, 5 (Sep. 1991), 481–492. DOI:https://doi.org/10.3758/BF03212182.

[30] Shelton, A.L., Clements-Stephens, A.M., Lam, W.Y., Pak, D.M. and Murray, A.J. 2012. Should social savvy equal good spatial skills? The interaction of social skills with spatial perspective taking. *Journal of Experimental Psychology: General*. 141, 2 (2012), 199–205. DOI:https://doi.org/10.1037/a0024617.

[31] Stronks, H.C., Mitchell, E.B., Nau, A.C. and Barnes, N. 2016. Visual task performance in the blind with the BrainPort V100 Vision Aid. *Expert Review of Medical Devices*. 13, 10 (Oct. 2016), 919–931. DOI:https://doi.org/10.1080/17434440.2016.1237287.

[32] Tarampi, M.R., Heydari, N. and Hegarty, M. 2016. A Tale of Two Types of Perspective Taking. *Psychological Science*. 27, 11 (Nov. 2016), 1507–1516. DOI:https://doi.org/10.1177/0956797616667459.

[33] Ward, J. and Meijer, P. 2010. Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and Cognition*. 19, 1 (Mar. 2010), 492–500. DOI:https://doi.org/10.1016/j.concog.2009.10.006.