NEUROSCIENCE
RESEARCH ARTICLE

M. V. Perrotta et al. / Neuroscience 458 (2021) 77–86

Deciphering Sounds Through Patterns of Vibration on the Skin

Michael V. Perrotta, [a] **Thorhildur Asgeirsdottir** [a] **and David M. Eagleman** [a,b]*

[a] *Neosensory, 4 West 4th Street, Suite 301, San Mateo, CA 94402, USA*

[b] *Department of Psychiatry and Behavioral Sciences, Stanford University, 401 Quarry Road, Stanford, CA 94304, USA*

Abstract—Sensory substitution refers to the concept of feeding information to the brain via an atypical sensory pathway. We here examined the degree to which participants (deaf and hard of hearing) can learn to identify sounds that are algorithmically translated into spatiotemporal patterns of vibration on the skin of the wrist. In a three-alternative forced choice task, participants could determine the identity of up to 95% and on average 70% of the stimuli simply by the spatial pattern of vibrations on the skin. Performance improved significantly over the course of 1 month. Younger participants tended to score better, possibly because of higher brain plasticity, more sensitive skin, or better skills at playing digital games. Similar results were obtained with pattern discrimination, in which a pattern representing the sound of one word was presented to the skin, followed by that of a second word. Participants answered whether the word was the same or different. With minimal difference pairs (distinguished by only one phoneme, such as ''house'' and ''mouse''), the best performance was 83% (average of 62%), while with non-minimal pairs (such as ''house'' and ''zip'') the best performance was 100% (average of 70%). Collectively, these results demonstrate that participants are capable of using the channel of the skin to interpret auditory stimuli, opening the way for low-cost, wearable sensory substitution for the deaf and hard of hearing communities.

Key words: sensory substitution, haptics, sound identification, hearing loss, deafness.

INTRODUCTION

Sensory substitution, the method of passing sensory information through a different sensory pathway, has been documented for decades (Eagleman, 2020). Typically, this involves translating visual information from a video feed into touch on the skin (Bach-y-Rita et al., 1969), into touch on the tongue (Danilov and Tyler, 2005; Grant et al., 2016), or into a soundscape (Amedi et al., 2007; Auvray et al., 2007; Ward and Meijer, 2010; Striem-Amit et al., 2012). With such approaches, blind users can learn to identify visual objects and scenes.

Sensory substitution devices have also been created and tested for different modalities, including balance-to-touch to mitigate balance disorders (Tyler et al., 2003) and touch-to-touch to give somatosensation to someone with nerve damage (e.g. from leprosy; Bach-y-Rita, 1999) or who uses a prosthetic limb (Riso, 1999).

Sound-to-touch sensory substitution devices have also been developed to aid people with hearing loss and deafness (Weisenberger and Miller, 1987; Weisenberger et al., 1987, 1991; Bernstein et al., 1989; Bernsteinet al., 1991; Weisenberger and Russell, 1989; Eberhardt et al., 1990; Weisenberger and Kozma-Spytek, 1991; Weisenberger and Percy, 1995; Auer et al., 1998; Reed and Delhorne, 2003; Fletcher et al., 2019); however, previous versions have been limited by size and computational speed.

Such devices can help people with sensory impairments by letting them access the sensory information they are otherwise unable to. Although sensory substitution has been experimentally fruitful, it has suffered from lack of practicality due to size, expense, or inconvenience of wearing (Bach-y-Rita, 1983; Maidenbaum et al., 2014).

In that light, we have developed a practical, convenient device for sensory substitution: a wristband consisting of four vibratory motors, a microphone to capture sound, and a sophisticated processing system to convert audio in real time to spatiotemporal patterns of vibration; more in Methods below and in Novich and Eagleman (2015). With this approach, participants with deafness or hearing loss can learn to identify sound categories through the vibrations alone.

We here test the efficacy of this device as a sound-to-touch sensory substitution device. We set out to quantify whether such a device, worn on the wrist, and containing just four motors, would be able to convey enough information to a user that they could learn to

*Correspondence to: D. M. Eagleman, Department of Psychiatry and Behavioral Sciences, Stanford University, 401 Quarry Road, Stanford, CA 94304, USA.
E-mail address: davideagleman@stanford.edu (D. M. Eagleman).

https://doi.org/10.1016/j.neuroscience.2021.01.008
0306-4522/

differentiate similar patterns of vibration and identify sound categories from the vibrations.

Knowing that improved performance with a sensory substitution device presumably relies on neuroplastic changes (Rauschecker, 1995; Bach-y-Rita and Kercel, 2003; Bach-Y-Rita, 2004; Bubi et al., 2010; Proulx et al., 2014; Eagleman, 2020) and presuming that real-world use of the device should be useful here, we tested deaf and hard of hearing participants over the course of a month and had them take home the wristband during this time, allowing them to use it however they wished in their daily life.

We here analyze the performance of our participants, assessing whether users can identify sounds and differentiate similar patterns of vibration.

EXPERIMENTAL PROCEDURES

Participants

We tested 18 participants (eight male) ranging in age from 23 to 84 (median = 59). No participant had any prior experience with the wristband. All participants had severe or profound hearing loss. Three additional participants dropped out before the end of the study; their data is not included in this analysis. Demographic data of the 18 participants is in Table 1.

Wristband

Device. The Neosensory Buzz wristband consists of four vibratory motors built into the strap of the wristband (Fig. 1A). The motors are linear resonant actuators (LRAs), vibrating at 175 Hz in a sine wave (confirmed by measuring the fundamental and harmonic frequencies at a variety of vibration intensities), capable of rising from 0 to 50% of their amplitude within 30 ms. We set a minimum intensity to avoid low voltages which might result in an inconsistent vibration. The motor amplitude can be controlled with an eight bit resolution, meaning the motors can be controlled at 256 different amplitudes. At the highest amplitude, each motor vibrates at 1.7 G_{RMS} (16.6 m/s^2). The motors are separated from one another at a distance of 18.2 mm and 19.2 mm for the small and large wristband sizes, respectively (center-to-center distances). Each motor pad contacts the wearer's skin on a rectangular area that measures 8.2 mm by 8.5 mm.

The top of the wristband is a module that contains the power button, user setting buttons, a microphone, and a microcontroller. The microphone captures audio in 16 ms chunks and sends each 16 ms chunk of audio to the microcontroller. The microcontroller processes the audio data through Neosensory's sound-to-touch algorithm and vibrates the motors according to the output of the algorithm.

Similar approaches of translating sound information into vibrations have been researched previously. For instance, Fletcher et al. (2019) found that vibration signals can increase speech recognition when used in conjunction with cochlear implants. While Fletcher et al. represented sound frequencies via changing vibration frequencies of their vibrating motors (from 50 to

Table 1. Demographic data. Hearing loss age indicates the age in years at which the participant started to lose hearing (or 0 if the participant was born deaf or with hearing loss). Lip reading self-rating indicates the level of proficiency in reading lips the participant reports, with 1 being "extremely bad" and 7 being "extremely good". ASL usage indicates the amount the participant reports using American Sign Language (ASL) with friends, family, and colleagues, where 1 is "none at all" and 5 is "a great deal". Tech indicates if the participant currently uses hearing aids (HA), cochlear implants (CI), or neither (None). Hearing loss values are decibels of hearing loss at six pure tones in the left and the right ears. Hearing loss values are measured without cochlear implants or hearing aids. Note that 90 dB of hearing loss is the most amount of hearing loss the test can detect and indicates 90 dB or more of hearing loss

	Age	Hearing loss age	Lip reading self-rating	ASL usage	Tech	Hearing loss (dB)											
						250 Hz		500 Hz		1000 Hz		2000 Hz		4000 Hz		8000 Hz	
						L	R	L	R	L	R	L	R	L	R	L	R
B1	23	4	4	5	None	90	90	90	90	90	90	90	90	90	90	90	90
B2	32	12	6	5	None	90	90	90	90	90	90	90	90	90	90	90	90
B3	35	3	5	1	HA	90	53	90	60	90	59	90	52	90	60	90	90
B4	39	0	7	5	None	90	90	90	90	90	90	90	90	90	90	90	90
B5	42	0	5	5	HA	90	90	90	90	90	90	90	90	90	90	90	90
B6	53	14	5	1	HA	27	18	34	36	41	39	63	37	90	34	90	90
B7	54	45	5	2	HA	14	10	21	20	33	42	63	57	48	42	35	33
B8	56	32	6	1	None	90	90	90	90	90	90	90	90	90	90	90	90
B9	58	6	7	2	HA	90	90	90	90	90	90	90	90	90	90	90	90
B10	59	40	6	1	HA	90	90	90	40	90	90	90	90	90	90	90	90
B11	59	5	1	3	CI	90	90	90	90	90	90	90	90	90	90	90	90
B12	59	0	6	5	None	90	90	90	90	90	90	90	90	90	90	90	90
B13	62	4	7	5	HA	90	90	90	90	90	90	90	90	90	90	90	90
B14	62	57	5	1	None	12	4	17	14	24	24	45	38	47	41	29	18
B15	66	3	1	4	None	90	90	90	90	90	90	90	90	90	90	90	90
B16	69	24	6	2	HA	90	90	90	90	90	71	90	90	90	90	90	90
B17	80	17	7	3	CI	90	90	90	90	90	90	90	90	90	90	90	90
B18	84	75	4	1	HA	22	22	38	18	46	30	48	38	90	90	90	90

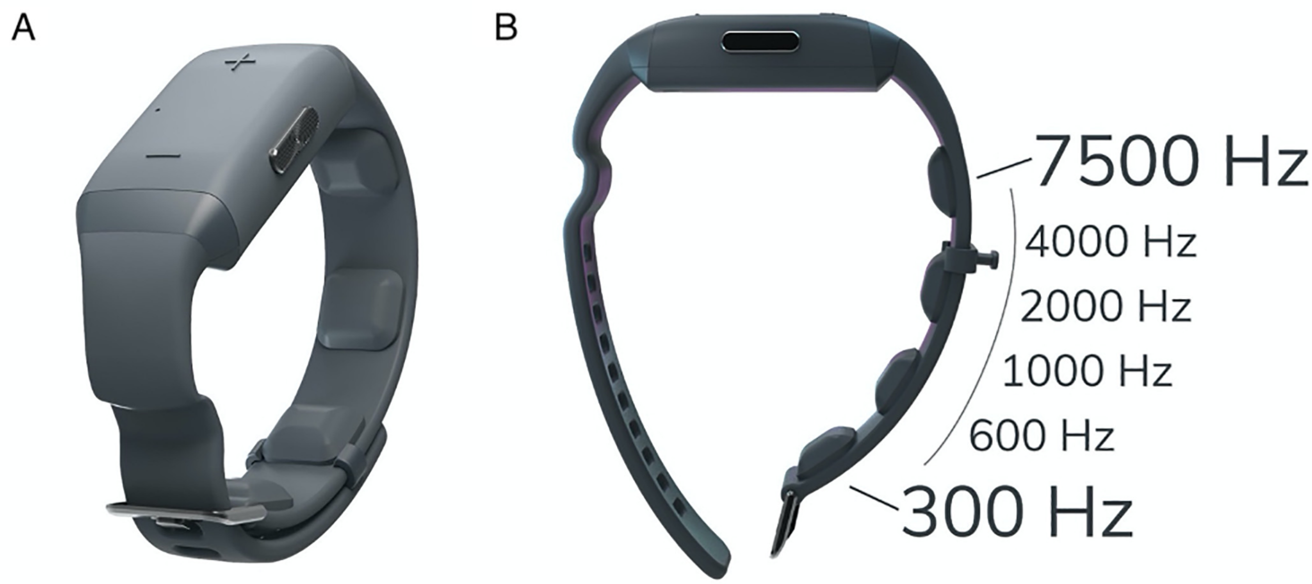

Fig. 1. The Neosensory Buzz. **(A)** A wristband with four vibratory motors built into the strap. **(B)** The sound captured by the Neosensory Buzz is played out through vibrating motors. A haptic illusion is used such that the four motors are able to represent a continuous frequency space. The haptic illusion is such that, when two adjacent motors are turned on, an illusory point between the two motors is felt. The perceived location of the vibration represents sound between 300 and 7500 Hz. The frequency space is scaled logarithmically.

230 Hz), the motors used in this present study vibrate at a constant frequency while we represent frequency information from the audio across the physical space on the wrist. Our motors vibrate only at 175 Hz, which is the vibration frequency to which skin has the lowest threshold (i.e. highest sensitivity; Verrillo, 1980).

Algorithm. The algorithm is frequency-based, using a discrete Fourier transform to analyze the amplitudes of different frequencies present in the sound in bins from 300 to 7500 Hz (Fig. 1B). Updated frequency amplitudes are calculated every 16 ms – a duration that is short enough to capture quick temporal differences in environmental sounds but long enough to allow the motors to reach their target vibration amplitude before the next frame. After calculating the frequency amplitudes, the algorithm chooses which frequency, if any, should be represented on the wristband (no more than one frequency is represented for any given 16 ms frame). The algorithm chooses the frequency with the greatest amplitude, so long as it is above its own running mean. That is, if a frequency bin contains the greatest amplitude in a given frame but has an amplitude less than that of its running mean, the next loudest frequency bin will be chosen by the algorithm instead. The running mean for any bin is calculated using exponential smoothing with a smoothing factor of 0.03, such that the running mean reaches ~63% of a new value in 33 frames (or 1.056 s). This results in the wristband not representing constant, loud hums, such as an AC unit. A frequency bin must also have an amplitude of at least 20 dB SPL.

The amplitude of the vibration is related to the amplitude of the chosen frequency bin. A minimum amplitude vibration corresponds to a frequency bin amplitude just above that frequency bin's running mean. A maximum vibration amplitude corresponds to the algorithm's current dynamic ceiling, which is a changing value based on the amplitude of recent sounds in any frequency bin. When a loud sound happens, the dynamic ceiling jumps up to the level of that loud sound and then gradually falls (if the sound is no longer as loud). By means of this dynamic ceiling, quiet sounds in a quiet environment can be felt with significant amplitude but the same quiet sounds in a loud environment will register as weak. This is comparable to dynamic range compression.

Any vibration amplitude between the minimum and maximum values is scaled on an exponential curve. This is to account for Weber's Law (Geldard, 1957), which dictates that two high-amplitude stimuli require a greater difference in amplitude than two low-amplitude stimuli for a participant to recognize a difference.

The algorithm then translates this frequency and amplitude into a motor output by mapping frequency onto one of 256 different spatial locations, accomplished by leveraging a well-known haptic illusion (Alles, 1970; Rahal et al., 2009; Luzhnica et al., 2017). Specifically, an illusory location is a point on the wrist that is felt by the wearer of the wristband even when a motor is not located directly at that point. We stimulate these illusion locations by turning on two motors, one on either side of the illusion location, at specific amplitudes such that the wearer feels as if a single point somewhere between the two motors is vibrating.

The range of frequencies is represented on the wristband in logarithmic space (see Fig. 1B). This gives more spatial resolution among the lower frequencies compared to the higher frequencies. This mapping was chosen because humans perceive frequencies on a logarithmic scale, such that a 200 Hz tone, a 400 Hz tone, an 800 Hz tone, and a 1600 Hz tone are perceived to be evenly spaced.

The algorithm is designed in this way with the aim of representing environmental sounds such that very different sounds feel very different and very similar sounds feel very similar.

Tests

Sessions. Each participant came into our laboratory three times: on day 0 (the day they first received their wristband), day 14, and day 28. During each session participants completed three tasks, in the following order: Minimum Threshold task, Pattern Discrimination task, and Sound Identification task. Participants were instructed to wear the wristband at a snug but comfortable tightness so that all motor pads contacted the skin. Once over the course of the study, participants also completed a demographic questionnaire and a hearing test. Between sessions, participants took the wristband home and were instructed to wear the wristband as often as they liked, as long as they wore it for at least four hours each day.

Note that no control condition is included in this study due to the impossibility of a placebo-effect.

Demographics. Participants completed a questionnaire about their sex, age, and hearing loss profile (e.g., when did you lose your hearing, how much of the time do you use ASL to communicate with friends and family).

Audiogram. We measured participants' hearing loss using the Mimi hearing test (apps.apple.com) and Beyerdynamic Aventho wired on-ear headphones. Data for each participant is in Table 1.

Sound identification. Participants performed a task to determine if they could identify sounds based on the patterns of vibrations on their wrist. In each trial, the wristband presented a pattern of vibrations that lasted between 1.6 and 7.2 s (see Supplementary Materials Table S1 for a list of all files and their durations). No audible sound was played to cause the vibrations; instead, the appropriate vibration pattern was constructed from audio recordings prior to the task and was transmitted to the wristband via Bluetooth. Then the participant chose from a list of three options (e.g., *dog bark, siren, running water*) the one they believed corresponded with the pattern. This list was visible to the participant as soon as they started to feel the vibrations. Once the participant answered, they were shown which answer was correct.

The patterns of vibration (motor activations) were created by passing audio recordings of different sounds through the sound-to-touch algorithm (Fig. 2). Sounds were chosen from a list of 14 sound categories: baby crying, car horn, car passing, clapping, clock alarm, coughing, dog bark, door knock, laughing, ringtone, running water, siren, smoke alarm, and speech. Each of the 14 sound categories had five sound recordings associated with it for a total of 70 audio recordings, each of which was translated into vibrational patterns. In each trial, one of the five patterns for the target sound category was chosen, such that the participant would not reencounter the exact same vibration pattern but instead was presented with similar but different vibration patterns for a given sound category.

Participants completed 56 trials of the Sound Identification task in each of the three sessions. The 14 sounds were evenly distributed across trials, such that each sound was the correct answer in exactly four trials each session. The two incorrect answers were selected at random from the remaining 13 sounds.

Pattern discrimination

Participants performed a task to determine whether they could distinguish different patterns of vibration on their

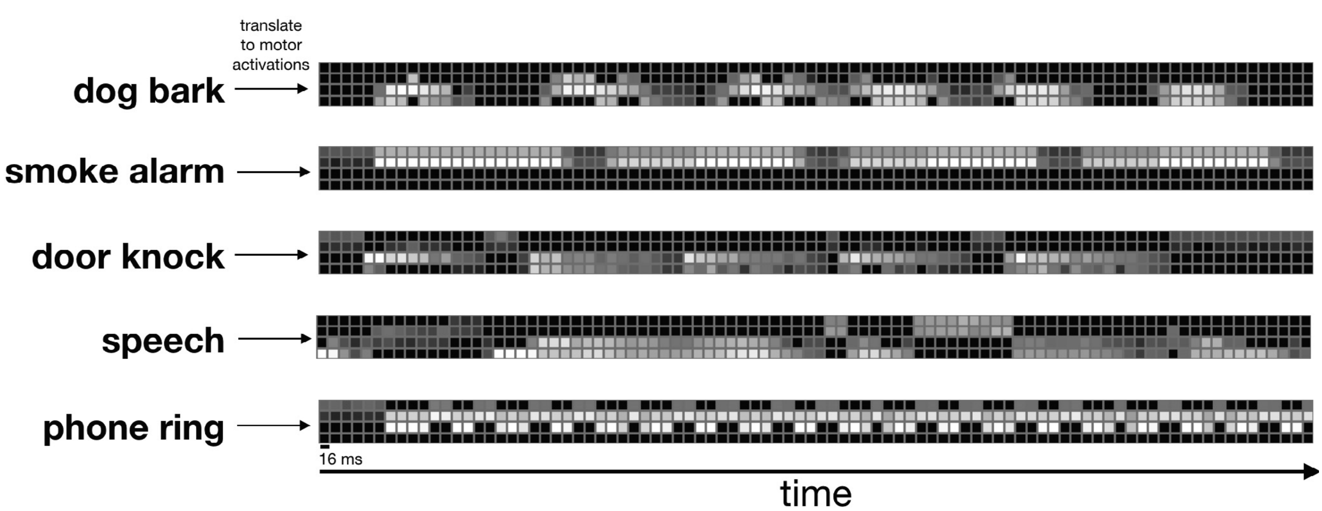

Fig. 2. Sound identification task. Sound files were algorithmically translated into patterns of motor activation. When two motors are active, an illusory position in between their locations is perceived. Five examples of 70 stimuli are shown.

wrist. In each trial, the participant was presented with two vibration patterns of ~1 s, with 300–600 ms of silence between vibration patterns. The participant then answered whether the two vibration patterns were identical or different. Feedback told the participant whether their answer was correct or incorrect.

The vibration patterns were created by passing audio recordings of spoken words through the sound-to-touch algorithm offline and capturing the motor outputs associated with each word's audio recording (Fig. 3).

For half of the trials, the two vibration patterns were identical (i.e. created from the same audio recording of a single word). For the other half, the two vibration patterns were created from audio recordings of two different words. Of these trials, half were "minimal pairs" and half were "non-minimal pairs". In a "minimal pair" trial, the two words used to create the vibration patterns differed from each other by only one phoneme (e.g., bees/cheese, rip/zip, bad/bed). For each minimal pair trial, one of 262 minimal pairs was randomly chosen. In non-minimal pair trials, the two words used to create the vibration patterns were chosen randomly from the 524 words that made up the 262 minimal pairs (e.g. bees/zip or rip/bed). See Supplementary Materials Tables S2 and S3 for a list of the words, the duration of each audio file for each word, and the list of minimal pairs of words.

It should be noted that, while words were used to create vibration patterns that are similar to one another, this task does not aim to test a participant's ability to understand speech. Instead, the task aims to determine a participant's ability to discriminate between two similar but different patterns. By using minimal and non-minimal pairs of words, we can straightforwardly test two levels of difficulty of pattern discrimination. However, it is important to note that the participants were not aware that the patterns they were feeling were words, as the patterns were delivered silently via Bluetooth. In other words, the participants knew nothing about the patterns they were discriminating, and this is because we were not testing whether this algorithm allowed participants to identify speech. Rather, we were simply quantifying the degree to which patterns could be discriminated, regardless of their origin.

Participants completed 60 trials of the Pattern Discrimination task in each of the three sessions.

Minimum threshold

Participants performed a task which determined the lowest detectable amplitude of vibration on their skin. In this task, participants felt pulses of vibrations at a single location on their wrist. The pulses were each 32 ms long and separated by 300 ms. The participant was instructed to press and hold a button as soon as s/he could feel any vibration, and to let go of the button as soon as the pulsing vibrations became too weak to feel. When the button was pressed, the vibrations became less intense; when the button was not pressed, the vibrations became more intense. Fig. 4 shows how the vibration amplitudes change as a participant indicates s/he can or cannot feel the vibrations. Each time the participant pressed or released the button, the amount of change between each amplitude decreased, giving the opportunity to more precisely indicate when vibrations were being felt. Once the participant pressed and released the button six times, the trial ended.

The minimum threshold for a given trial was calculated as the average of the amplitudes at which the participant pressed and released the button. These averages exclude the first three presses and releases and only include the last three presses and releases (as indicated by the red pluses in Fig. 4), since the rate of change of the amplitudes is lowest during the end of the trial and thus this reduces the error introduced by a participant being slow to react to their sensation.

We quantified the minimum threshold for five locations on each participant. Three of the locations were individual motors on the wristband while the other two were "illusion locations", created by vibrating two motors simultaneously (see Algorithm section above). A participant's final minimum threshold was calculated as the average of their minimum threshold at all five locations.

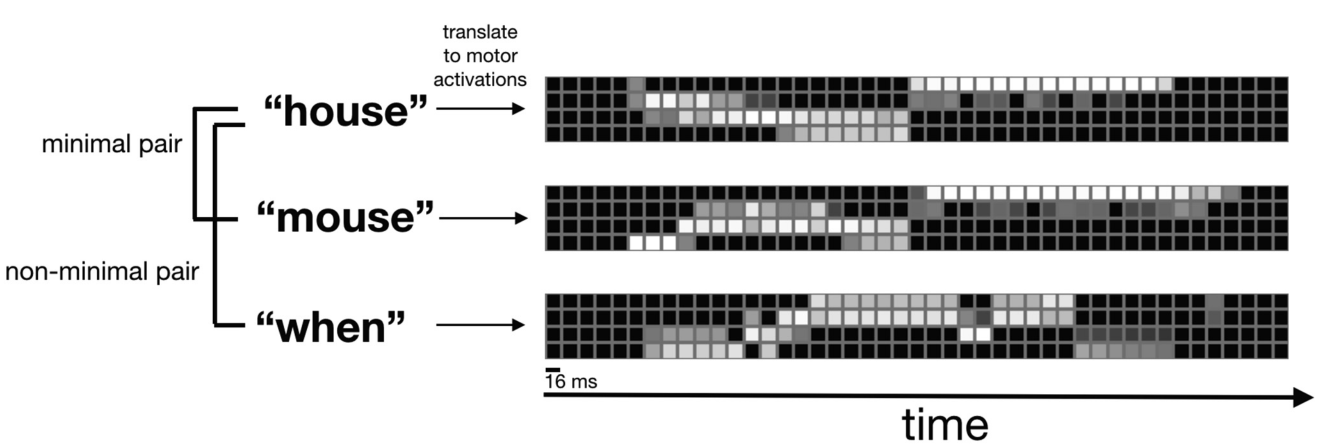

Fig. 3. Pattern discrimination task. Spoken words are converted to sequences of motor activations on the wristband. Three examples of 524 stimuli are shown.

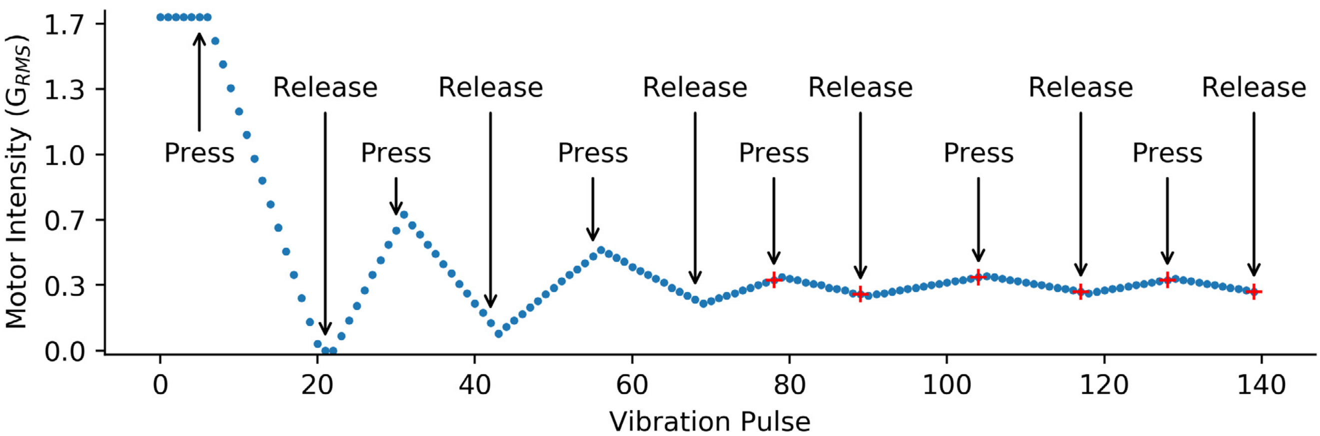

Fig. 4. Testing for minimum threshold – representative example. When presented with a train of pulsing vibrations, participants pressed and held a button when the vibrations were strong enough to feel and released the button when the vibrations were too weak to feel. The vibrations became weaker after the button was pressed and stronger when it was released. The rate of change of the vibration amplitude decreased throughout the trial. The six red crosses indicate the values averaged to determine a participant's minimum threshold for a single trial. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The minimum threshold amplitudes are measured in root-mean-square acceleration (G_{RMS}). At its maximum amplitude, the wristband vibrates at 1.7 G_{RMS}.

Institutional review board

The study protocol was approved by Solutions IRB, an independent institutional review board accredited by the Association for the Accreditation of Human Research Protection Programs, Inc. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

RESULTS

Sound identification

Remarkably, in their first session, 11 out of 18 of the participants were able to identify more than 62% of sounds simply by feeling these sounds as vibration patterns. The highest score a participant received in any of the three sessions was 94.6% (chance performance was 33.3%) (Fig. 5A).

Over the course of one month there was a significant improvement. We fit a logistic mixed model (estimated using ML and Nelder-Mead optimizer) to predict participant performance on a given trial with the

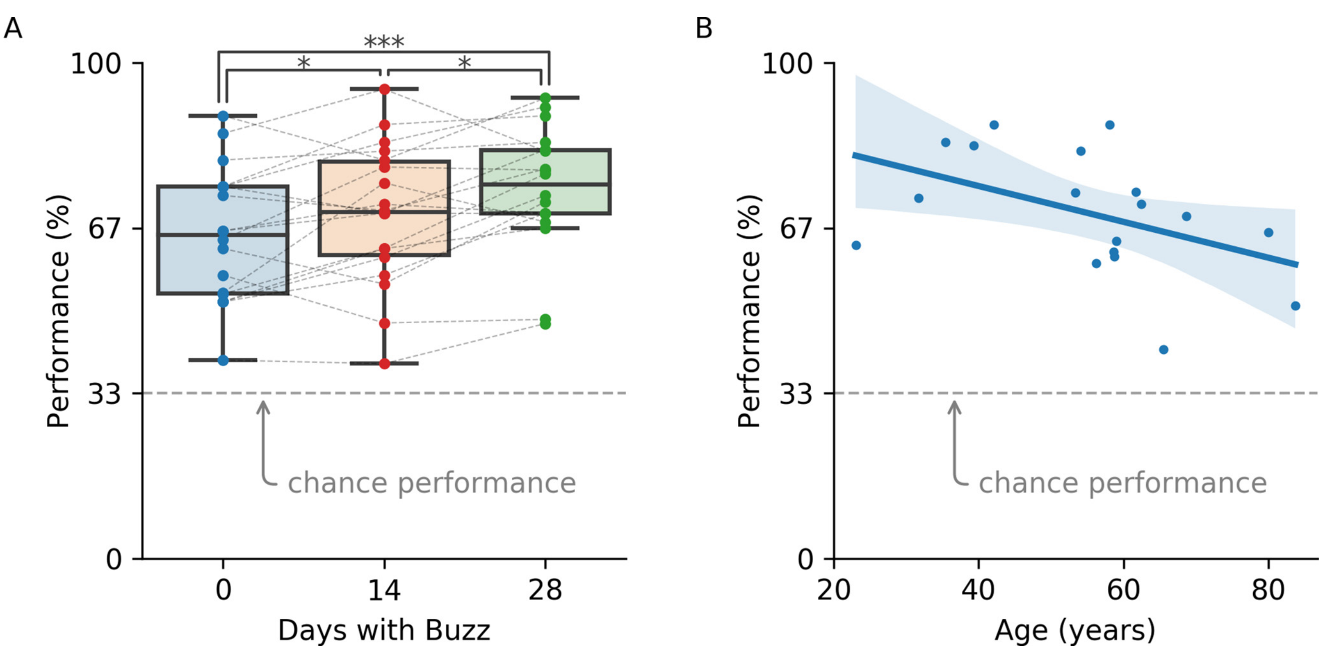

Fig. 5. Sound identification performance. **(A)** Participants scored significantly higher on day 28 than they did on day 0 ($p < 0.001$). Box plots use Tukey's original configuration (McGill et al., 1978) with outliers defined as falling more than 1.5 times the interquartile range outside of the box. **(B)** Participants who are older scored lower than their younger counterparts ($p = 0.028$). Performance averaged over the three sessions. Shaded area shows a 95% confidence interval for the linear regression.

participant's current session and age (formula = user_correct ~ session + age). The model included the participant's ID as a random effect (formula = ~1 | user_id). Standardized parameters were obtained by fitting the model on a standardized version of the dataset. Within this model, the session number correlates positively with performance (beta = 0.24, SE = 0.05, std. beta = 0.19, $p < 0.001$) and the effect of age correlates negatively (beta = −0.02, SE = 8.49e−03, std. beta = −0.28, $p < 0.05$).

We also fit a logistic mixed model to predict participant performance on a given trial with the participant's age at hearing loss onset. However, no significant effect was found.

Pattern discrimination

We found participants were able to discriminate between similar vibrational patterns played through the wristband, with one participant scoring an average of 87.8% on non-minimal pair trials (Fig. 6A). Being a two-alternative forced choice task, participants would be expected to score at 50% if they could not discriminate between patterns (i.e. chance). Participants performed significantly above chance for both the minimal and non-minimal pairs conditions (one-tailed, single-sample Student's *t*-test: non-minimal pairs $t(17) = 8.61$, $p < 0.001$; minimal-pairs $t(17) = 6.69$, $p < 0.001$). Participants scored significantly higher on non-minimal pair trials than they did on minimal-pair trials ($t(17) = 3.78$, $p < 0.001$). We did not measure a statistically significant improvement in score from the first session to the third; we therefore averaged the data over the three sessions (Day 0, 14, 28). Further studies will determine whether a longer term of use improves discrimination.

Because two different vibration patterns can differ in length within a single trial (the mean duration difference over all trials was 94 ms), we investigated participants' scores on "Different" trials where there was a 0 ms difference between the lengths of the two patterns. Of the 235 trials with a 0 ms duration difference, 131 trials were answered correctly. A one-tailed binomial test shows the probability of correctly answering a trial with no difference in duration is significantly greater than chance ($x = 131$, $n = 235$, $p = 0.0448$).

We also investigated the relationship between age and performance on the pattern discrimination task. We did not find a correlation of significance at the alpha value of 0.05 ($r(16) = -0.31$, $p = 0.075$; Fig. 6B); however, the data generally suggest that participants who are older are more likely to score lower.

Minimum threshold

When taking the average minimum threshold score over the three sessions, we did not find a significant correlation between minimum threshold and age ($r(16) = 0.29$, $p = 0.122$; one-tailed; Fig. 7). We also did not find a significant correlation between minimum threshold and performance on sound identification ($r(16) = -0.31$, $p = 0.11$; one-tailed) or pattern discrimination ($r(16) = -0.43$, $p = 0.04$; one-tailed).

DISCUSSION

The purpose of this study was to investigate a pragmatic sound-to-touch sensory substitution device. Our results show that deaf and hard of hearing participants are able to identify sound categories by feeling vibrations on their wrists, and in some cases can score nearly perfectly in a three-alternative forced choice task of sound

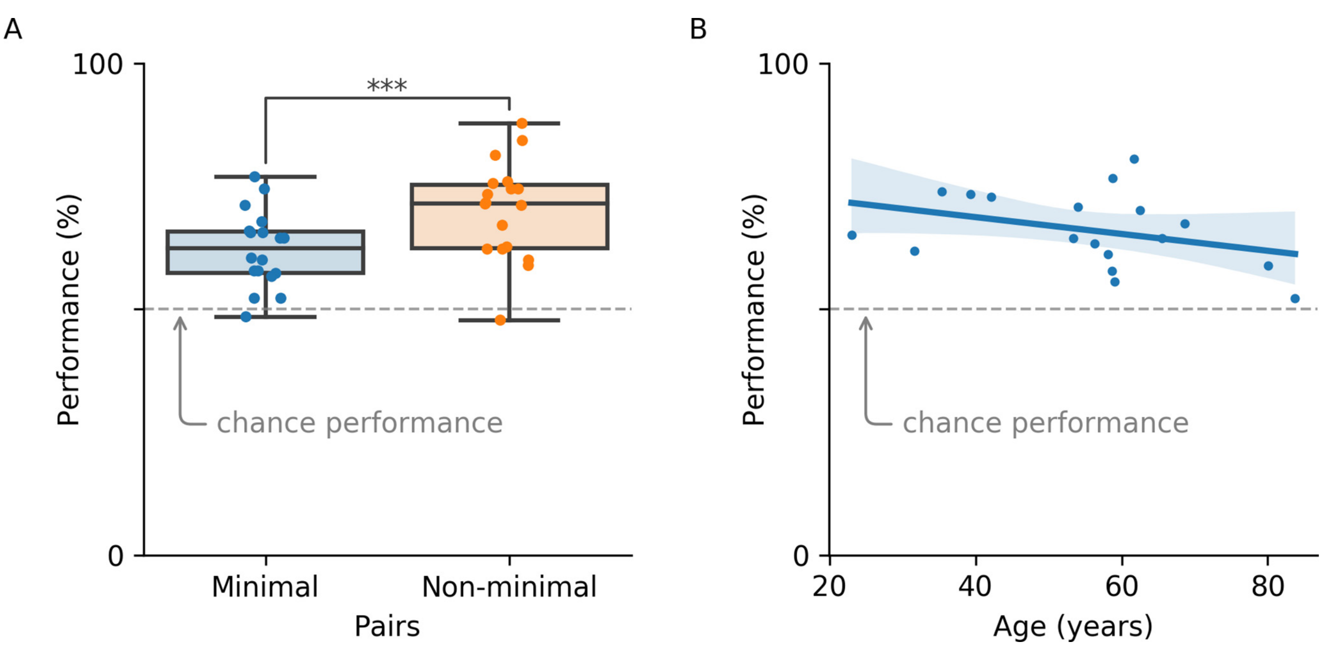

Fig. 6. Pattern discrimination performance. **(A)** Participants scored significantly higher on non-minimal pairs than they did on minimal pairs ($p < 0.001$). **(B)** Participants who are older tended to score lower than their younger counterparts (non-significant, $p = 0.075$).

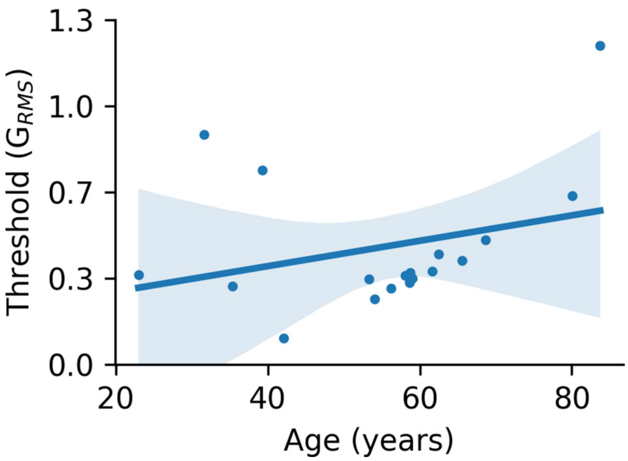

Fig. 7. Minimum threshold. Participants who are older tended to have a higher minimum threshold amplitude than their younger counterparts (non-significant, $p = 0.122$).

identification. Moreover, participants performed above chance on their first day, even before any experience wearing the wristband around sounds in their everyday lives. Performance then increased significantly over time.

Note that, while some participants were born deaf and would never have heard some of the signal sounds, they nevertheless scored highly. This could be because (1) someone who is deaf can gain a sense of the auditory signal coming from a sound through multimodal cues (e.g. seeing a door knock) and (2) the participants can learn about new auditory information by wearing the wristband over the course of the month that they participated in this study.

On tests of pattern discrimination, participants could significantly detect differences between two words that differed by only a single phoneme (minimal pairs) and by more than a single phoneme (i.e. non-minimal pairs). They could detect these differences even in the absence of duration differences of the vibration patterns.

When looking at correlations between performance and age, we only found significant results for the sound identification task. However, the low but insignificant p-values for the pattern discrimination task and minimum threshold task suggest that it is possible we will find a significant effect when testing with a greater number of participants.

There are several possibilities for the age-performance correlation. First, previous studies show significant decreases in vibration thresholds with age (Verrillo, 1980; Gescheider et al., 1996; Deshpande et al., 2008). It stands to reason that as the skin loses sensitivity a participant receives less haptic information (because they cannot feel the weak vibrations) and thus has more difficulty identifying a sound category that caused a pattern of vibrations or discriminating two similar patterns of vibration. A second reason may be that older participants are worse at learning due to decreased neuroplasticity (Kempermann et al., 2002; Lu et al., 2004). Although decreased learning would not by itself explain a correlation between age and score on day 0 (which we did see), it could explain why younger participants improved faster. Other possible reasons why older participants scored worse might be that they had lower attention (Wright and Elias, 1979; McDowd and Birren, 1990; Bolton and Staines, 2012), were less familiar with computer-based tests and games, or had worse two-point discrimination on the skin (Shimokata and Kuzuya, 1995; Bowden and McNulty, 2013).

Finally, some users ask why we use multiple motors instead of a single motor. Note that a single motor would not provide the same performance due to sounds that are distinguishable only through frequency components. For instance, the sound categories of running water, car honk, siren, and phone ring can all consist of a continuous, equal-amplitude sound. However, users are able to distinguish between these categories because of the differences in frequency components: a car honk is fairly steady, sirens alternate between frequencies but at a much slower pace than a phone ring, and running water is spread noisily throughout the frequencies.

We conclude that a practical, wrist-worn sensory substitution device with four vibrating motors is capable of providing enough information to a deaf or hard of hearing user that they are able to identify sound categories through patterns of vibrations.

Our next steps are to run studies that quantify performance over a longer period of time and with a larger number of participants. Studies over a matter of months will allow us to assess performance changes over time and to determine how high performance can get. Many studies of cochlear implants wait six months after implantation before running sound identification tasks and continue to see improvements over years following implantation (Grant et al., 1999; Reed and Delhorne, 2005; Shafiro et al., 2015; Strelnikov et al., 2018). This suggests we might continue to see improvement in performance with the wristband over the course of years.

Future studies will also investigate the neural correlates of learning with the wristband. We hypothesize that deaf participants who use the wristband will begin to show activation in their auditory cortex when feeling patterns of vibration on their wrist – in part because that will be the cortical territory available for takeover (Finney et al., 2001; Bola et al., 2017; Eagleman, 2020). A non-exclusive hypothesis is that activation may be seen in multisensory cortical areas, such as the caudal auditory belt cortex (Kayser et al., 2005). Although we cannot have wearers participate in functional magnetic resonance imaging (fMRI) with the wristband (due to magnetic components in the device), we can either (1) investigate cortical activation with brain imaging techniques that do not rely on magnetic fields, such as functional near-infrared spectroscopy (fNIRS) (Ferrari and Quaresima, 2012) or (2) create an fMRI-safe version of the wristband that stimulates sensations on the wrist without needing magnetic components.

In conclusion, we have demonstrated the usefulness of an inexpensive, self-contained, wrist-worn device for the sensory substitution of sound.

CONFLICTS OF INTEREST

All authors were employed by company Neosensory, Inc.

ACKNOWLEDGEMENTS

We thank Scott Novich and several engineers from Neosensory for their background contributions to this work. MVP and DME designed the experiments and wrote the manuscript. MVP and TA ran the experiments and analyzed the data. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Alles D (1970) Information transmission by phantom sensations. IEEE Trans Man Mach Syst 11:85–91. https://doi.org/10.1109/tmms.1970.299967.

Amedi A, Stern WM, Camprodon JA, Bermpohl F, Merabet L, Rotman S, et al. (2007) Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. Nat Neurosci 10:687–689. https://doi.org/10.1038/nn1912.

Auer Jr ET, Bernstein LE, Coulter DC (1998) Temporal and spatio-temporal vibrotactile displays for voice fundamental frequency: an initial evaluation of a new vibrotactile speech perception aid with normal-hearing and hearing-impaired individuals. J Acoust Soc Am 104:2477–2489. https://doi.org/10.1121/1.423909.

Auvray M, Hanneton S, O'Regan JK (2007) Learning to perceive with a visuo—auditory substitution system: localisation and object recognition with "The Voice" Available at:. Perception 36:416–430. Available from: https://journals.sagepub.com/doi/abs/10.1068/p5631.

Bach-y-Rita P (1983) Tactile vision substitution: past and future. Int J Neurosci 19:29–36. https://doi.org/10.3109/00207458309148643.

Bach-y-Rita P (1999) Theoretical aspects of sensory substitution and of neurotransmission-related reorganization in spinal cord injury. Spinal Cord 37:465–474. https://doi.org/10.1038/sj.sc.3100873.

Bach-Y-Rita P (2004) Tactile sensory substitution studies Available at:. Ann N Y Acad Sci 1013:83–91. Available from: https://pdfs.semanticscholar.org/8f2e/1b6f57b1d5f299daed783641903fa904ee3d.pdf.

Bach-y-Rita P, Collins CC, Saunders FA, White B, Scadden L (1969) Vision substitution by tactile image projection. Nature 221:963–964. https://doi.org/10.1038/221963a0.

Bach-y-Rita P, Kercel W, S. (2003) Sensory substitution and the human–machine interface. Trends Cogn Sci 7:541–546. https://doi.org/10.1016/j.tics.2003.10.013.

Bernstein LE, Demorest ME, Coulter DC, O'Connell MP (1991) Lipreading sentences with vibrotactile vocoders: Performance of normal-hearing and hearing-impaired subjects. J Acoust Soc Am 90:2971–2984. https://doi.org/10.1121/1.401771.

Bernstein LE, Eberhardt SP, Demorest ME (1989) Single-channel vibrotactile supplements to visual perception of intonation and stress. J Acoust Soc Am 85:397–405. https://doi.org/10.1121/1.397690.

Bola Ł, Zimmermann M, Mostowski P, Jednoróg K, Marchewka A, Rutkowski P, Szwed M (2017) Task-specific reorganization of the auditory cortex in deaf humans. Proc Natl Acad Sci U S A 114(4):E600–E609. https://doi.org/10.1073/pnas.1609000114.

Bolton DAE, Staines WR (2012) Age-related loss in attention-based modulation of tactile stimuli at early stages of somatosensory processing. Neuropsychologia 50:1502–1513. https://doi.org/10.1016/j.neuropsychologia.2012.03.002.

Bowden JL, McNulty PA (2013) Age-related changes in cutaneous sensation in the healthy human hand. Age 35:1077–1089. https://doi.org/10.1007/s11357-012-9429-3.

Bubic A, Striem-Amit E, Amedi A (2010) Large-scale brain plasticity following blindness and the use of sensory substitution devices. In: Multisensory object perception in the primate brain (J Kaiser, MJ Naumer, eds) pp 351–380 New York, NY: Springer New York doi:10.1007/978-1-4419-5615-6_18.

Danilov Y, Tyler M (2005) Brainport: an alternative input to the brain. J Integr Neurosci 4:537–550. https://doi.org/10.1142/s0219635205000914.

Deshpande N, Metter EJ, Ling S, Conwit R, Ferrucci L (2008) Physiological correlates of age-related decline in vibrotactile sensitivity. Neurobiol Aging 29:765–773. https://doi.org/10.1016/j.neurobiolaging.2006.12.002.

Eagleman DM (2020) Livewired: the inside story of the ever-changing brain. New York: Pantheon.

Eberhardt SP, Bernstein LE, Demorest ME, Goldstein MH (1990) Speechreading sentences with single-channel vibrotactile presentation of voice fundamental frequency. J Acoust Soc Am 88:1274–1285. https://doi.org/10.1121/1.399704.

Ferrari M, Quaresima V (2012) A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 63:921–935. https://doi.org/10.1016/j.neuroimage.2012.03.049.

Finney EM, Fine I, Dobkins KR (2001) Visual stimuli activate auditory cortex in the deaf. Nat Neurosci 4(12):1171–1173. https://doi.org/10.1038/nn763.

Fletcher MD, Hadeedi A, Goehring T, Mills SR (2019) Electro-haptic enhancement of speech-in-noise performance in cochlear implant users. Sci Rep 9:1–8. https://doi.org/10.1038/s41598-019-47718-z.

Geldard FA (1957) Adventures in tactile literacy. Available at. Available from: https://books.google.com/books/about/Adventures_in_tactile_literacy.html?hl=&id=zBCwoAEACAAJ.

Gescheider GA, Edwards RR, Lackner EA, Bolanowski SJ, Verrillo RT (1996) The effects of aging on information-processing channels in the sense of touch: III. Differential sensitivity to changes in stimulus intensity. Somatosens Mot Res 13:73–80. https://doi.org/10.3109/08990229609028914.

Grant GD, Cheng AK, Niparko JK (1999) Meta-analysis of pediatric cochlear implant literature. Ann Otol Rhinol Laryngol 108:124–128. https://doi.org/10.1177/00034894991080s425.

Grant P, Spencer L, Arnoldussen A, Hogle R, Nau A, Szlyk J, et al. (2016) The functional performance of the BrainPort V100 device in persons who are profoundly blind. J Vis Impair Blind 110:77–88. https://doi.org/10.1177/0145482X1611000202.

Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. Neuron 48:373–384. https://doi.org/10.1016/j.neuron.2005.09.018.

Kempermann G, Gast D, Gage FH (2002) Neuroplasticity in old age: sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. Ann Neurol 52:135–143. https://doi.org/10.1002/ana.10262.

Lu T, Pan Y, Kao S-Y, Li C, Kohane I, Chan J, et al. (2004) Gene regulation and DNA damage in the ageing human brain. Nature 429:883–891. https://doi.org/10.1038/nature02661.

Luzhnica G, Stein S, Veas E, Pammer V, Williamson J, Smith RM (2017) Personalising vibrotactile displays through perceptual sensitivity adjustment. In: Proceedings of the 2017 ACM international symposium on wearable computers - ISWC. p. '17.. https://doi.org/10.1145/3123021.3123029.

Maidenbaum S, Arbel R, Buchs G, Shapira S, Amedi A (2014) Vision through other senses: Practical use of sensory substitution devices as assistive technology for visual rehabilitation. In: 22nd Mediterranean conference on control and automation. https://doi.org/10.1109/med.2014.6961368.

McDowd JM, Birren JE (1990) Aging and attentional processes. Handb Psychol Aging 222–233. https://doi.org/10.1016/b978-0-12-101280-9.50019-x.

McGill R, Tukey JW, Larsen WA (1978) Variations of box plots. The American Statistician 32:12–16. https://doi.org/10.1080/00031305.1978.10479236.

Novich SD, Eagleman DM (2015) Using space and time to encode vibrotactile information: toward an estimate of the skin's

achievable throughput. Exp Brain Res 233:2777–2788. https://doi.org/10.1007/s00221-015-4346-1.
Proulx MJ, Brown DJ, Pasqualotto A, Meijer P (2014) Multisensory perceptual learning and sensory substitution. Neurosci Biobehav Rev 41:16–25. https://doi.org/10.1016/j.neubiorev.2012.11.017.
Rahal L, Cha J, El Saddik A (2009) Continuous tactile perception for vibrotactile displays. In: 2009 IEEE international workshop on robotic and sensors environments. https://doi.org/10.1109/rose.2009.5355986.
Rauschecker JP (1995) Compensatory plasticity and sensory substitution in the cerebral cortex. Trends Neurosci 18:36–43. https://doi.org/10.1016/0166-2236(95)93948-w.
Reed CM, Delhorne LA (2003) The reception of environmental sounds through wearable tactual Aids. Ear Hear 24:528–538. https://doi.org/10.1097/01.AUD.0000100207.97243.88.
Reed CM, Delhorne LA (2005) Reception of environmental sounds through cochlear implants. Ear Hear 26:48–61. https://doi.org/10.1097/00003446-200502000-00005.
Riso RR (1999) Strategies for providing upper extremity amputees with tactile and hand position feedback – moving closer to the bionic arm. Technol Health Care 7:401–409. https://doi.org/10.3233/thc-1999-7604.
Shafiro V, Sheft S, Kuvadia S, Gygi B (2015) Environmental sound training in cochlear implant users. J Speech Lang Hear Res 58:509–519. https://doi.org/10.1044/2015_JSLHR-H-14-0312.
Shimokata H, Kuzuya F (1995) Two-point discrimination test of the skin as an index of sensory aging. Gerontology 41:267–272. https://doi.org/10.1159/000213693.
Strelnikov K, Collett E, Gaillard P, Truy E, Déguine O, Marx M, et al. (2018) Categorisation of natural sounds at different stages of auditory recovery in cochlear implant adult deaf patients. Hear Res 367:182–194. https://doi.org/10.1016/j.heares.2018.06.006.
Striem-Amit E, Guendelman M, Amedi A (2012) "Visual" acuity of the congenitally blind using visual-to-auditory sensory substitution. PLoS ONE 7:e33036. https://doi.org/10.1371/journal.pone.0033136.
Tyler M, Danilov Y, Bach-Y-Rita P (2003) Closing an open-loop control system: vestibular substitution through the tongue. J Integr Neurosci 2:159–164. https://doi.org/10.1142/s0219635203000263.
Verrillo RT (1980) Age related changes in the sensitivity to vibration. J Gerontol 35:185–193. https://doi.org/10.1093/geronj/35.2.185.
Ward J, Meijer P (2010) Visual experiences in the blind induced by an auditory sensory substitution device. Conscious Cogn 19:492–500. https://doi.org/10.1016/j.concog.2009.10.006.
Weisenberger JM, Craig JC, Abbott GD (1991) Evaluation of a principal-components tactile aid for the hearing-impaired. J Acoust Soc Am 90:1944–1957. https://doi.org/10.1121/1.401674.
Weisenberger JM, Heidbreder AF, Miller JD (1987) Development and preliminary evaluation of an earmold sound-to-tactile aid for the hearing-impaired Available at:. J Rehabil Res Dev 24:51–66. Available from: https://www.ncbi.nlm.nih.gov/pubmed/3585784.
Weisenberger JM, Kozma-Spytek L (1991) Evaluating tactile aids for speech perception and production by hearing-impaired adults and children Available at:. Am J Otol 12(Suppl):188–200. Available from: https://www.ncbi.nlm.nih.gov/pubmed/2069181.
Weisenberger JM, Miller JD (1987) The role of tactile aids in providing information about acoustic stimuli. J Acoust Soc Am 82:906–916. https://doi.org/10.1121/1.395289.
Weisenberger JM, Percy ME (1995) The transmission of phoneme-level information by multichannel tactile speech perception aids. Ear Hear 16:392–406. https://doi.org/10.1097/00003446-199508000-00006.
Weisenberger JM, Russell AF (1989) Comparison of two single-channel vibrotactile aids for the hearing-impaired. J Speech Hear Res 32:83–92. https://doi.org/10.1044/jshr.3201.83.
Wright LL, Elias JW (1979) Age differences in the effects of perceptual noise. J Gerontol 34:704–708. https://doi.org/10.1093/geronj/34.5.704.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroscience.2021.01.008.

(Received 23 August 2020, Accepted 5 January 2021)
(Available online 16 January 2021)