# Enhancing Touch Sensibility by Sensory Retraining in a Sensory Discrimination Task via Haptic Rendering

**Eduardo Villar Ortega [1,*], Efe Anil Aksöz [1,2], Karin A. Buetler [1] and Laura Marchal-Crespo [1,3]**

[1] *Motor Learning and Neurorehabilitation Laboratory, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland*
[2] *Department of Engineering and Information Technology, Bern University of Applied Sciences, Burgdorf, Switzerland*
[3] *Department of Cognitive Robotics, Delft University of Technology, Delft, Netherlands*

Correspondence*:
Corresponding Author
eduardo.villarortega@unibe.ch

## ABSTRACT

Stroke survivors are commonly affected by somatosensory impairment, hampering their ability to interpret somatosensory information. Somatosensory information has been shown to critically support movement execution in healthy individuals and stroke survivors. Despite the detrimental effect of somatosensory impairments on performing activities of daily living, somatosensory training –in stark contrast to motor training– does not represent standard care in neurorehabilitation. Reasons for the neglected somatosensory treatment are the lack of high-quality research demonstrating the benefits of somatosensory interventions on stroke recovery, the unavailability of reliable quantitative assessments of sensorimotor deficits, and the labor-intensive nature of somatosensory training that relies on therapists guiding the hands of patients with motor impairments. To address this clinical need, we developed a virtual reality-based robotic texture discrimination task to assess and train touch sensibility. Our system incorporates the possibility to robotically guide the participants' hands during texture exploration (i.e., passive touch) and no-guided free texture exploration (i.e., active touch). We ran a three-day experiment with thirty-six healthy participants who were asked to discriminate the odd texture among three visually identical textures –haptically rendered with the robotic device– following the method of constant stimuli. All participants trained with the passive and active conditions in randomized order on different days. We investigated the reliability of our system using the Intraclass Correlation Coefficient (ICC). We also evaluated the enhancement of participants' touch sensibility via somatosensory retraining and compared whether this enhancement differed between training with active vs. passive conditions. Our results showed that participants significantly improved their task performance after training. Moreover, we found that training effects were not significantly different between active and passive conditions, yet, passive exploration seemed to increase participants' perceived competence. The reliability of our system ranged from poor (in active condition) to moderate and good (in passive condition), probably due to the dependence of the ICC on the between-subject variability, which in a healthy population is usually small. Together, our virtual reality-based robotic haptic system may be a key asset for evaluating and retraining sensory loss with minimal supervision, especially for brain-injured patients who require guidance to move their hands.

**Keywords: Haptic rendering, sensory rehabilitation, active exploration, passive exploration, touch, texture discrimination.**

## 1 INTRODUCTION

Stroke is the most common acquired brain injury that causes persisting and long-term disability in adults (Adamson et al., 2004). Between 65 % and 85 % of stroke survivors suffer from somatosensory impairment (Carey et al., 1993; Chia et al., 2019), hampering individuals' ability to interpret somatosensory information (Pumpa et al., 2015), and thus, their ability to perform skillful movements independently (Schabrun and Hillier, 2009; Taylor et al., 2021). Importantly, somatosensory impairment increases patients' hospitalization time (Sommerfeld and von Arbin, 2004) and limits the recovery of sensorimotor function (Meyer et al., 2014). Despite the negative impact of somatosensory impairment on upper limb functionality and recovery (Yilmazer et al., 2019; Zandvliet et al., 2020), somatosensory training is not the standard of care following stroke (Schabrun and Hillier, 2009; Serrada et al., 2019) and generally receives less attention than motor training (Yilmazer et al., 2019). The lack of time for therapy and limited access to somatosensory training guidelines are some factors that may contribute to the lack of attention to sensory rehabilitation (Pumpa et al., 2015).

Somatosensory interventions are therapeutic techniques performed by a therapist designed to retrain sensory function (O'Tool, 2017). Somatosensory interventions can be classified as sensory retraining –i.e., interpretation of a stimulus– and sensory stimulation –i.e., afferent stimulation (Doyle et al., 2010). **Sensory retraining** involves the patients' interpretation of stimuli, which are usually provided by the therapist. An example of sensory retraining intervention is the tactile discrimination test (TDT), a conventional approach to evaluate and train touch sensibility in clinical settings (Carey et al., 1993). TDT is performed by asking the patient to tactually explore gratings textures, also known as active touch. The therapist may also guide the patient's paretic hand (i.e., passive touch) when the patient has a severe motor deficit. **Sensory stimulation**, by contrast, relies on the therapist providing a stimulus, e.g., transcutaneous electrical nerve stimulation, while the patient does not move and is asked to simply feel the stimulus without an active motor or cognitive reaction (Schabrun and Hillier, 2009; Yilmazer et al., 2019).

Somatosensory interventions have shown promising results in enhancing sensory discrimination –i.e., the skill to discern and interpret specific sensory stimuli (O'Tool, 2017)– in stroke survivors (Taylor et al., 2021). Moreover, sensory retraining interventions have been found beneficial for the recovery of motor function in stroke patients (Yilmazer et al., 2019) and improvement of somatosensory function, especially those interventions based on the discrimination of textures, proprioceptive discrimination tests, and tactile object recognition (Carey et al., 2011; Elangovan et al., 2017; Yeh et al., 2021). Yet, in the last decade, six reviews concluded that there is insufficient empirical evidence regarding the effectiveness of sensory retraining interventions on the recovery of sensorimotor function after a brain injury (Schabrun and Hillier, 2009; Taylor et al., 2021; Yilmazer et al., 2019; Serrada et al., 2019; Turville et al., 2019; Doyle et al., 2010). These reviews cited poor quality of study designs, variations in outcome measurements (Carlsson et al., 2021), small sample sizes, and inadequate statistical power to detect meaningful differences between control and treatment (Taylor et al., 2021) as limiting factors to draw clear conclusions. Two recent reviews about the effectiveness of somatosensory interventions concluded that high-quality research is necessary to determine whether sensory retraining is effective in stroke rehabilitation (Yilmazer et al., 2019; Serrada et al., 2019).

Quantitative reliable assessment of the sensorimotor performance is needed to evaluate if patients are achieving functional rehabilitation gains after somatosensory interventions (Taylor et al., 2021; Turville et al., 2019). However, conventional somatosensory assessments may present variations in results, especially when the assessment is performed by different clinicians (Lincoln et al., 1991). A systematic, qualitative, and objective assessment of touch sensibility would facilitate diagnosis, prognosis, and the selection of adequate somatosensory treatments according to the patients' touch sensibility (Pumpa et al., 2015).

Robotics is a promising technology to quantitatively assess somatosensory function and provide somatosensory training. Compared to other conventional treatments, robotic devices are capable of delivering precise and reproducible stimuli (Zbytniewska et al., 2021). Further, robots can physically guide the patients' limbs during sensorimotor training (Basalp et al., 2021), facilitating the admission of patients with severe motor impairments into the training and enhancing their motivation and engagement during repetitive and intensive practice (Rowe et al., 2017). However, despite their potential, the usage of robots to assess and treat somatosensory function is, to date, mainly neglected (Ballardini et al., 2018; Handelzalts et al., 2021). Although research efforts have been made to assess and enhance proprioceptive function with robots –e.g., Kenzie et al. (2017); Zbytniewska et al. (2021); Elangovan et al. (2017); Yeh et al.

(2021); Cappello et al. (2015)–, fewer efforts have been done into assessing and training tactile function (Ballardini et al., 2018). Currently, there is a clinical need to develop robotic systems to assess and train touch sensibility in patients with limited motor function that within this project we aim to meet.

When designing a robotic device for assessment and training of touch sensibility, especially for those patients with severe motor impairments who require robotic assistance to move their paretic limbs, it is important to first understand the differences in touch sensibility perception when a patient is assisted or not. The perception of touch sensibility may differ depending on the mode of touch: active or passive touch. Fundamental studies of touch sensation refer to **active touch** as the *action* of touching, e.g., by actively moving the limbs. In contrast, **passive touch** refers to two different processes: 1) the act of being touched, while the limb does not move/is not passively moved (Lederman and Klatzky, 2009), and 2) the act of touching while being assisted by an external agent (e.g., by a therapist) (Symmons et al., 2004). It is not yet fully understood how the nervous system processes active and passive touch. According to Pertovaara et al., active touch relies on the afferent-induced mechanism and the motor command signals, whereas passive touch relies mainly on the afferent-induced mechanism (Pertovaara et al., 1994). Further, active touch involves participants choosing their exploration strategy, notably the intention, planning, preparation, and execution of the movements (Van Doorn et al., 2012). On the contrary, passive exploration is considered to minimize any involvement of decision-making processes (Van Doorn et al., 2012), allowing participants to focus on the perception of the stimulus (Magee and Kennedy, 1980). Consistently, Van Doorn et al. (2012) found an increase in attentional networks activity in the parietal lobe in active touch compared to passive touch using functional Magnetic Resonance Imaging (fMRI).

In their review, Symmons et al. (2004) attributed differences in sensory perception between active and passive modalities to the task characteristics and the nature of the stimulus, rather than the exploration mode. For example, Magee and Kennedy (1980) found passive exploration to be better in the discrimination of dot-pattern shapes when compared to active exploration, while Richardson et al. (1981) found no differences between active and passive touch in discriminating embossed-dots mazes. Vega-Bermudez et al. (1991) associated the differences between active and passive touch to two main causes: 1) the experimenter failing to provide equivalent somatosensory information in both modalities and 2) differences in the sensory neural mechanisms underlying tactual pattern recognition behavior. Thus, for comparing active versus passive touch, the experimenter should provide the same stimulus in both conditions, including kinesthetic information regarding the movement of the limb. Furthermore, other subjective factors such as motivation, might play a role in the differences between passive and active touch. Active engagement during training has been associated with higher motivation (Colombo et al., 2007), while high motivation is associated with an increase in motor adjustments based on sensory signals (Lezkan et al., 2018). However, passive touch may allow participants to better focus on the task (i.e., to the sensory input), enhancing their perceived competence (Wenk et al., 2022).

This study aims to evaluate a novel robotic intervention to assess and train tactile function and, when needed, provide robotic assistance to guide the hand during passive touch. We developed a sensory discrimination task to characterize and treat the acuity of touch sensibility via the perception of virtual textures rendered by a haptic robotic device (Fig. 1). The novelty of our approach relies on the provision of the haptic rendering forces from the virtual textures that are independent of the normal forces that participants exert against the virtual surface, and thus, providing more controlled stimuli within and between participants.

We ran a three-day within-subject experiment with 37 healthy participants who actively and passively (i.e., assisted by the robot) explored a set of virtual haptically rendered textures and selected the odd texture among three visually identical textures. The first session consisted of two initial baselines. The remaining two sessions comprised three phases: baseline, training, and retention, performed each with passive or active touch in randomized order. In this paper, we evaluated: 1) the system reliability, 2) the change in participants' touch sensibility after somatosensory training, 3) differences in touch discrimination changes pre-post training between active and passive conditions, and 4) differences between passive and active touch conditions on participants' motivation. We hypothesized that passive exploration would have a higher reliability coefficient than active exploration since the provision of stimuli is more controlled. We also hypothesized that participants would improve their tactile acuity of textures after training. Moreover, the improvement of touch sensibility after training would not differ between active and passive conditions in our controlled set-up. Finally, we expected active exploration to generally enhance participants' effort,

**Table 1.** Participants demographics. Thirty-seven (one participant excluded) participants were recruited for the experiment.

| Characteristics | Participants (N= 36) no. (%) |
|---|---|
| Gender | |
| Female | 18 (50) |
| Male | 18 (50) |
| Age | |
| $<$25 yr | 4 (11.111) |
| 25-35 yr | 25 (69.444) |
| $\geq$ 35 yr | 7 (19.444) |
| Handedness | |
| Right | 34 (94.444) |
| Left | 2 (5.556) |

pressure, and enjoyment/interest during the task (Lezkan et al., 2018), yet, haptic guidance may specifically increase the self-reported level of perceived competence (Wenk et al., 2022).

# 2 METHODS

## 2.1 Participants

Thirty-seven healthy participants gave written informed consent to participate in the study. One participant was excluded from the analysis due to a hardware failure during the second session and could not participate in the third session. Thus, 36 participants (50 % females) completed the experiment, see demographic information in Table 1. The study was approved by the local ethical committee (Swiss Cantonal Ethics Committee; Basec ref: 2018-01179) and the Swiss Agency for Therapeutic Products (Swissmedic ref: 100000432), and conducted in compliance with the Declaration of Helsinki. The participants' hand dominance was assessed with the Edinburgh Handedness Inventory (Bryden, 1977).

The sample size was calculated using the R package "sensR" (Christensen and Brockhoff, 2020). First, we ran a first pilot experiment with seven healthy participants, in which the average number of correct responses after 40 trials was 25. Second, we used the average number of correct responses, assumed the desired power of 0.95, a type I error of alpha equal to 0.05, and a probability guess of 1/3 (i.e., triangle test) to compute the sample size. The result of the sample size computation was 34.

## 2.2 Experimental set-up

The experimental set-up (Fig. 1) consisted of a 24 inch monitor (S24E650, Samsung, South Korea), a robotic device (Delta.3, Force Dimension, Switzerland), a passive arm weight support (SaeboMAS mini, Saebo, USA), noise-canceling headphones (WH-1000XM4, Sony, Japan), and a custom-made response box with a push-button.

During the experiment, participants were seated at a desk on a comfortable chair with backrest. Their dominant arms were placed using Velcro® straps in a passive arm weight support device attached to the table. The passive arm weight support was used to reduce fatigue during the experiment. The weight compensation level was adjusted to each participants' arm weight and kept constant during the three sessions. The location of the monitor, robot, arm-weight support, and response box were adjusted to the handedness of each participant before the start of the experiment and kept constant during the whole experiment. Participants performed the experiment with their dominant hand. Right- (left-)handed participants had the monitor and response box on the left (right) side of their sagittal plane and the arm-weight support on their right (left) side.

Participants were asked to hold the end effector of the haptic device with their dominant hand at all times. Right- (left-)handed participants had the robot located on the right- (left-)side of their sagittal plane and aligned to the shoulder of the dominant arm. The chair height was adjusted such that the participant's

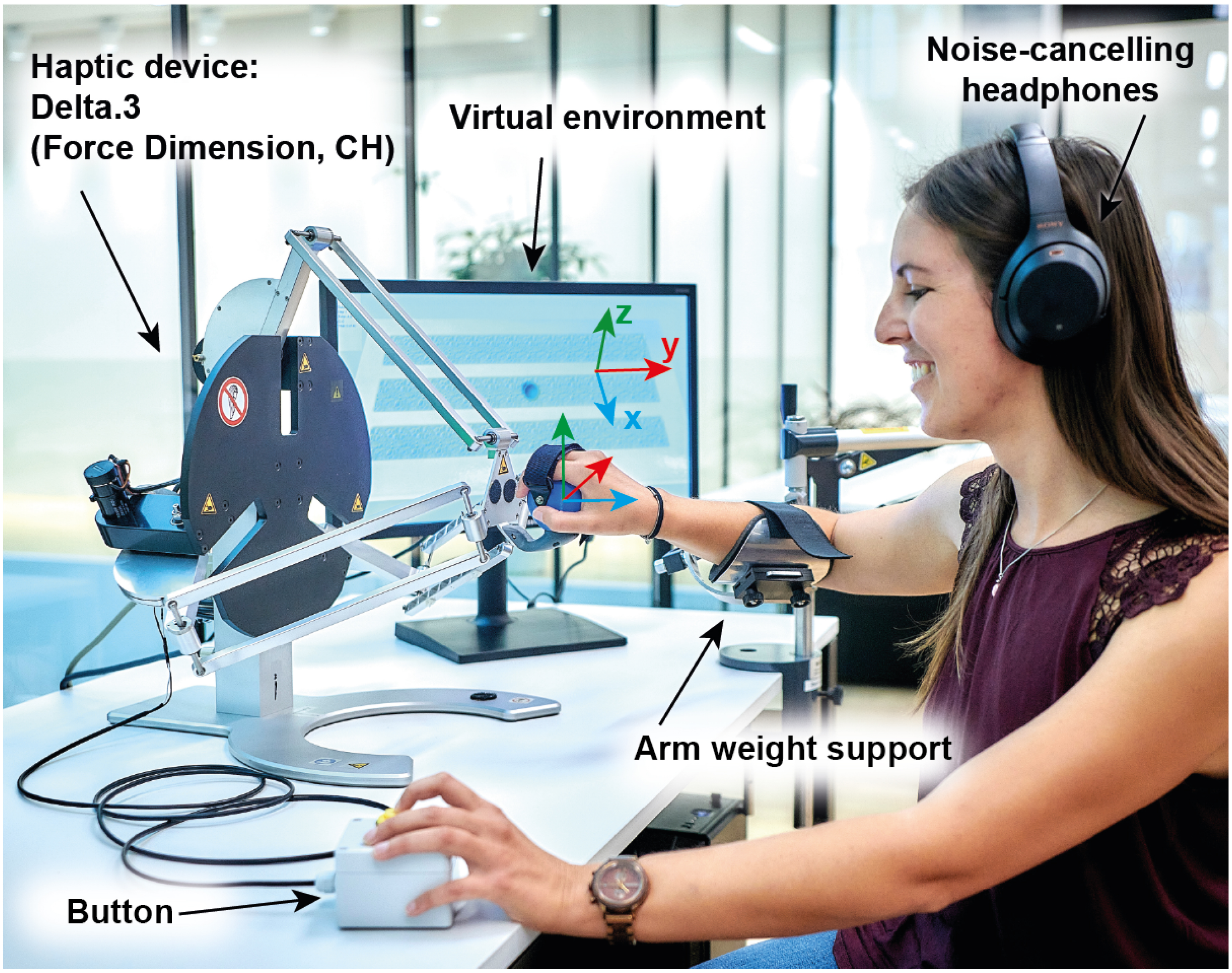


**Figure 1.** Experimental set-up. The Delta.3 robot was located on a table next to a LED monitor, which showed the virtual environment. Participants wore noise-canceling headphones and used an arm-weight support system, which was adjusted to each participant's individual arm weight. Note that the monitor is located on the right-side of a right-handed user only for illustration purposes (during the experiment it was located on the left side of the robot).

shoulder was not below the robot end effector in the center of the workspace. We exchanged the robot commercial end effector with a new 3D-printed handle to improve participants' comfort. The participants' hands were secured to the robot end effector using a Velcro® strap. The participants were wearing active noise-canceling headphones and we delivered white-noise during the experiment to mask any auditory cues from the robot actuators.

## 2.3 The haptic exploration task

We designed a virtual environment to assess and train participants' touch discrimination using haptically rendered virtual textures (i.e., stimuli) using Unity (Unity Technologies, USA). During the experiment, the participants were asked to discriminate the odd texture among three visually identical textures. We asked participants to select the odd texture by pressing a custom-made button on the response box with their non-dominant hand when they were on top of the texture they believed was different from the other two. If a participant pressed the button outside a texture, we registered the last texture explored as their response.

**Table 2.** The set of experimental stimuli. The standard stimulus $S_t$ with spatial frequency $f_{St}$ was kept constant during the experiment. The spatial frequency of the comparison stimulus $f_{Co}$ was varied every trial along with the more coarse and less coarse textures sets. Each *Co* stimulus was presented five times during a block of 40 trials.

| $f_{St}$ $(m^{-1})$ | $f_{Co}$ $(m^{-1})$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | More coarse | | | | Less coarse | | | |
| 164 | 100 | 116 | 132 | 148 | 180 | 196 | 212 | 228 |

The virtual textures were displayed horizontally on the monitor. Each texture had the same dimensions in the physical (robot) workspace (0.176 m x 0.02 m), an area large enough to allow participants to move the robot end effector tangentially across the texture. The textures were located in parallel and separated 0.01 m from each other along the $x$-axis (blue axis in Fig. 1).

We rendered the virtual textures as sinusoidal gratings (see section 2.5 and Fig. 2) using the haptic device Delta.3. During each trial, participants explored the textures either with active or passive touch conditions (see section 2.6). Participants were allowed to explore each texture as many times as wanted and switch between textures when desired.

## 2.4 Tactile stimuli

The stimuli consisted of virtual sinusoidal gratings along the robot end effector $y$-axis (Fig. 2). The interaction forces between the participants' hands and the gratings were rendered by the Delta.3 robotic device and generated by the following equation:

$$F_g = C \sin(2\pi f y_{EE}). \tag{1}$$

The grating forces $F_g$ depended on the robot end effector position along the $y$-axis, $y_{EE}$, and the spatial frequency $f$ of the grating, i.e., the reciprocal of the spatial period defined as the distance between two consecutive crests (Fig. 2). The constant $C$ = 3 N determined the amplitude of the sinusoidal and did not change between textures.

Several different virtual textures were generated (see Table 2). They included eight *comparison stimuli* (*Co*) and one *standard stimulus* (*St*), which differed between them in terms of the value of the spatial frequency $f$. The standard stimulus was fixed during the experiment and employed as a basis for quantitative comparison against the set of comparison stimuli, i.e., stimuli with varying physical attributes. In every trial –defined as a single discrimination attempt of the odd texture in a set of three textures– the three virtual textures consisted of two types of stimuli, the *St* and a random stimulus selected from the set of *Co*. We employed the triangle testing method for sensory discrimination (described in Bi (2015), page 3), i.e., two of the textures were equal with possible combinations sets: St/St/Co, St/Co/St, Co/St/St, Co/Co/St, Co/St/Co, St/Co/Co. The presentation order of the stimuli followed the method of constant stimuli (described in Gescheider (1997), page 46).

The *St* was fixed through all trials, while the *Co* was varied in each trial, from the pool of preselected *Co*. Two preselected pools of *Co* were created, which spanned two ranges of textures: more coarse and less coarse textures (Table 2). The more coarse textures had a spatial frequency that ranged from 100 to 148 $m^{-1}$, and the less coarse textures ranged from 180 to 228 $m^{-1}$. Each spanned set of the *Co* (i.e., more and less coarse sets) consisted of four different stimuli with equal inter-space distance between consecutive *Co* that was set to 16 $m^{-1}$. The most coarse texture (100 $m^{-1}$) corresponded to a spatial period of 10 mm, whereas the least coarse texture (228 $m^{-1}$) corresponded to a spatial period of 4.38 mm. The *St* was set to 164 $m^{-1}$ –the mean of all *Co* spatial frequencies– and kept constant throughout the experiment. We chose the spatial frequency of the *St* to be the average of all *Co* spatial frequencies to avoid any bias toward either of the textures roughness directions.

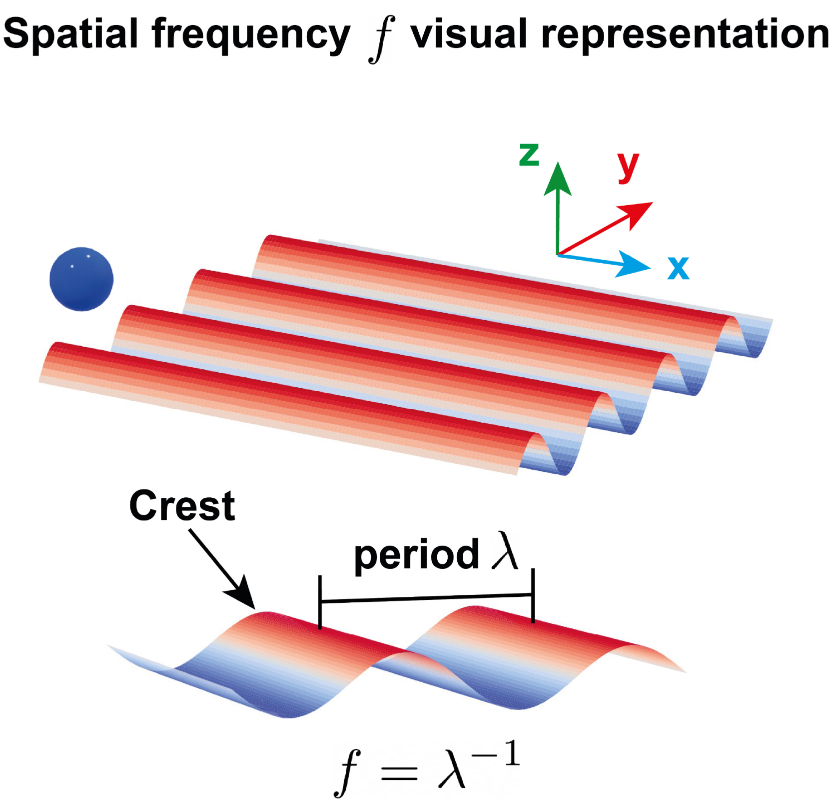


**Figure 2.** Visual representation of a virtual sinusoidal grating. Each texture consisted of adjacent crests along the $y$-axis. The spatial frequency $f$ was defined as the inverse of the spatial period $\lambda$, i.e., the distance between two consecutive crests. The blue ball illustrates the position of the robot end effector, $y_{EE}$, in the virtual environment. The scale and size of the texture and the ball are only for visual purposes and do not represent the scale and size used in the experiment.

The *Co* spatial frequency values in Table 2 had various levels of discrimination difficulty, i.e., the closer they were to the $f_{St}$ the more similar they were perceived and more difficult to differentiate with respect to the $S_t$ became. The values of the spatial frequencies of *Co* and *St* were selected after running a first pilot experiment with seven healthy participants such that they were within the range used in literature (Campion and Hayward, 2005; Cholewiak et al., 2010), considering the resolution of the robot (i.e., 0.02 mm), and stimuli that were not judged as too easy nor too difficult by the participants.

## 2.5 Haptic rendering of virtual textures

The virtual textures –visually represented in Fig. 3– were rendered using the grating force calculated in equation 1 (see section 2.4). The textures were rendered (i.e., through the force $F_{rd}$) only on the $y$-direction and participants only perceived them when they were in contact with the texture, i.e., when the position of the robot end effector was below the virtual table height ($z_{EE} < z_{tbl} = 0.001\ m$) and within the perimeter of the virtual textures in the $xy$-plane.

$$F_{rd} = \begin{cases} F_g & \text{if } contact \text{ is } True \\ 0 & \text{else.} \end{cases} \tag{2}$$

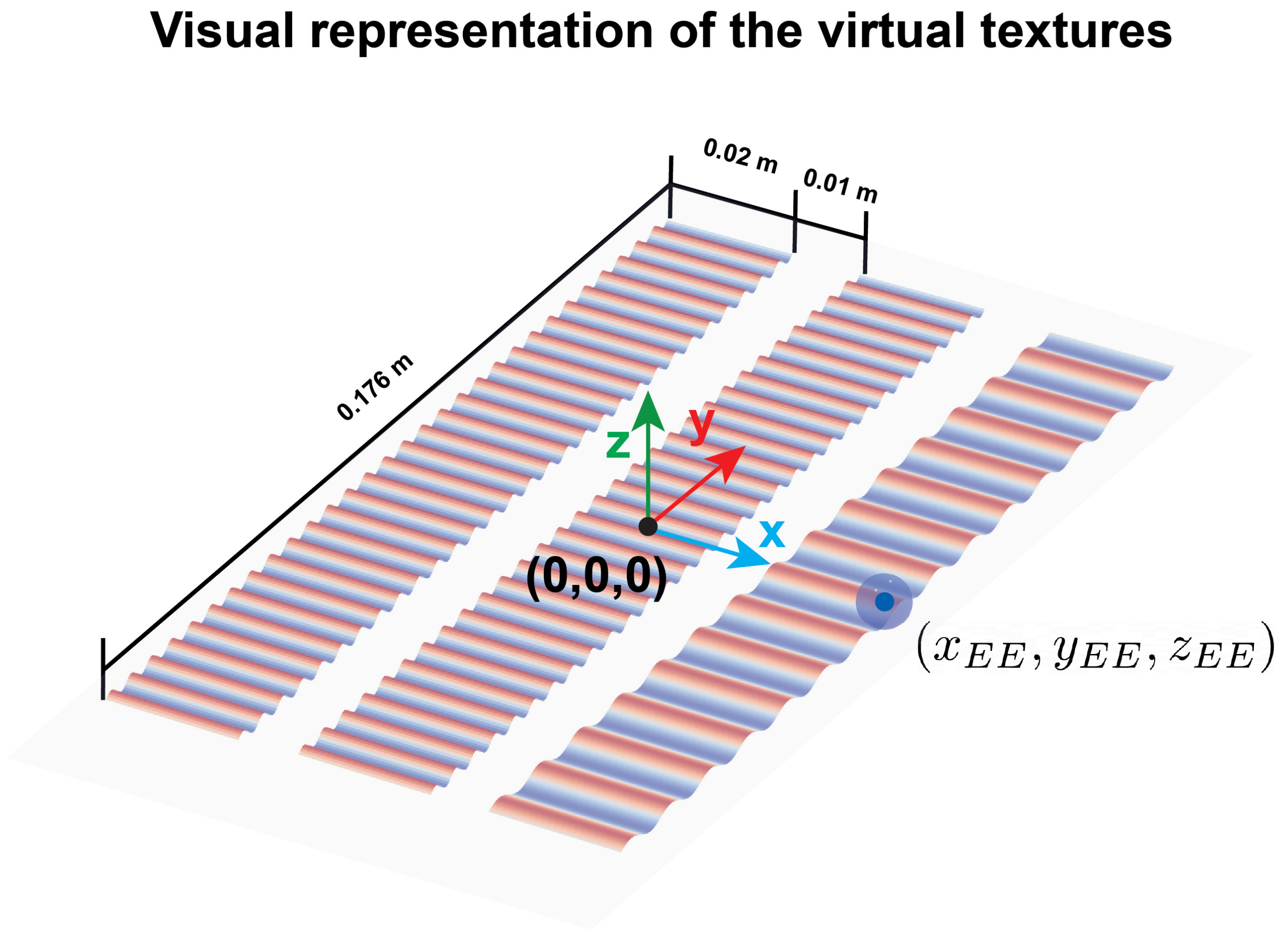


**Figure 3.** Visual representation of the three virtual textures. The coordinate center of the workspace is denoted by (0,0,0) and was located in the center of the second texture. The robot end effector is depicted with a blue ball whose center had the coordinates ($x_{EE}$, $y_{EE}$, $z_{EE}$). The scale and size of the textures and the ball are only for visual purposes and do not represent the scale and size used in the experiment.

The three virtual textures laid on top of a haptic table of $0.20\ m$ x $0.10\ m$ whose surface was rendered by the robot using a Proportional-Derivative (PD) controller:

$$F_z = \begin{cases} K_z(z_{tbl} - z_{EE}) + B_z(-\dot{z}_{EE}) & \text{if } z_{EE} < z_{tbl} \\ 0 & \text{else,} \end{cases} \tag{3}$$

where the rendered force in the vertical direction $F_z$ was proportional ($K_z = 1960\ N/m$) to the difference between the end effector vertical position $z_{EE}$ and the height of the virtual table $z_{tbl} = 0.001\ m$ when the robot end effector height was below the height of the virtual table ($z_{EE} < z_{tbl}$). The force in the vertical direction was zero otherwise. We added a damping element ($B_z = 28\ N.s/m$) to avoid excessive oscillations when the robot end effector was in contact with a rigid virtual surface (Colgate and Brown, 1994). The robot was transparent in the $x$-direction at all times.

## 2.6 Exploration conditions: active and passive touch

Participants explored the virtual textures under two different conditions: with active and passive touch. In this study, we employ the definition of passive touch provided by Symmons et al. (2004), i.e., the act of touching an object while being assisted by an external agent (Symmons et al., 2004). In our case, this external agent was the robot, which physically guided the participants' dominant hands during the exploration of the virtual textures. The haptic guidance $F_{hg}$ was provided using the PD controller described in equation 4:

$$F_{hg} = \begin{cases} \ddot{y_R} + K_{hg}(y_R - y_{EE}) + B_{hg}(\dot{y_R} - \dot{y_{EE}}) & \text{if } contact \text{ is } True \\ 0 & \text{else,} \end{cases} \quad (4)$$

where $y_{EE}$ was the $y$ coordinate of the robot end effector position, i.e., the axis along the perceived textures, and $\dot{y_{EE}}$ its derivative. The stiffness $K_{hg}$ was set to 300 $N/m$, and damping $B_{hg}$ to 60 $N.s/m$. The reference trajectory –defined by $\ddot{y_R}$, $\dot{y_R}$, and $y_R$– was obtained following the cycloidal motion law (described in section 1 of the supplementary material).

Participants were instructed to not oppose to the haptic guidance force and move along with the robot. They could move between textures by either exiting the textures sides in the $XY$-plane or by lifting the end effector ($z$-axis). Therefore, they were instructed not to lift the end effector while the guidance force was on.

Taking together equations 2, 3, and 4, the total force applied by the robot at the end effector was:

$$F_{Total} = F_{rd} + F_z + F_{hg}. \quad (5)$$

## 2.7 Study protocol

Fig. 4 illustrates the experimental protocol of the within-subject experimental design. Participants completed three sessions, performing one session per day. There was a minimum of one to a maximum of two days between sessions.

### 2.7.1 First session

The first session started with a **familiarization** phase followed by two initial baselines (iBL), which included one baseline per condition (i.e., active and passive touch). During the familiarization phase, all participants familiarized themselves with the robot and the experimental stimuli. Participants were invited to explore a single texture of 100 $m^{-1}$, i.e., the more coarse texture in Table 2. During familiarization, we also provided visual feedback that mapped the haptic sensation (see Fig. 5A) to facilitate the understanding of the virtual gratings. The dark color in Fig. 5A represents the grooves, while the light blue color represents the texture crests. Subsequently, in a single familiarization trial, we asked them to select the odd texture among three textures that looked identical (see Fig. 5B). The texture combination for all participants was 228 $m^{-1}$, 228 $m^{-1}$, and 100 $m^{-1}$, respectively, from Table 2.

The first session included two **initial baseline** tests (iBL), one performed with the active and one with the passive condition. Half of the participants (randomly selected) performed the first iBL with the active condition and the second iBL with passive condition. The other half performed the iBLs in the contrary order. Each iBL included 40 trials, where each of the eight different comparison stimuli (*Co*) in Table 2 was presented a total of five times. The order of presentation of the stimuli was randomized for each participant and each condition.

Correct responses in each trial were registered following the criteria:

$$Y_i = \begin{cases} 1 & \text{if the response is correct} \\ 0 & \text{else,} \end{cases} \quad (6)$$

where $Y_i$ represents the correctness of the response for each trial and stimuli $i \in \{1, 8\}$. The total number of correct responses was shown to the participant after finishing each iBL block. We saved the responses for all the 40 trials per iBL to compute the probability of correct responses (see section 2.8).

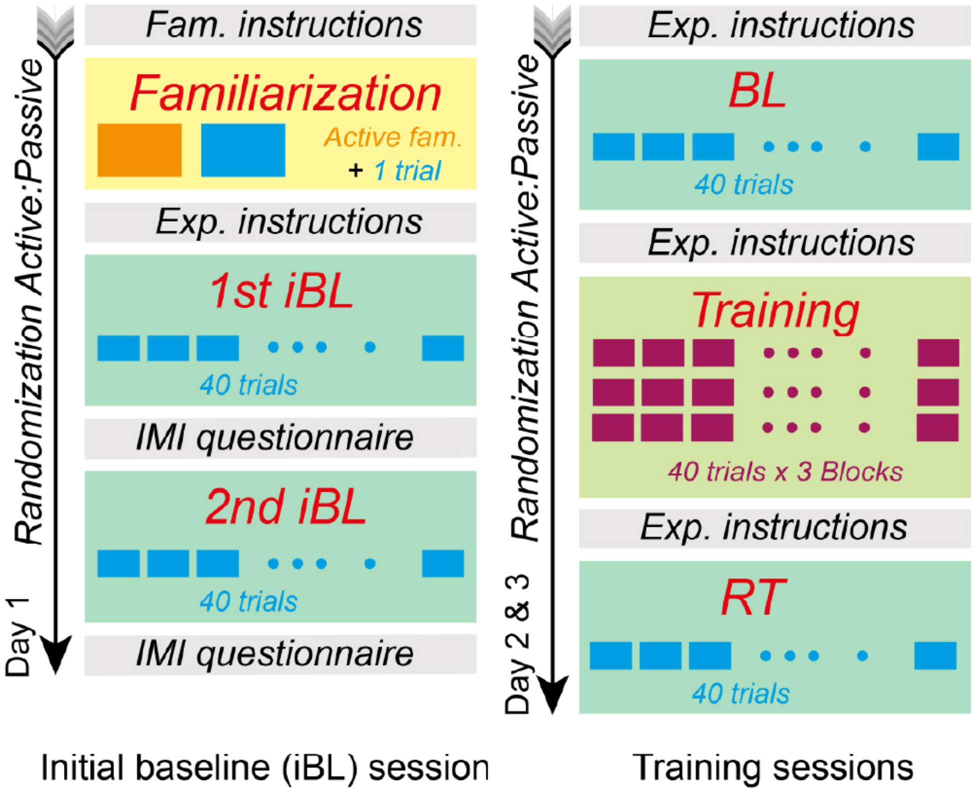


**Figure 4.** Experimental protocol. Participants completed three sessions on different days. The first session included familiarization and an initial baseline (iBL) for each condition. Half of the participants (randomly allocated) performed the first iBL in the active condition, whereas the second half had the passive condition first. We performed a second randomization to define the order of the training conditions during sessions 2 & 3. Each second and third session consisted of baseline (BL), training, and retention (RT) with either, passive or active conditions.

### 2.7.2 Sessions 2 & 3: Training

The sessions in the following experimental days included **baseline** (BL), **training**, and **retention** (RT) phases (Fig. 4). Half of the participants (randomly allocated) performed the second session in the passive condition and the other half in the active condition. This was reversed in the third session, i.e., participants who performed the second session in the passive condition continued the third session with the active condition and vice versa. The BL and RT were consistent with the iBL session, i.e., the eight stimuli comparisons (*Co*) in Table 2 were presented a total of five times each. We randomized the stimuli presentation order for each phase (BL and RT).

The training phase consisted of 120 trials, grouped into three blocks of 40 trials each. Participants could rest for 2 minutes between training blocks. Each 40-trial block consisted of eight different comparison stimuli presented five times each. The comparison stimuli during training differed from those used in the baseline and retention phases (see section 2.7.3). The stimuli presentation order was randomized for the first training block and repeated for the second and third training blocks.

During training, after each trial, we provided terminal visual feedback to the participants (Fig. 6). This visual feedback consisted of 1) a texture turning green (i.e., the odd texture) and the two others red (i.e., incorrect responses), and 2) black parallel lines along the $x$-axis that represented the location of the texture grooves. Participants had then the opportunity to re-explore the textures as desired.

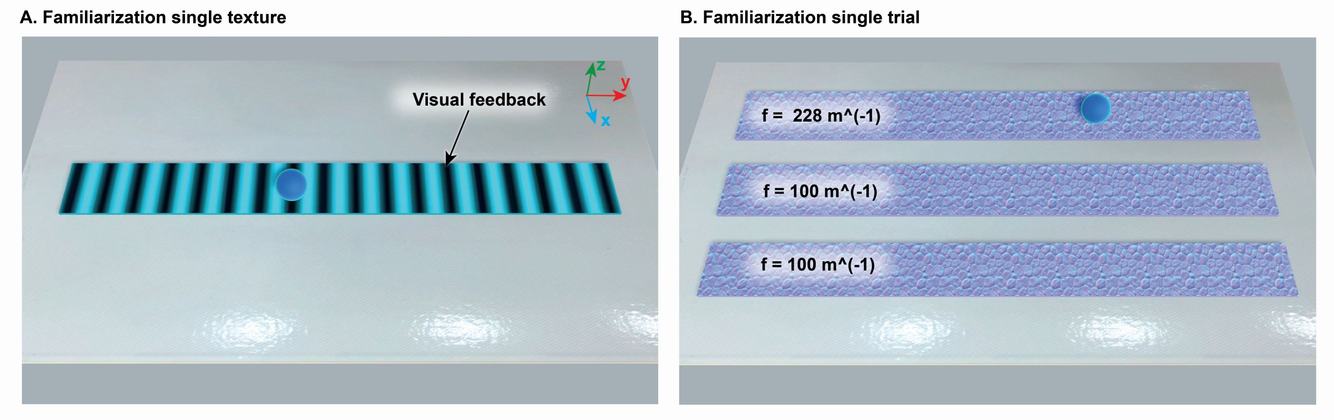


**Figure 5.** Virtual environments during familiarization. **A:** The visual feedback that mapped the grating rendering was shown to help participants understand the haptic stimuli provided during the experiment. The dark color represented grooves, while the light blue color represented texture crests. Participants were encouraged to actively explore the texture as much as desired. **B:** After the single texture familiarization, participants were asked to discriminate the odd texture among three virtual textures with identical visualization. The texture combination used for all participants was 228 $m^{-1}$, 100 $m^{-1}$, 100 $m^{-1}$, respectively.

### 2.7.3 Comparison stimuli during training

After completion of BL, we used a psychometric function to fit the probability of the participants having a correct response in BL based on the stimulus intensity –i.e., the absolute difference between the spatial frequency of the *Co* ($f_{Co}$) and the spatial frequency of the *St* ($f_{St}$), divided by the spatial frequency of the *St*). We employed the logistic function:

$$F(x|\alpha, \beta) = \frac{1}{1 + exp(-(\alpha + \beta x))}, \tag{7}$$

where $\alpha$ and $\beta$ are the *intercept* and *slope* of the logistic function, respectively. Two logistic regressions were computed after the BL data were collected, one for the more coarse textures and a second one for the less coarse textures. Once the logistic functions were fitted to the BL data, we computed the point of subjective equality (PSE) by selecting the probability of a positive response $\pi$ = 0.50 –i.e., the point at which two stimuli are perceived as one– for each participant and for each texture set, i.e., more and less coarse sets, $PSE_{mc}$ and $PSE_{lc}$, respectively). We then used these calculated values to create two new spanned ranges of comparison stimuli that were employed during training, i.e., we adapted the set of *Co* that were employed during training to each participants' performance during BL. Information about this process can be found in the supplementary material 2.

The fitting of the logistic functions did not always converge, i.e., the $PSE$ was not within the range of the BL difference ratio (i.e., between 0.39 (-) and 0.098 (-)). In those cases, the same *Co* stimuli used for BL and RT were employed during training.

## 2.8 Outcome variables

### 2.8.1 Task performance

#### Probability of correct responses

The participants' texture discrimination performance was assessed using the probability of correct responses, calculated following the equation:

**Visual feedback training block**

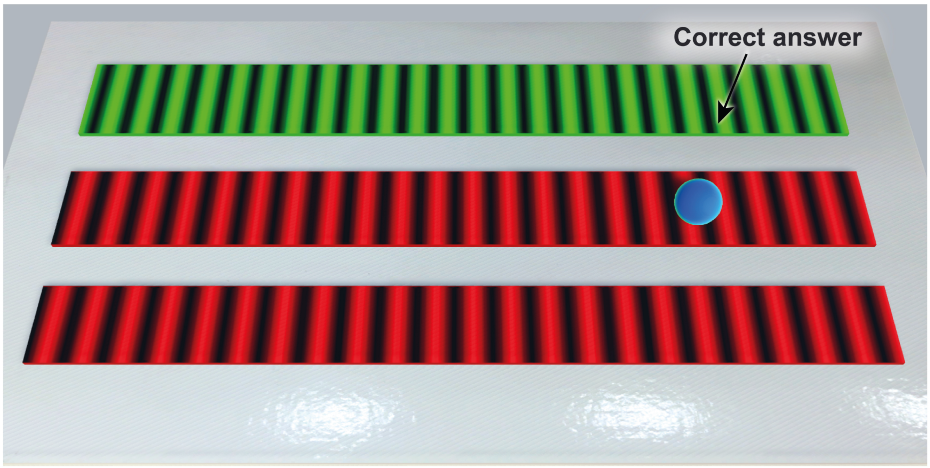


**Figure 6.** The terminal visual feedback provided after each trial during training. The green color represented the correct response (i.e., the odd texture), whereas the red color represented the twin textures (i.e., incorrect responses). The black lines were located along the grooves of the textures. Participants were allowed to re-explore the textures as desired.

$$p_i = \frac{1}{n_i}\sum_{i=1}^{n_i} Y_i. \tag{8}$$

The probability of correct responses was calculated for each comparison stimulus, denoted by the subindex $i \in \{1, 8\}$. The total number of times the response was correct was divided by the number of times $n_i$ the *Co* stimulus was presented. For the iBL, BL, RT, and each training block of 40 trials, the total number of times each stimulus was presented was $n_i = 5$. The probability of correct responses for each stimulus was then averaged across all eight *Co* stimuli (four from more and four from less coarse textures) for each participant.

### Point of subjective equality

The participants' psychometric task performance during BL and RT was also assessed using the point of subjective equality (PSE) (see equation S4). Compared to the probability of correct responses, the PSE is a performance metric that is more robust against participants' guessed responses. We averaged the PSE values for both the more coarse and less coarse textures. We only included PSE values that were within the range of the difference ratios provided during the experiment. The PSE values included were between 0.39 (-) and 0.098 (-). If a participant correctly responded to all trials, i.e., $p_i = 1$, the PSE score for that specific phase (i.e., BL or RT) was set to the minimum value of the difference ratio (i.e., 0.098 (-) during BL and RT). The PSE could not be calculated when the logistic function did not converge. In those cases, we excluded from the data analysis paired cases in which PSE was not calculated either for BL or RT for a participant.

### 2.8.2 Kinematic outcomes

We also evaluated the participants' texture exploratory behavior during BL and RT. In particular, per each trial, we calculated the scanning duration –i.e., the average time participants spent in contact with the three textures and moving faster than 0.01 m/s–, the path length –i.e., the path covered by the end effector over the texture averaged for the three textures– and the mean scanning speed –i.e., the mean end effector speed in $y$-direction.

#### 2.8.3 Motivation outcomes

We assessed the participants' subjective motivation after completing each active and passive iBL in session 1. Participants responded to 12 items selected from four subscales (i.e., Effort/Importance, Perceived Competence, Interest/Enjoyment, and Pressure/Tension) of the original Intrinsic Motivation Inventory (IMI) (Ryan et al., 1990). Three items per subscale were included (see supplementary material 3, table S1 for the selected items). Participants rated each item using a 7-point Likert scale, with 1 indicating "not at all true" and 7 denoting "very true". We averaged the answers of the three items for each subscale.

### 2.9 Data processing and statistical analysis

#### 2.9.1 System reliability

We estimated the system test-retest reliability –i.e., the correlation between two measurements from the same participant under the same conditions at distinct time points (Koo and Li, 2016; Salter et al., 2005)– by comparing the probability of correct responses and PSE scores between iBL (day 1) and BL (day 2) for each condition (passive and active touch).

We used the Intraclass Correlation Coefficient (ICC). The ICC reflects the degree of correlation and agreement between the participants' baseline (day 1 vs. day 2) measurements. The ICC value was estimated using the Python $pingouin.intraclass\_corr$ function and selecting the output of average random raters (McGraw and Wong, 1996), i.e., considering an absolute agreement with multiple measurements. Reliability was considered excellent when $ICC > 0.90$, good when $0.75 < ICC \leq 0.90$, moderate when $0.5 < ICC \leq 0.75$ and poor otherwise (Koo and Li, 2016).

To analyze the ICC for each condition, we allocated participants who performed the active condition on the second session to $ICC_{active}$ group, whereas those who performed the passive condition on the second day were allocated to the $ICC_{passive}$ group. Each group included 18 participants. We then compared the baseline in the second session BL of each condition group to their corresponding iBL in session 1 –note that on day 1, all participants performed an active iBL and a passive iBL. We did not consider BL on day 3 to avoid any training effects from the second day training. Absolute test-retest reliability was visually inspected using Bland-Altman plots for active and passive conditions for the probability of correct responses and the PSE.

#### 2.9.2 Training effects

For each task performance and kinematic outcome, we calculated the mean of the participants' BL and RT trials per condition –i.e., passive and active touch. The normal distribution of the outcome variables was verified using QQ-plots and the Shapiro-Wilk test.

To study the training effects on the task performance and kinematic outcome variables for each condition, we performed repeated-measures one-way ANOVAs –with factor *time*: BL and RL – for data with normal distribution. We analyzed non-normal data with Friedman tests.

We also evaluated the differences in training effects on touch sensibility between conditions –i.e., condition effects: Passive vs. Active– by comparing the pre-post changes (RT-BL) in task performance between active and passive conditions using repeated-measures one-way ANOVAs for data with normal distribution, and Friedman test for non-normal distributed data.

#### 2.9.3 Motivation

To analyze how active and passive conditions affected participants' motivation, we performed separated repeated measures one-way ANOVAs for each subscale –with factor *condition*: Active and Passive– with normal-distributed data, and Friedman test for non-normal distributed data.

The assumption of sphericity was met for all tests. We used the Python package "Pingouin" (Vallat, 2018) for all the statistical tests. Finally, all statistical tests were set at a significance level of $\alpha = 0.05$.

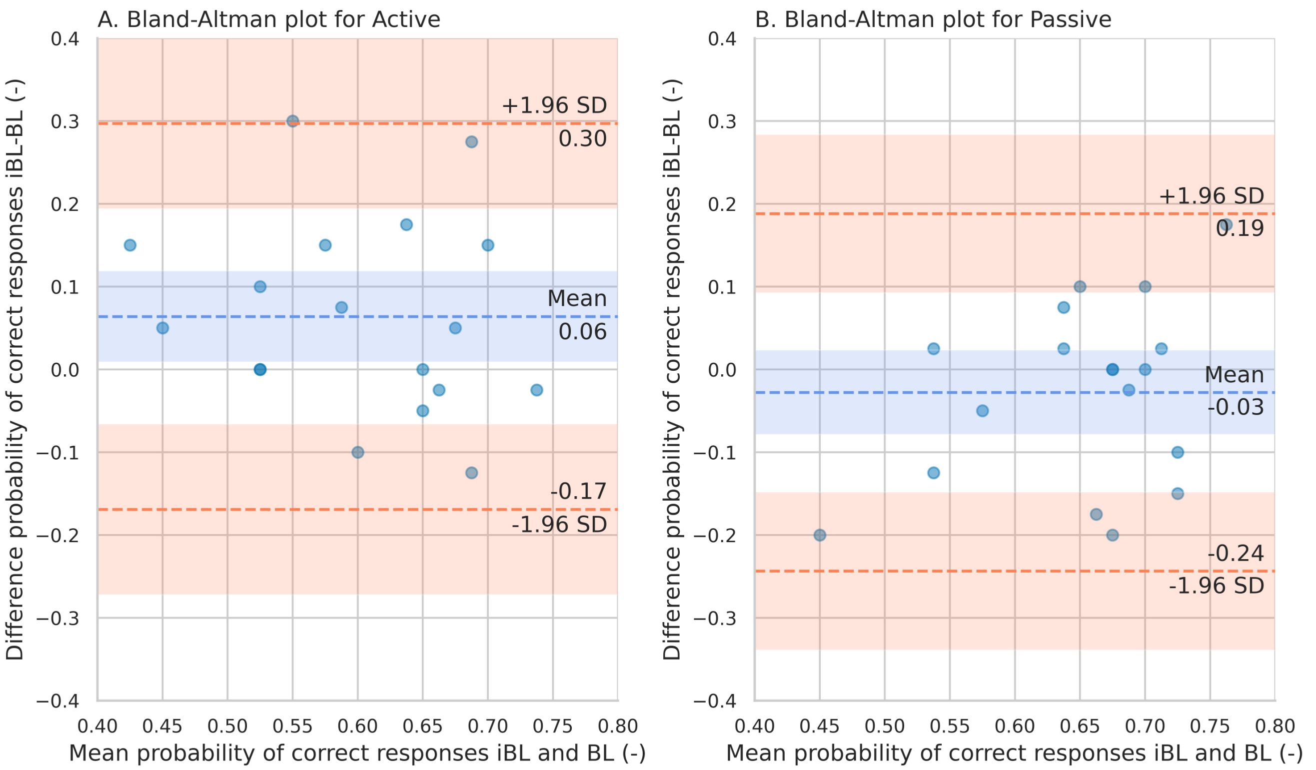


**Figure 7.** Bland-Altman plots for the probability of correct responses in iBL and BL for the Active condition (**A**) and Passive condition (**B**). The blue dashed line represents the mean difference between iBL and BL. The lower and upper orange dashed lines represent the lower and upper 95 % confidence limits, respectively.

## 3 RESULTS

### 3.1 System Reliability

The ICC values for the probability of correct responses were computed for each condition. The active condition ($ICC(2,k) = 0.497,\ \ 95\,\%\,CI\ [-0.170, 0.800],\ \ p = 0.055$), and passive condition ($ICC(2,k) = 0.518,\ \ 95\,\%\,CI\ [-0.260, 0.820],\ \ p = 0.069$) showed poor and moderate reliability, respectively. Similarly, the ICC values for the PSE were computed for each condition. The active ($ICC(2,k) = -1.885,\ \ 95\,\%\,CI\ [-12.530, 0.200],\ \ p = 0.955$) had a poor reliability, whereas the passive condition ($ICC(2,k) = 0.795,\ \ 95\,\%\,CI\ [0.010, 0.950],\ \ p = 0.025$) showed a good reliability.

The Bland-Altman plots for the probability of correct responses and PSE for active and passive conditions are shown in Fig. 7 and Fig. 8, respectively. From Fig. 7A it can be observed that in just one participant the difference between iBL and BL measurements for the probability of correct responses is over the upper bound of the 95 % CI, while the mean of the differences between baselines is around 0.06 (-). From Fig. 7B, it can be seen that the difference between iBL and BL is closer to zero (-0.03 (-)) compared to the active condition. Fig. 8A shows that there was a participant whose PSE difference between iBL and BL for the active condition was outside of the CI. On the contrary, all data points in Fig. 8B for the PSE in passive condition were within the 95 % CI with zero mean.

### 3.2 Task performance

Participants significantly improved the probability of correct responses from BL to RT in the active ($X^2(1,35) = 4.235$, $p = 0.039$; Fig. 9A, Table 3) and passive conditions ($F(1,35) = 15.564$, $p < 0.001$;

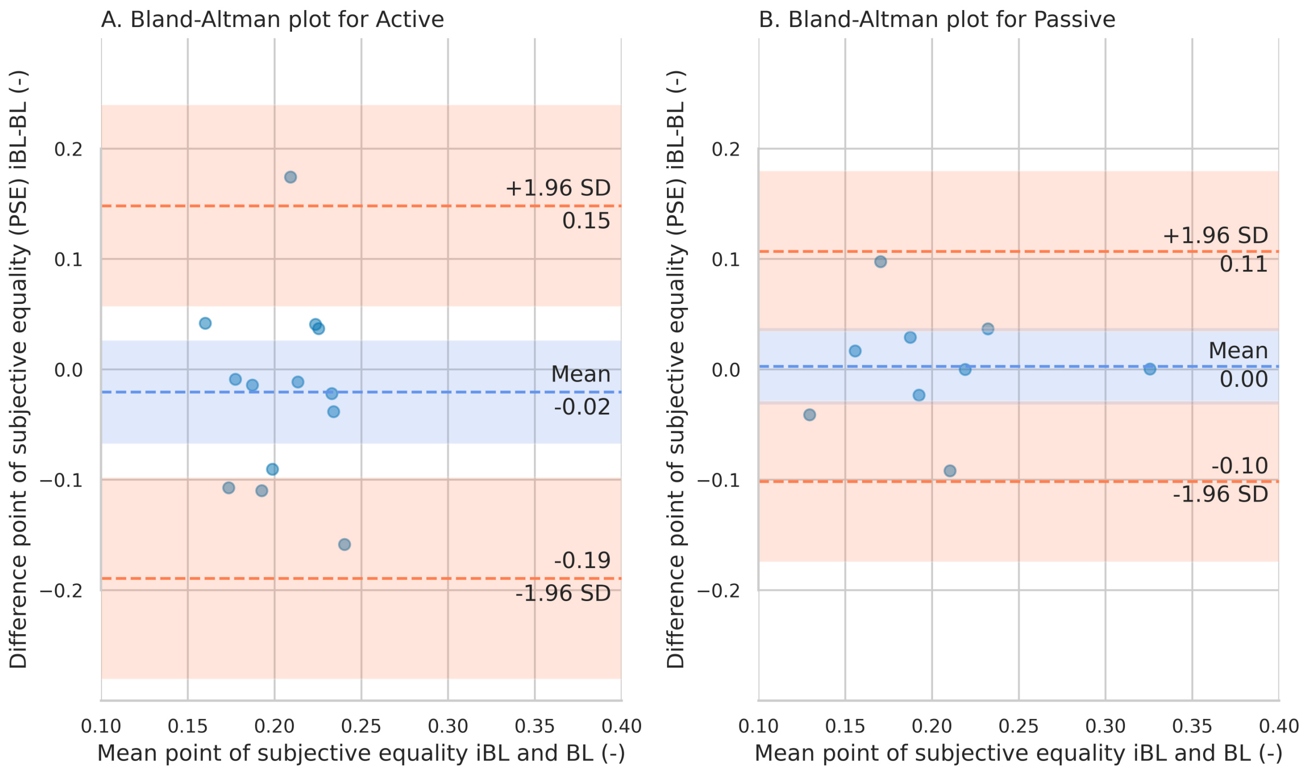


**Figure 8.** Bland-Altman plots for the point of subjective equality in iBL and BL for the Active condition (**A**) and passive condition (**B**). The blue dashed line represents the mean difference between iBL and BL. The lower and upper orange dashed lines represent the lower and upper 95 % confidence limits, respectively.

Fig. 9A, Table 3). We found no significant differences between the active and passive conditions in the pre-post training changes in the probability of correct responses (Table 4).

Participants did not significantly improve their PSE values from BL to RT in any of the conditions (Fig. 9B, Table 3). There were also no significant differences in the RT-BL changes between the conditions (Table 4).

## 3.3 Kinematic outcomes

We did not find significant changes in the scanning duration or path length from BL to RT in any of the conditions, neither we found differences in the RT-BL changes between conditions (Fig. 10A & B, Table 3 & 4). However, we found that participants performed faster exploratory movements after training with both conditions (active: $F(1, 35) = 12.121$, $p = 0.001$; passive: $F(1, 35) = 17.989$, $p < 0.001$; Fig. 10C, Table 3). The increase in scanning speed was significantly higher in the active condition compared to the passive condition ($F(1, 35) = 8.017$, $p = 0.008$, Fig. 10C, Table 4).

## 3.4 Motivation

We did not find significant differences between active and passive conditions in the first experimental session in the IMI subscales Effort/Importance, Interest/Enjoyment, and Pressure/Tension (Table 5).

However, we found a significant difference between active and passive conditions in the Perceived Competence subscale ($F(1, 35) = 9.701$, $p = 0.004$, Table 5).

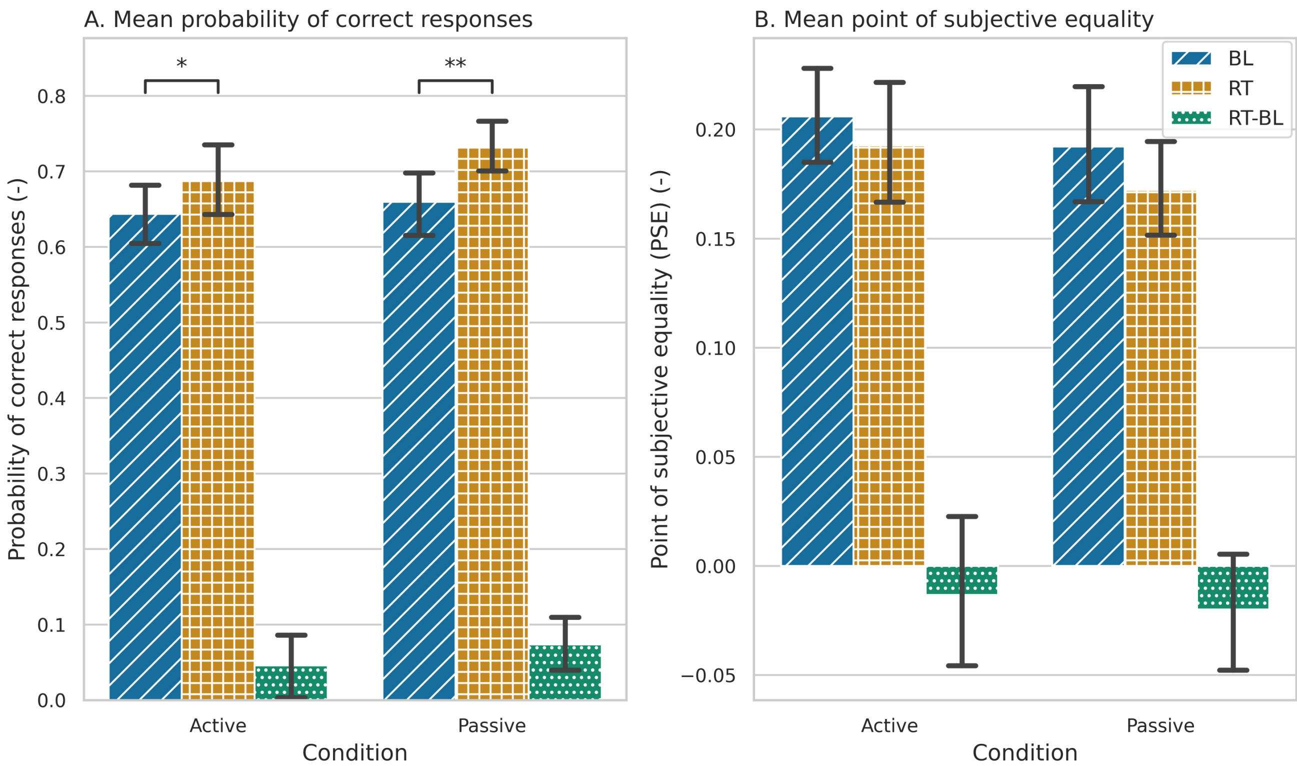


**Figure 9.** Task performance during baseline (BL), retention (RT), and the changes between baseline and retention (RT-BL) for the active and passive conditions. **A**: Probability of correct responses. **B**: Point of subjective equality (PSE). Statistically significant differences are marked by $*(p < 0.05)$, $**(p < 0.001)$. The error bars represent the confidence intervals.

**Table 3.** Overview of results for task performance and kinematic outcome variables in active and passive conditions during the baseline (BL) and retention (RT) phases. Mean and standard deviation (in brackets) are reported when the data were normally distributed and median and mean absolute deviation (in brackets) when the data were non-normal. Statistically significant values are highlighted in bold.

| Independent Variables | Training Effects | | | | | |
|---|---|---|---|---|---|---|
| | Active | | | Passive | | |
| | BL | RT | *p*-value | BL | RT | *p*-value |
| Task performance | | | | | | |
| Probability of correct responses (-) | 0.644 (0.120) | 0.725 (0.116) | **0.039** | 0.660 (0.124) | 0.733 (0.105) | **<0.001** |
| Point of subjective equality (-) | 0.206 (0.058) | 0.193 (0.073) | 0.450 | 0.192 (0.068) | 0.172 (0.053) | 0.163 |
| Kinematic outcomes | | | | | | |
| Scanning duration (s) | 4.418 (1.537) | 4.125 (1.624) | 0.153 | 3.969 (1.569) | 3.954 (1.514) | 0.931 |
| Path length (m) | 0.804 (0.281) | 0.842 (0.338) | 0.356 | 0.820 (0.259) | 0.788 (0.221) | 0.343 |
| Scanning speed (m/s) | 0.175 (0.043) | 0.194 (0.053) | **0.001** | 0.172 (0.009) | 0.176 (0.007) | **<0.001** |

## 4 DISCUSSION

We developed and evaluated a novel virtual reality-based haptic system to assess and train touch discrimination. The novelty of our system is that the haptic rendering forces of the textures are applied tangentially to the hand and do not depend on the normal force that participants exert on the texture surface, as is the case when using real textures. Thus, by disentangling the tangential and normal forces, we aimed to provide more controlled stimuli within and between participants.

Thirty-six healthy participants were asked to discriminate virtual textures using active and passive conditions, i.e., with the robot not guiding their movements vs. the robot guiding their hands. We evaluated

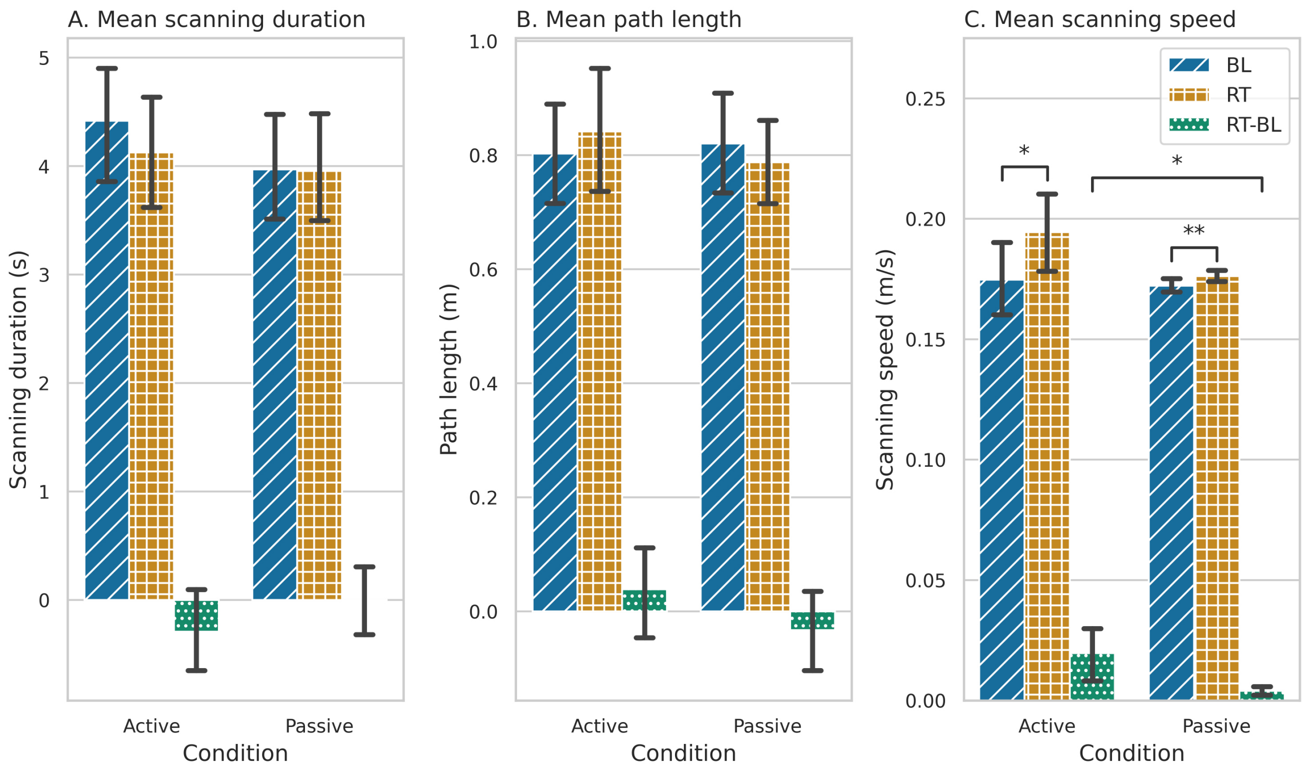


**Figure 10.** Kinematic outcomes during baseline (BL), retention (RT), and the changes between baseline and retention (RT-BL) for the active and passive conditions. **A**: Scanning duration. **B**: Path length. **C**: Scanning speed. Statistically significant differences are marked by $*p < 0.05$, $**p < 0.001$. The error bars represent the confidence intervals.

**Table 4.** Overview of results comparing the pre-post changes from baseline (BL) to retention (RT) in active versus passive conditions. Mean and standard deviation (in brackets) are reported when the data were normally distributed and median and mean absolute deviation (in brackets) when the data were non-normally distributed. Statistically significant values are highlighted in bold.

| Independent Variables | Active vs Passive | | |
|---|---|---|---|
| | Active RT-BL | Passive RT-BL | p-value |
| Task performance | | | |
| Probability of correct responses (-) | 0.046 (0.124) | 0.074 (0.112) | 0.323 |
| Point of subjective equality (-) | 0.013 (0.087) | 0.020 (0.063) | 0.285 |
| Kinematic outcomes | | | |
| Scanning duration (s) | -0.293 (1.203) | -0.015 (1.028) | 0.261 |
| Path length (m) | 0.039 (0.248) | -0.033 (0.206) | 0.151 |
| Scanning speed (m/s) | 0.020 (0.034) | 0.004 (0.006) | **0.008** |

the reliability of our system and the potential to train tactile sensibility in a within-subjects three-session experiment.

## 4.1 The system reliability

We evaluated our system reliability by comparing the baseline tests on two different days. We found that the reliability with respect to the probability of correct responses ranged from poor to moderate in active and passive conditions, respectively. The reliability relative to the PSE differed between conditions; although our system showed good reliability in the passive condition, the reliability in the active condition

**Table 5.** The results for each motivation subscale comparing the differences between initial baseline (iBL) active and initial baseline passive. Mean and standard deviation (in brackets) are reported when the data were normally distributed and median and mean absolute deviation (in brackets) when the data were non-normal. Statistically significant values are highlighted in bold.

| | Active vs Passive | | |
|---|---|---|---|
| Independent Variables | Active | Passive | |
| | iBL | iBL | p-value |
| Motivation | | | |
| Effort/Importance (-) | 5.213 (1.130) | 5.028 (1.018) | 0.108 |
| Perceived Competence (-) | 2.833 (1.085) | 3.296 (1.041) | **0.004** |
| Interest/Enjoyment (-) | 4.574 (1.407) | 4.694 (1.470) | 0.245 |
| Pressure/Tension (-) | 2.000 (1.061) | 2.000 (0.877) | 0.368 |

was rather poor. This is in line with our hypothesis of better reliability in the passive condition as by guiding the movements with the robot, we provided a more controlled texture exploration.

It is recommended that the test-retest ICC value should be at least 0.90 if the system is aimed to assess or evaluate the treatment of a patient (Salter et al., 2005). Our system did not reach excellent reliability in any of the performance metrics evaluated with 36 healthy participants (18 subjects per condition). Yet, several studies have shown poor to moderate reliability for sensory assessment in a healthy population, e.g., when using robotic devices (Rinderknecht et al., 2018), and physical textures (Ofek et al., 2018). It has also been observed lower reliability values in robotic sensory assessments in the unimpared limb compared to the impared limb in stroke patients (Rinderknecht et al., 2018). This has been suggested to be due to the dependency of the ICC on the between subject variability, that in a healthy population is usually small. Thus, we expect that including brain-injured patients in future studies would result in higher reliability values, especially when testing the impaired limb – e.g., excellent reliability in a tactile discrimination test that required stroke patients to discriminate differences in physical finely graded ridged surfaces was observed in an experiment with 35 patients (Carey et al., 1997).

There may be other reasons for the limited reliability observed in our system. First, contrary to Carey et al. (1997), we did not account for a possible lack of familiarization in the first session. Although we did include a short familiarization phase, it only included one trial. Thus, probably between the two sessions, learning occurred. Furthermore, the familiarization was only performed with the active condition. Second, in our three-day experiment, participants experienced both conditions in the first session. Some participants, therefore, first had an active and then passive baseline condition on day one, followed by an active condition on day two, while some participants had an active and then passive baseline on the first day, followed by a passive condition on the second day. Thus, in some cases, there was a baseline with the other condition between two baselines of the same type (iBL and BL) that could have served as “training”, potentially hampering the reliability. Thus, our results support the idea that more extensive familiarization trials should be performed prior to clinical assessments. Finally, the variance of the outcome measurements between sessions may be a result from the spontaneous fluctuations (i.e., noise) of the somatosensory system (Ribeiro et al., 2016).

The ICC has been shown to be sensitive to intra- and inter-subject variability and that reporting the CI for reliability is also important (Germanotta et al., 2018). Yet, reporting ICC values together with the CI alone might provide insufficient information for a reliability analysis (Rankin and Stokes, 1998). Therefore, to extend our system reliability analysis and visually inspect the agreement between the first (iBL) and second measurements (BL), we performed Bland-Altman plots for both performance metrics. Bland-Altman plots on the proportion of correct responses and PSE can be employed to observe the magnitude of measurement error of each test-retest difference (iBL-BL). In the case of the passive condition, the difference between the two measurements was relative low and all the data points laid within the limit of confidence (chosen as 95 %) in both performance metrics. However, in the active condition some data points laid outside the CI boundaries in both performance metrics. Thus, using the robot to guide the movements during passive texture exploration seems a more reliable assessment tool than allowing participants to freely explore. This finding will guide next experiments with brain-injured patients. It seems reasonable to not include active touch during follow-up experiments with patients. This will reduce the duration of future experiments, leveraging more promising techniques while preventing patients from engaging in too exhausting interventions.

### 4.2 Passive and active robotic somatosensory training enhances the discrimination of virtual textures

Consistent with our hypothesis, we found an improvement in virtual texture discrimination after training with both active and passive conditions. Participants significantly improved their tactile acuity of virtual textures, as reflected in a significant increase in the probability of correct responses from baseline to retention after three training blocks of active or passive conditions. Although we also observed a decrease in the PSE after training with active and passive conditions, the difference did not reach significance. This might be due to the lower number of sample points included in the PSE statistical analysis, compared to the larger sample in the probability of correct responses analysis. Several BL and RT trials (active: 22 out of 72, and passive: 30 out of 72) had to be removed from the statistical analysis as the fitting of the logistic function did not converge.

Our results are in line with several studies that evaluated the potential of sensory retraining strategies, e.g., Elangovan et al. (2017); Ballardini et al. (2018); Carey et al. (1993). Most studies to date focused on training proprioception rather than tactile perception. For example, Elangovan et al. and Yeh et al. found enhanced proprioceptive function after an active proprioception retraining intervention that required healthy participants (Elangovan et al., 2017) and stroke patients (Yeh et al., 2021) to balance a virtual ball on a virtual table. Improvements in tactile perception were also found in a few studies that evaluated tactile training interventions in healthy (e.g., Ballardini et al. (2018)) and stroke patients (e.g., Carey et al. (1993)). Ballardini et al. found enhanced tactile sensitivity after a sensory retraining task that required healthy participants to discriminate and replicate skin-brushed stimuli that were applied by the end effector of a robotic device on the palm of their dominant hand (Ballardini et al., 2018). In each training trial, participants were asked to actively move their non-dominant hands and reproduce the stimulus that they previously experienced at several target locations on the palm of their dominant hands. Further, Carey et al. found clinically and statistically significant improvements in the ability to discriminate differences in tactile stimuli after 13–16 weeks of training (Carey et al., 1993). Carey et al. assessed and trained touch sensibility using a set of fine plastic gratings, which differed only by their spatial periods. Textures were presented in sets of three, with and odd texture among two identical ones, and –as in with our experiment– patients were asked to select the odd texture. However, unlike our work, Carey et al. did not adjust the difficulty of the training based on the specific baseline performance of the patients. As a novelty, our sensory retraining strategy followed an adaptive intervention that adjusted the difficulty of the task to meet the specific needs of the participants. However, we did not compare differences on the effect of training with adaptive vs. fixed difficulty levels on tactile discrimination, and therefore, conclusions about the suitability of our adaptive training program cannot be driven.

Carey et al. (1993) and Sathian and Zangaladze (1997) reported that training effects appear to be stimulus-specific and task-specific. Therefore, improvements in texture discrimination due to our intervention may not be transferable to other types of tactile perception tasks –e.g., recognition of haptic letters (Vega-Bermudez et al., 1991). Further, we cannot rule out that the performance improvements observed after our tactile discrimination training may result from a better understanding of the task rather than reflect improvements in individual touch sensibility. The addition of visual feedback after each training trial could have helped participants to better understand the task and, consequently, improve their performance. However, it should be noted that most of the participants trained with spatial frequencies that were different from those presented during baseline and retention as we adapted the training comparison stimuli based on the participants' individual PSE at baseline. More precisely, for the active training sessions, the logistic regression for the less coarse textures converged in 25 of the 36 participants (and therefore assembling a new set of training comparison stimuli), and for the more coarse textures in 23 of 36 participants. For the passive training sessions, the logistic regression converged in 24 of 36 participants for both less and more coarse textures. The fact that participants trained with stimuli different from those presented at baseline and retention minimizes the possibility that our findings reflect stimulus-specific training effects.

Although we found significant differences in the participants' discrimination performance after training, these did not seem to be related to pre-post training changes in the textures exploration behavior. We did not find significant changes in the scanning duration nor the path length on the textures after training. However, we did find a small, albeit significant, increase in the scanning speed after training for both active and passive conditions. Yet, several researchers showed that tactile perception is invariant to changes in exploration speed (Lamb, 1983; Lederman, 1983; Boundy-Singer et al., 2017).

### 4.3 No differences between active and passive conditions in training effects on touch sensibility

As hypothesized, we did not find differences in the effect of active vs. passive conditions on pre-post changes regarding the probability of correct responses nor PSE values. Our results are consistent with previous findings, which revealed no differences between active and passive conditions (Lederman, 1981; Vega-Bermudez et al., 1991; Lamb, 1983). Lederman (1981) found no differences between active and passive conditions when their participants estimated magnitudes of the roughness of metal gratings (Lederman, 1981). Vega-Bermudez et al. (1991) found no differences in tactile recognition of letters between active and passive conditions (Vega-Bermudez et al., 1991). Further, Lamb (1983) found no differences between active and passive conditions in discriminating between plastic strips with raised dots (Lamb, 1983). However, contrary to our experiment, in all these studies (Lederman, 1981; Vega-Bermudez et al., 1991; Lamb, 1983) participants received the passive stimulation with their arms and hands immobilized, e.g., by using a drum stimulation. Thus, the active conditions did not provide an advantage over the passive condition due to the addition of, for example, kinesthetic cues associated with the active movement.

Vega-Bermudez et al. reported that the majority of experiments that found differences between active and passive conditions employed tasks in which proprioceptive information represented a critical component for the success of the sensory task (Vega-Bermudez et al., 1991). In our study, we compared the active condition to a passive condition in which the robot guided the participants' hands and thus also provided proprioceptive information during the tactile discrimination task. Our findings, therefore, extend previous studies by suggesting that there are no differences between active and passive touch even when controlling for kinesthetic information between conditions.

A potential rationale behind the lack of differences between passive and active conditions is that in both conditions participants explored the textures using indirect contact, i.e., through the robot end-effector. In the haptics field, tactile texture perception has been investigated in two modes of touch contact: direct and indirect contact. Direct contact refers to participants touching (or being touched by) an object with (in) their bare skin –e.g., their fingertips–, whereas indirect contact refers to touching objects using an intermediary link –e.g., gloves, tools, or robotic devices (Lederman et al., 1999; Klatzky and Lederman, 2006). During direct touch, the roughness of the texture is spatially coded by the central nervous system using tactile information sensed by mechanoreceptors on the glabrous skin. Temporal cues are also temporally coded using vibration cues (Ryan et al., 2021). The weight of the relative use of spatial and temporal cues seems to vary depending on the spatial period of the texture (Klatzky and Lederman, 2006). However, when people explore a texture indirectly through a tool, the roughness perception of textures seems to be mainly coded via temporal cues (Klatzky and Lederman, 2006). During indirect touch, the spatial information of the texture is usually no longer available as the spatial deformation of the skin relates to the shape of the probe rather than the properties of the scanned texture (Yoshioka et al., 2007). Consequently, in our indirect touch experiment, participants received similar vibratory cues to perform the sensory task under both conditions, which may explain the lack of significant differences between the active and passive touch conditions.

### 4.4 Participants' motivation

We hypothesized that haptic guidance during training would improve participants' perceived competence. As expected, we found that participants reported significantly higher competence during passive than active exploration, which may be a positive indicator for using passive exploration with brain-injured patients. The robotic guidance during training might allow participants to focus on the sensory input, without the need to think about how to explore the textures (instead, the robot takes the "responsibility"), therefore, increasing their subjectively reported perceived competence. Physically and cognitively impaired patients could likely even further benefit from this guidance, as it allows them to focus on the task instead of the exploration strategy, potentially increasing the effects of sensory training. We also hypothesized that during active exploration, participants would report higher levels in effort, pressure, and interest/enjoyment than during passive exploration, as exploring the textures themselves may make the training more engaging, but may also be associated with the challenge to choose an optimal exploration strategy. Contrary to our hypothesis, we found no differences between active and passive conditions on the IMI subscales Effort/Importance, Interest/Enjoyment, and Pressure/Tension. This result may indicate that participants remained engaged in the task under both conditions, even though they did not need to actively move along the textures during training with the passive condition. However, overall, we found high Effort/Importance reportings, whereas the perceived competence was quite low, indicating that the experiment was rather challenging. Hence, in future experiments and especially with brain-injured patients, it is important to lower the difficulty of the task, e.g., by reducing the number of stimulus comparisons or augmenting the interstimulus distance to simplify the task. Additionally, attention may influence the perception of virtual textures.

### 4.5 Study limitations and future work

Our study suffers from several limitations. First, participants were allowed to freely explore the textures, without any time limitation. We decided to allow for "free" exploration while measuring the kinematic data –i.e., scanning duration, end effector path length, and scanning speed– during exploration to evaluate the effect of training with the different conditions on the exploration strategies after training. Second, the long duration of the training blocks might have triggered the so-called paresthesia, i.e., the sensation experienced as a numbness or tingling sensation on the skin (O'Tool, 2017) and a result of excessive sensory stimulation without long enough resting periods. Yet, we still found improvements in the discrimination of the virtual textures right after the training finished (in short-term retention). Third, contrary to (Carey et al., 1993) and (Ballardini et al., 2018), we did not blindfold the participants, i.e., we did not occlude the vision of the hands and/or robot. Instead, we provided visual feedback using a virtual environment to motivate participants. Yet, the use of the virtual environment to provide visual feedback on the participant's hand position with respect to the virtual textures, which was located in a different space than the robot/hand movements, probably limited participants to look at their own hands as they were focused on the screen visualization.

Although our system showed moderate to good reliability values with 18 healthy participants per condition, the reliability evaluation would benefit from including more healthy participants, and especially brain-injured patients. It has been observed lower reliability values in healthy compared to stroke patients (Rinderknecht et al., 2018; Ofek et al., 2018), and thus, we expect to observe higher reliability values when assessing texture discrimination in a brain-injured population. Moreover, in future experiments, we plan to test the feasibility and acceptability of our system with brain-injured population, as the majority of robotic devices in neurorehabilitation are tested with healthy population instead of patients. Yet, several changes must be performed to our protocol before bringing our system to the clinics. First, we suggest performing the experiment with brain-injured patients only with the passive condition, as our results with healthy participants suggest that training effects would not differ between conditions and the passive condition shows higher reliability. Second, the training duration should be reduced to prevent paresthesia, e.g., by reducing the number of trials and repetitions per stimuli. Third, we may need to adapt the location of the monitor to account for patients with visual neglect, as it was found that visual neglect might interfere with the assessment of somatosensory impairment probably due to attention deficit (Meyer et al., 2016). Further, we may need to adjust the level of discrimination difficulty accordingly to the patients' deficits. Finally, to increase the system reliability, we suggest increasing the familiarization time, controlling the time between sessions (with a minimum of 24 h between sessions) and the time of the day the intervention is delivered.

## 5 CONCLUSION

Despite the high prevalence of sensory deficits after stroke, somatosensory treatment is currently neglected in neurorehabilitation interventions. Crucially, there is a lack of high-quality research demonstrating benefits of somatosensory (re)training on stroke recovery and a need for reliable quantitative assessments of sensorimotor deficits. Further, to date, somatosensory assessments and interventions are labor-intensive and require therapists to guide the paretic limbs of the patients. The goal of this study was to develop and evaluate the reliability of a novel virtual reality-based robotic texture discrimination task that allows to assess and train touch sensibility with (i.e., passively) and without (i.e., actively) guidance for potential clinical application.

In our sample of healthy young participants with expected low between-subject variability, our system showed poor (in active condition) to moderate/good (in passive condition) reliability. Furthermore, we found that participants significantly improved their task performance after training and that these training effects did not differentiate between active and robotic-guided passive exploration. Similarly, both conditions did not differ in motivation, except that passive touch sensibility training was associated with increased perceived competence.

Together, our virtual reality-based robotic haptic system may be a key asset for the evaluation and retraining of sensory loss with minimal supervision, especially for brain-injured patients who require guidance to move their limbs.

## CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## AUTHOR CONTRIBUTIONS

EVO, EAA, KB, and LM-C designed the study and wrote the manuscript. EVO performed the analysis of the dataset. EVO prepared the experimental setup and programmed the virtual environment and robot. EVO collected the experimental data. All authors edited and revised the manuscript and approved the submitted version.

## FUNDING

This work was supported by the Swiss National Science Foundation through the grant PP00P2 163800. This work was also supported by SENACYT and IFARHU, the Panamanian Government.

## ACKNOWLEDGMENTS

We would like to thank all members of the Motor Learning and Neurorehabilitation Laboratory for their kind support during this project, especially to Dr. Özhan Özen and his support in the development of the experimental software, and Raphael Rätz for helping us designing and 3D-printing the robot end effector.

The authors also thank Prof. René Müri for his advice regarding the experimental setup prior to data collection. We would also like to thank Dr. Yasemin Vardar for the valuable discussions during the document drafting. Finally, a sincere gratitude goes to our participants.

## REFERENCES

Adamson, J., Beswick, A., and Ebrahim, S. (2004). Is stroke the most common cause of disability? *Journal of Stroke and Cerebrovascular Diseases* 13, 171–177. doi:10.1016/j.jstrokecerebrovasdis.2004.06.003

Ballardini, G., Carlini, G., Giannoni, P., Scheidt, R. A., Nisky, I., and Casadio, M. (2018). Tactile-STAR: A Novel Tactile STimulator And Recorder System for Evaluating and Improving Tactile Perception. *Frontiers in Neurorobotics* 12, 12. doi:10.3389/fnbot.2018.00012

Basalp, E., Wolf, P., and Marchal-Crespo, L. (2021). Haptic Training: Which Types Facilitate (re)Learning of Which Motor Task and for Whom? Answers by a Review. *IEEE Transactions on Haptics* 14, 722–739. doi:10.1109/TOH.2021.3104518

Bi, J. (2015). *Sensory discrimination tests and measurements: statistical principles, procedures, and tables* (Chichester, West Sussex: Wiley Blackwell). OCLC: 1039309670

Boundy-Singer, Z. M., Saal, H. P., and Bensmaia, S. J. (2017). Speed invariance of tactile texture perception. *Journal of Neurophysiology* 118, 2371–2377. doi:10.1152/jn.00161.2017

Bryden, M. (1977). Measuring handedness with questionnaires. *Neuropsychologia* 15, 617–624. doi:10.1016/0028-3932(77)90067-7

Campion, G. and Hayward, V. (2005). Fundamental Limits in the Rendering of Virtual Haptic Textures. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (Pisa, Italy: IEEE), 263–270. doi:10.1109/WHC.2005.61

Cappello, L., Elangovan, N., Contu, S., Khosravani, S., Konczak, J., and Masia, L. (2015). Robot-Aided Assessment of Wrist Proprioception. *Frontiers in Human Neuroscience* 9

Carey, L., Macdonell, R., and Matyas, T. A. (2011). SENSe: Study of the Effectiveness of Neurorehabilitation on Sensation: A Randomized Controlled Trial. *Neurorehabilitation and Neural Repair* 25, 304–313. doi:10.1177/1545968310397705

Carey, L. M., Matyas, T. A., and Oke, L. E. (1993). Sensory loss in stroke patients: Effective training of tactile and proprioceptive discrimination. *Archives of Physical Medicine and Rehabilitation* 74, 602–611. doi:10.1016/0003-9993(93)90158-7

Carey, L. M., Oke, L. E., and Matyas, T. A. (1997). Impaired Touch Discrimination After Stroke: A Quantiative Test. *Neurorehabilitation and Neural Repair* 11, 219–232. doi:10.1177/154596839701100404

Carlsson, H., Rosén, B., Björkman, A., Pessah-Rasmussen, H., and Brogårdh, C. (2021). SENSory relearning of the UPPer limb (SENSUPP) after stroke: development and description of a novel intervention using the TIDieR checklist. *Trials* 22, 430. doi:10.1186/s13063-021-05375-6

Chia, F. S., Kuys, S., and Low Choy, N. (2019). Sensory retraining of the leg after stroke: systematic review and meta-analysis. *Clinical Rehabilitation* 33, 964–979. doi:10.1177/0269215519836461

Cholewiak, S., Kwangtaek Kim, Tan, H., and Adelstein, B. (2010). A Frequency-Domain Analysis of Haptic Gratings. *IEEE Transactions on Haptics* 3, 3–14. doi:10.1109/TOH.2009.36

[Dataset] Christensen, R. H. B. and Brockhoff, P. B. (2020). sensr—an r-package for sensory discrimination. R package version 1.5-2. http://www.cran.r-project.org/package=sensR/

Colgate, J. and Brown, J. (1994). Factors affecting the Z-Width of a haptic display. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation* (San Diego, CA, USA: IEEE Comput. Soc. Press), 3205–3210. doi:10.1109/ROBOT.1994.351077

Colombo, R., Pisano, F., Mazzone, A., Delconte, C., Micera, S., Carrozza, M. C., et al. (2007). Design strategies to improve patient motivation during robot-aided rehabilitation. *Journal of NeuroEngineering and Rehabilitation* 4, 3. doi:10.1186/1743-0003-4-3

Doyle, S., Bennett, S., Fasoli, S. E., and McKenna, K. T. (2010). Interventions for sensory impairment in the upper limb after stroke. *Cochrane Database of Systematic Reviews* doi:10.1002/14651858.CD006331.pub2

Elangovan, N., Cappello, L., Masia, L., Aman, J., and Konczak, J. (2017). A robot-aided visuo-motor training that improves proprioception and spatial accuracy of untrained movement. *Scientific Reports* 7, 17054. doi:10.1038/s41598-017-16704-8

Germanotta, M., Cruciani, A., Pecchioli, C., Loreti, S., Spedicato, A., Meotti, M., et al. (2018). Reliability, validity and discriminant ability of the instrumental indices provided by a novel planar robotic device for upper limb rehabilitation. *Journal of NeuroEngineering and Rehabilitation* 15, 39. doi:10.1186/s12984-018-0385-8

Gescheider, G. A. (1997). *Psychophysics: the fundamentals* (Mahwah, N.J: L. Erlbaum Associates), 3rd ed edn.

Handelzalts, S., Ballardini, G., Avraham, C., Pagano, M., Casadio, M., and Nisky, I. (2021). Integrating Tactile Feedback Technologies Into Home-Based Telerehabilitation: Opportunities and Challenges in Light of COVID-19 Pandemic. *Frontiers in Neurorobotics* 15, 617636. doi:10.3389/fnbot.2021.617636

Kenzie, J. M., Semrau, J. A., Hill, M. D., Scott, S. H., and Dukelow, S. P. (2017). A composite robotic-based measure of upper limb proprioception. *Journal of NeuroEngineering and Rehabilitation* 14, 114. doi:10.1186/s12984-017-0329-8

Klatzky, R. L. and Lederman, S. J. (2006). The perceived roughness of resistive virtual textures: I. rendering by a force-feedback mouse. *ACM Transactions on Applied Perception* 3, 1–14. doi:10.1145/1119766.1119767

Koo, T. K. and Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* 15, 155–163. doi:10.1016/j.jcm.2016.02.012

Lamb, G. D. (1983). Tactile discrimination of textured surfaces: psychophysical performance measurements in humans. *The Journal of Physiology* 338, 551–565. doi:10.1113/jphysiol.1983.sp014689

Lederman, S. J. (1981). The perception of surface roughness by active and passive touch. *Bulletin of the Psychonomic Society* 18, 253–255. doi:10.3758/BF03333619

Lederman, S. J. (1983). Tactual roughness perception: Spatial and temporal determinants. *Canadian Journal of Psychology/Revue canadienne de psychologie* 37, 498–511. doi:10.1037/h0080750

Lederman, S. J. and Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception & Psychophysics* 71, 1439–1459. doi:10.3758/APP.71.7.1439

Lederman, S. J., Klatzky, R. L., Hamilton, C. L., and Ramsay, G. I. (1999). Perceiving Surface Roughness via a Rigid Probe: Effects of Exploration Speed and Mode of Touch. *Haptics-e, The electronic journal of haptics research*

Lezkan, A., Metzger, A., and Drewing, K. (2018). Active Haptic Exploration of Softness: Indentation Force Is Systematically Related to Prediction, Sensation and Motivation. *Frontiers in Integrative Neuroscience* 12

Lincoln, N., Crow, J., Jackson, J., Waters, G., Adams, S., and Hodgson, P. (1991). The unreliability of sensory assessments. *Clinical Rehabilitation* 5, 273–282. doi:10.1177/026921559100500403

Magee, L. E. and Kennedy, J. M. (1980). Exploring pictures tactually. *Nature* 283, 287–288. doi:10.1038/283287a0

McGraw, K. O. and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1, 30–46. doi:10.1037/1082-989X.1.1.30

Meyer, S., De Bruyn, N., Lafosse, C., Van Dijk, M., Michielsen, M., Thijs, L., et al. (2016). Somatosensory Impairments in the Upper Limb Poststroke: Distribution and Association With Motor Function and Visuospatial Neglect. *Neurorehabilitation and Neural Repair* 30, 731–742. doi:10.1177/1545968315624779

Meyer, S., Karttunen, A. H., Thijs, V., Feys, H., and Verheyden, G. (2014). How Do Somatosensory Deficits in the Arm and Hand Relate to Upper Limb Impairment, Activity, and Participation Problems After Stroke? A Systematic Review. *Physical Therapy* 94, 1220–1231. doi:10.2522/ptj.20130271

Ofek, H., Alperin, M., Knoll, T., Livne, D., and Laufer, Y. (2018). Assessment of texture discrimination ability at the sole of the foot in subjects with chronic stroke compared with young and elderly subjects with no neurological deficits: a reliability and validity study. *Disability and Rehabilitation* 40, 1960–1966. doi:10.1080/09638288.2017.1323019

O'Tool, M. T. (ed.) (2017). *Mosby's pocket dictionary of medicine, nursing & health professions* (St. Louis, Missouri: Elsevier), eighth edition edn. OCLC: ocn930664561

Pertovaara, A., Helminen, R.-R., and Mansikka, H. (1994). The movement-induced modulation in discriminability between cutaneous nonpainful stimuli depends on test stimulus intensity. *Experimental Brain Research* 101, 506–512. doi:10.1007/BF00227343

Pumpa, L. U., Cahill, L. S., and Carey, L. M. (2015). Somatosensory assessment and treatment after stroke: An evidence-practice gap. *Australian Occupational Therapy Journal* 62, 93–104. doi:10.1111/1440-1630.12170

Rankin, G. and Stokes, M. (1998). Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clinical Rehabilitation* 12, 187–199. doi:10.1191/026921598672178340

Ribeiro, M. J., Paiva, J. S., and Castelo-Branco, M. (2016). Spontaneous Fluctuations in Sensory Processing Predict Within-Subject Reaction Time Variability. *Frontiers in Human Neuroscience* 10

Richardson, B. L., Wuillemin, D. B., and Mackintosh, G. J. (1981). Can passive touch be better than active touch? A comparison of active and passive tactile maze learning. *British Journal of Psychology* 72, 353–362. doi:10.1111/j.2044-8295.1981.tb02194.x

Rinderknecht, M. D., Lambercy, O., Raible, V., Büsching, I., Sehle, A., Liepert, J., et al. (2018). Reliability, validity, and clinical feasibility of a rapid and objective assessment of post-stroke deficits in hand proprioception. *Journal of NeuroEngineering and Rehabilitation* 15, 47. doi:10.1186/s12984-018-0387-6

Rowe, J. B., Chan, V., Ingemanson, M. L., Cramer, S. C., Wolbrecht, E. T., and Reinkensmeyer, D. J. (2017). Robotic Assistance for Training Finger Movement Using a Hebbian Model: A Randomized Controlled Trial. *Neurorehabilitation and Neural Repair* 31, 769–780. doi:10.1177/1545968317721975

Ryan, C. P., Bettelani, G. C., Ciotti, S., Parise, C., Moscatelli, A., and Bianchi, M. (2021). The interaction between motion and texture in the sense of touch. *Journal of Neurophysiology* 126, 1375–1390. doi:10.1152/jn.00583.2020

Ryan, R. M., Connell, J. P., and Plant, R. W. (1990). Emotions in nondirected text learning. *Learning and Individual Differences* 2, 1–17. doi:10.1016/1041-6080(90)90014-8

Salter, K., Jutai, J., Teasell, R., Foley, N., Bitensky, J., and Bayley, M. (2005). Issues for selection of outcome measures in stroke rehabilitation: ICF activity. *Disability and Rehabilitation* 27, 315–340. doi:10.1080/09638280400008545

Sathian, K. and Zangaladze, A. (1997). Tactile learning is task specific but transfers between fingers. *Perception & Psychophysics* 59, 119–128. doi:10.3758/BF03206854

Schabrun, S. and Hillier, S. (2009). Evidence for the retraining of sensation after stroke: a systematic review. *Clinical Rehabilitation* 23, 27–39. doi:10.1177/0269215508098897

Serrada, I., Hordacre, B., and Hillier, S. L. (2019). Does Sensory Retraining Improve Sensation and Sensorimotor Function Following Stroke: A Systematic Review and Meta-Analysis. *Frontiers in Neuroscience* 13, 402. doi:10.3389/fnins.2019.00402

Sommerfeld, D. K. and von Arbin, M. H. (2004). The impact of somatosensory function on activity performance and length of hospital stay in geriatric patients with stroke. *Clinical Rehabilitation* 18, 149–155. doi:10.1191/0269215504cr710oa

Symmons, M. A., Richardson, B. L., and Wuillemin, D. B. (2004). Active versus passive touch : Superiority depends more on the task than the mode. *Touch, blindness, and neuroscience* , 179–185

Taylor, S. C., McKenna, C., Kent, J., Jess, M.-A., Robinson, J. R., Dixon, J., et al. (2021). The effect of sensory discrimination training on sensorimotor performance in individuals with central neurological conditions: A systematic review. *British Journal of Occupational Therapy* 84, 461–473. doi:10.1177/03080226211018150

Turville, M. L., Cahill, L. S., Matyas, T. A., Blennerhassett, J. M., and Carey, L. M. (2019). The effectiveness of somatosensory retraining for improving sensory function in the arm following stroke: a systematic review. *Clinical Rehabilitation* 33, 834–846. doi:10.1177/0269215519829795

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software* 3, 1026. doi:10.21105/joss.01026

Van Doorn, G. H., Dubaj, V., Wuillemin, D. B., Richardson, B. L., and Symmons, M. A. (2012). Cognitive Load Can Explain Differences in Active and Passive Touch. In *Haptics: Perception, Devices, Mobility, and Communication*, eds. D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, P. Isokoski, and J. Springare (Berlin, Heidelberg: Springer Berlin Heidelberg), vol. 7282. 91–102. doi:10.1007/978-3-642-31401-8_9

Vega-Bermudez, F., Johnson, K. O., and Hsiao, S. S. (1991). Human tactile pattern recognition: active versus passive touch, velocity effects, and patterns of confusion. *Journal of Neurophysiology* 65, 531–546. doi:10.1152/jn.1991.65.3.531

Wenk, N., Jordi, M. V., Buetler, K. A., and Marchal-Crespo, L. (2022). Hiding Assistive Robots During Training in Immersive VR Does Not Affect Users' Motivation, Presence, Embodiment, Performance, Nor Visual Attention. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30, 390–399. doi:10.1109/TNSRE.2022.3147260

Yeh, I.-L., Holst-Wolf, J., Elangovan, N., Cuppone, A. V., Lakshminarayan, K., Capello, L., et al. (2021). Effects of a robot-aided somatosensory training on proprioception and motor function in stroke survivors. *Journal of NeuroEngineering and Rehabilitation* 18, 77. doi:10.1186/s12984-021-00871-x

Yilmazer, C., Boccuni, L., Thijs, L., and Verheyden, G. (2019). Effectiveness of somatosensory interventions on somatosensory, motor and functional outcomes in the upper limb post-stroke: A systematic review and meta-analysis. *NeuroRehabilitation* 44, 459–477. doi:10.3233/NRE-192687

Yoshioka, T., Bensmaïa, S. J., Craig, J. C., and Hsiao, S. S. (2007). Texture perception through direct and indirect touch: An analysis of perceptual space for tactile textures in two modes of exploration. *Somatosensory & Motor Research* 24, 53–70. doi:10.1080/08990220701318163

Zandvliet, S. B., Kwakkel, G., Nijland, R. H. M., van Wegen, E. E. H., and Meskers, C. G. M. (2020). Is Recovery of Somatosensory Impairment Conditional for Upper-Limb Motor Recovery Early After Stroke? *Neurorehabilitation and Neural Repair* 34, 403–416. doi:10.1177/1545968320907075

Zbytniewska, M., Kanzler, C. M., Jordan, L., Salzmann, C., Liepert, J., Lambercy, O., et al. (2021). Reliable and valid robot-assisted assessments of hand proprioceptive, motor and sensorimotor impairments after stroke. *Journal of NeuroEngineering and Rehabilitation* 18, 115. doi:10.1186/s12984-021-00904-5

# Supplementary Material

## 1 THE CYCLOIDAL MOTION LAW

The reference trajectory –defined by $\ddot{y_R}$, $\dot{y_R}$, and $y_R$– was obtained following the cycloidal motion law (described in Biagiotti and Melchiorri (2008), page 44):

$$y_R(t) = y_{R0} + dS\left(\frac{t}{T} - \frac{\sin\left(2\pi\frac{t}{T}\right)}{2\pi}\right), \tag{S1}$$

from which

$$\begin{aligned} \dot{y_R}(t) &= \frac{dS}{T}\left(1 - \cos\left(2\pi\frac{t}{T}\right)\right) \\ \ddot{y_R}(t) &= \frac{2\pi dS}{T^2}\left(\sin\left(2\pi\frac{t}{T}\right)\right), \end{aligned} \tag{S2}$$

where $y_{R0}$ is the $y$-position of the end effector when the participant entered or landed into the texture. Participants could enter the texture from any of the four sides of each rectangular texture or lift the robot end effector and land on top of the texture. The displacement $dS$ was calculated each time the participant entered the texture as the maximum distance between $y_{R0}$ and the farthest short-edge of the texture (Fig. 3 in the main manuscript). Thus, participants were guided in the $y$-direction towards the end of the texture that was farthest away from the initial position. The $t$ represented the internal clock of the robot controller with a sampling rate of 4 kHz and $T$ was set to 1 second (i.e., the time needed to finish the movement). By using the cycloidal motion law we could ensure a smooth trajectory reference each time a participant entered a texture.

## 2 ADJUSTED COMPARISON STIMULI FOR TRAINING

Once the logistic functions were fitted to the BL data, we computed the point of subjective equality (PSE) by selecting the probability of a positive response $\pi = 0.50$ –i.e., the point at which two stimuli are perceived as one– for each participant and for each texture set, i.e., more and less coarse sets, $PSE_{mc}$ and $PSE_{lc}$, respectively). The PSEs were calculated from the inverse of the logistic function as:

$$x = F^{-1}(\pi|\alpha,\beta) = \frac{1}{\beta}\left(\log\frac{\pi}{1-\pi} - \alpha\right) \tag{S3}$$

$$PSE = F^{-1}(0.5|\alpha,\beta) = \frac{1}{\beta}\left(\log\frac{0.5}{0.5} - \alpha\right) = -\frac{\alpha}{\beta}. \tag{S4}$$

We used the $PSE_{mc}$ and $PSE_{lc}$ calculated for each participant to create two new spanned ranges of comparison stimuli that were employed during training, i.e., we adapted the set of *Co* that were employed during training to each participants' performance during BL.

We used the mathematical expressions in equations S5 to S8 to set the new frequency difference between consecutive comparison stimuli (fixed to 16 $m^{-1}$ during BL and RT) for the more coarse textures $f_{mc}$ and less coarse textures $f_{lc}$.

$$fpse_{mc} = |f_{St} - PSE_{mc} \cdot f_{St}| \tag{S5}$$

$$fpse_{lc} = |f_{St} + PSE_{lc} \cdot f_{St}| \tag{S6}$$

$$f_{mc} = \frac{|f_{St} - fpse_{mc}|}{n} \tag{S7}$$

$$f_{lc} = \frac{|f_{St} - fpse_{lc}|}{n}, \tag{S8}$$

where $n$ is the number of new comparison stimuli employed during training per coarse type, that was fixed again to $n = 4$. Thus, the new set of comparison stimuli for the more coarse textures was set to:

$$f_{Co_{more\ coarse}} = [f_{St} - f_{mc}, f_{St} - 2f_{mc}, f_{St} - 3f_{mc}, f_{St} - 4f_{mc}] \tag{S9}$$

while the new set of comparison stimuli for the less coarse textures was set to:

$$f_{Co_{less\ coarse}} = [f_{St} + f_{mc}, f_{St} + 2f_{mc}, f_{St} + 3f_{mc}, f_{St} + 4f_{mc}]. \tag{S10}$$

## 3 MOTIVATION QUESTIONNAIRE

**Table S1.** Subscales and items selected from the "Intrinsic Motivation Questionnaire" (Ryan et al., 1990).

| Subscale | Questions |
|---|---|
| Effort/Importance | I tried very hard on this activity. |
| | I put a lot of effort into this. |
| | It was important to me to do well at this task. |
| Perceived Competence | I am satisfied with my performance at this task. |
| | I think I am pretty good at this activity. |
| | I was pretty skilled at this activity. |
| Interest/Enjoyment | The task was fun to do. |
| | I thought this activity was quite enjoyable. |
| | I would describe this activity as very interesting. |
| Pressure/Tension | I felt very tense while doing this activity. |
| | I felt pressured while doing these. |
| | I was anxious while working on this task. |

## REFERENCES

Biagiotti, L. and Melchiorri, C. (2008). *Trajectory planning for automatic machines and robots* (Berlin Heidelberg: Springer)

Ryan, R. M., Connell, J. P., and Plant, R. W. (1990). Emotions in nondirected text learning. *Learning and Individual Differences* 2, 1–17. doi:10.1016/1041-6080(90)90014-8