Learning to Integrate an Artificial Sensory Device: From Decision Fusion to Sensory Integration

Mohammad-Ali Nikouei Mahani[1,2], Karin Maria Bausenhart[2], Rolf Ulrich[2], Majid Nili Ahmadabadi[1]

1 Cognitive Systems Lab, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran

2 Cognition and Perception, Department of Psychology, University of Tübingen, Tübingen, Germany

*Corresponding Author:

Email: nikouei@ut.ac.ir

Abstract

The present study examines how artificial tactile stimulation from a novel non-invasive sensory device is learned and integrated with information from another sensory system. Participants were trained to identify the direction of visual dot motion stimuli with a low, medium, and high signal-to-noise ratio. In bimodal trials, this visual direction information was paired with reliable symbolic tactile information. Over several blocks of training, discrimination performance in unimodal tactile test trials improved, indicating that participants were able to associate the visual and tactile information and thus learned the meaning of the symbolic tactile cues. Formal analysis of the results in bimodal trials showed that the information from both modalities was integrated according to two different integration policies. Initially, participants seemed to rely on a linear decision integration policy based on the metacognitive experience of confidence. In later learning phases, however, our results are consistent with a Bayesian integration policy, that is, optimal integration of sensory information. Thus, the present study demonstrates that humans are capable of learning and integrating an artificial sensory device delivering symbolic tactile information. This finding connects the field of multisensory integration research to the development of sensory substitution systems.

1. Introduction

Humans perceive the environment through multiple sensory inputs. Lacking or unreliable sensory inputs can cause a significant drop in the accuracy of perception. Some artificial sensory devices, such as substitution systems, can partially compensate the loss of a sensory modality [1,2]. Other artificial sensory devices are designed to improve perception by providing complementary or processed information[3,4]. Invasive artificial sensory systems exert a direct effect on the neuronal system[5] that can even lead to optimal integration of multisensory information[6]. However, invasive techniques can still gain from further development. Therefore, non-invasive sensory feedback devices such as wearable systems are potentially the best alternatives for invasive techniques because they are already developed for realistic applications. Although there is a large number of studies that addressed the technical aspects of non-invasive and wearable devices [7-9], the cognitive aspects of them are less well studied. In particular, it is still unclear whether and how the input from a non-invasive artificial sensory device can be integrated into the multisensory perceptual system.

Many studies reported that adult humans integrate multiple sensory modalities in an optimal Bayesian fashion. Bayesian integration, in comparison to unimodal perception, leads to a significant decrease in response time [10-12] and an increase in accuracy and reliability of perception [12-16]. However, most of the previous studies have only examined the integration of well-experienced sensory inputs. There are only a few studies that addressed whether the learning of novel sensory devices leads also to an optimal integration or the selection of a sensory modality. Dadarlat et al. showed that monkeys could optimally integrate unfamiliar multichannel intracortical microstimulation (ICMS) signals and proprioceptive input[6]. Nevertheless, this issue has not been addressed so far for non-invasive artificial devices.

In the present study, we thus investigated the integration of visual motion information with symbolic input from an unfamiliar wearable vibrotactile device. In several consecutive training phases, participants received synchronous static vibrotactile spatial patterns and visual random dot motion stimuli. They were asked to learn these visual-tactile associations. Following each training phase, the

trained stimuli were presented either unimodally or bimodally and participants were asked to report the associated direction of motion. Accuracy and self-reported confidence were assessed to examine whether or not participants can learn the symbolic meaning conveyed by the artificial wearable device and whether the information form the two inputs could be integrated.

To this end, participants performed a multisensory learning task which involves a novel artificial vibro-tactile device (see Figure 1**Error! Reference source not found.**). The experiment consisted of seven blocks, each involving a training phase followed by an evaluation phase. During the training phase, participants simultaneously received a dot motion and a novel vibro-tactile stimulation pattern, which did not involve any directional movement. Throughout the whole experiment, each motion direction was paired with a specific symbolic vibro-tactile pattern. Participants were asked to learn the associations between each specific motion direction and the corresponding vibro-tactile stimulation pattern. In the evaluation phase, participants received a unimodal visual motion dot stimulus, a unimodal vibro-tactile stimulus, or a multimodal stimulus combining both inputs (see materials and methods for more details). Participants were then asked to report motion direction, either directly from the motion of the dots, from the associated vibro-tactile pattern, or from both. They also rated the confidence of their decisions. The visual stimulus in both unimodal conditions and the multimodal visual-tactile condition had three levels of reliability: low, medium and high. The reliability of tactile stimuli was constant during the whole experiment.

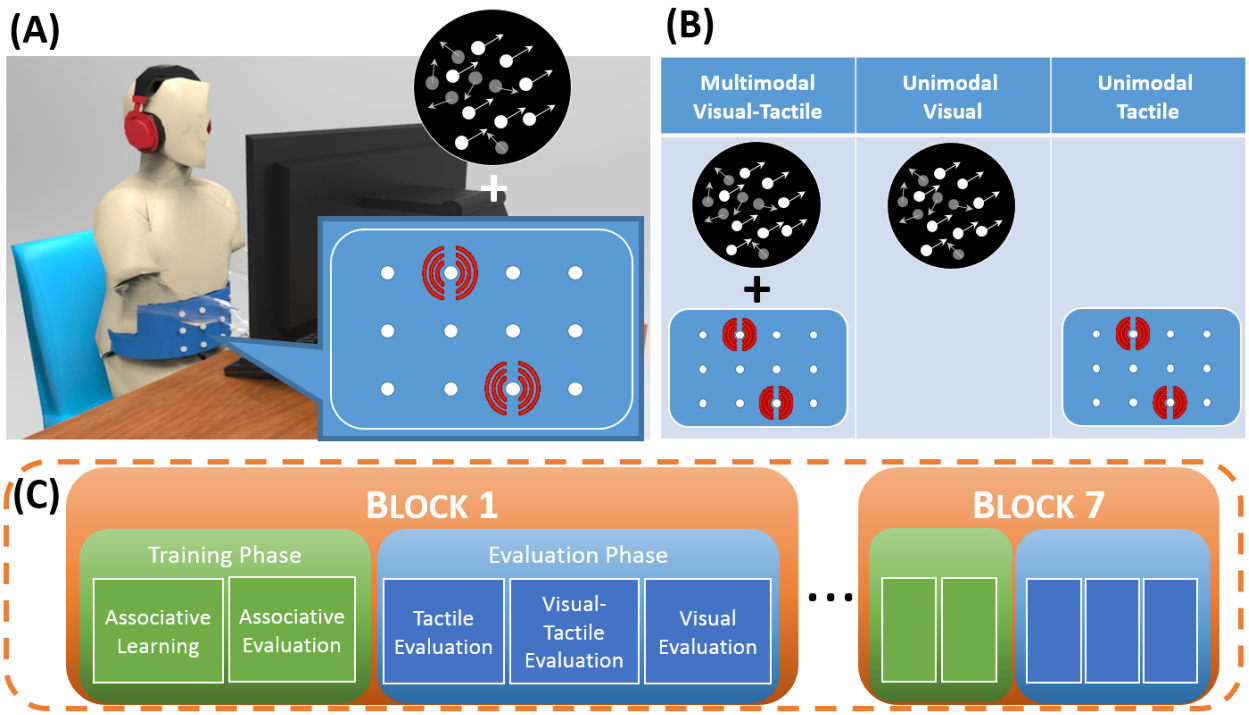

Figure 1. (A) Experimental setup. Participants received synchronous static vibrotactile spatial patterns and visual random dot motion stimuli. Vibroractile patterns were provided by a custom designed belt through a matrix of 3x4 tiny vibration motors. (B) Visual stimulus, tactile stimulus and synchronous visual-tactile stimulus. The reliability of visual dot motion stimulus was manipulated through three coherence levels, while the reliability of the vibrotactile stimulus was fixed. (C) The experiment consisted of seven blocks, each involving a training phase followed by an evaluation phase.

The impact of the artificial sensory stimulation on multisensory perception was assessed over the course of learning. We analyzed the accuracy of perception as well as the confidence in perceptual decisions in low, medium, and high reliability conditions.

2 Results

Figure 2 depicts the accuracy of perception in low, medium, and high reliability conditions. The accuracy of tactile perception increases over the course of the experiment, thus demonstrating that the visuo-tactile patterns provided by our artificial sensory device were efficiently learned and associated with the corresponding dot-motion patterns.

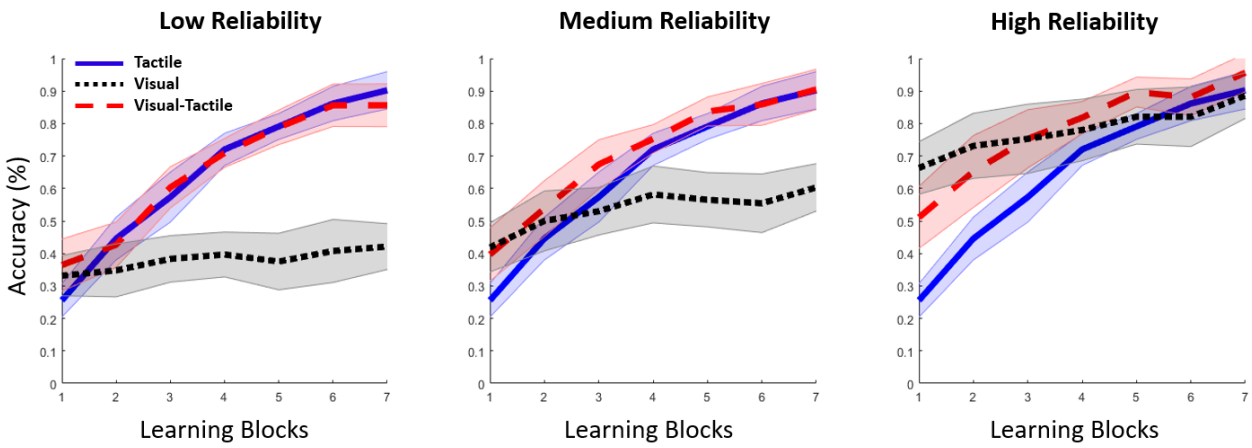

Figure 2. Accuracy of perception in low, medium, and high reliability conditions over the course of the learning procedure. For the sake of clarity, data from the tactile condition are re-plotted in each panel.

Separate two-way within-subject ANOVA with factors reliability condition (low, medium, and high) and block (block 1 to block 7) were performed on accuracy of visual and visual-tactile perceptions.

The accuracy of visual perception was significantly affected by reliability condition, $F(2,44) = 457.22$, $p < .001$, as well as by block, $F(6,132) = 9.48$, $p < .001$. The interaction of reliability condition and block was not significant, $F(12,264) = 1.47$, $p = .13$. Post-hoc Tukey tests on reliability condition showed significant differences between all reliability conditions (*p*s < .001). This result points to a successful reliability control of the visual motion stimuli.

Analysis of the visual-tactile perception also showed a significant effect of reliability condition on accuracy, $F(2,44) = 30.73$, $p < .001$. The effect of block on accuracy was also significant, $F(6,132) = 44.93$, $p < .001$. Post-hoc Tukey tests on reliability condition showed the same result as before, that is, accuracy increased from low to medium, and from medium to high reliability (all *ps* < .001). This effect of reliability condition was especially pronounced in the initial blocks of learning, as indicated by an interaction of both factors, $F(12,264) = 2.81$, $p = .001$. This shows that the information from both modalities is integrated differently depending on the reliability of the visual input.

Additional two-way within-subject ANOVAs with factors modality (visual, tactile, visual-tactile) and block (1 to 7) were conducted for each of the three reliability

conditions to assess learning and multisensory integration of the artificial sensory device.

In the low reliability condition, the effect of modality, $F(2,44) = 83.40$, $p < .001$, the effect of block, $F(6,132) = 52.76$, $p < .001$, and their interaction, $F(12,264) = 23.04$, $p < .001$, were significant. Post-hoc Tukey tests revealed that the visual stimuli were perceived less accurately than tactile ($p < .001$) and visual-tactile stimuli ($p < .001$). However, the difference between the accuracy of visual-tactile and tactile was not significant ($p = .81$). This result indicates a dominance of tactile stimulation on visual-tactile perception, when the reliability of the visual stimuli is low.

The same analysis in the medium reliability condition also showed effects of modality, $F(2,44) = 24.93$, $p < .001$, block, $F(6,132) = 52.42$, $p < .001$, and their interaction, $F(12,264) = 16.08$, $p < .001$. Although post-hoc Tukey tests indicated significant differences among all modalities (all *p*s < .007), the slope of accuracy across blocks was different between modalities. At the beginning of learning, the accuracy for the tactile stimuli was lower than for the visual stimuli, and the accuracy for visual-tactile stimulation was determined by the visual input. However, tactile perception got more accurate over the course of learning and as a result, its influence on the multimodal visual-tactile perception increased. In some learning blocks, the accuracy for visual-tactile stimulation was even higher than for the unimodal visual and tactile stimuli in isolation (see Figure 2). This indicates an effective integration of visual and tactile information (see the modeling section for a more elaborated account of this integration effect).

Finally, in the high reliability condition, accuracy was also affected by modality, $F(2,44) = 13.59$, $p < .001$, block, $F(6,132) = 47.80$, $p < .001$, and their interaction, $F(12,264) = 14.67$, $p < .001$. At the beginning of the experiment, the accuracy for visual-tactile stimulation was codetermined by both visual and tactile information, as indicated by intermediate accuracy, lying between the respective accuracies for visual and tactile unimodal stimulation. This was confirmed by a one-way within-subject ANOVA, which showed a significant difference of accuracy between the different modalities in the first block, $F(2,44) = 55.21$, $p < .001$. Post-hoc tests showed significant differences among all modalities (*p*s < .001). Most interestingly, this suggests that participants integrated the visual and tactile

information from the beginning of learning, even though this impaired performance in comparison to the unimodal visual condition.

3 Modeling

The potential mechanisms underlying the integration of the two modalities were investigated by computational modeling in the subsequent section. In order to shed light on the mechanisms behind the perceptual learning of this artificial sensory device, two computational models are proposed and fitted to the experimental data. The first model focuses on decision integration and the second one on multisensory integration.

3.1 Decision Integration Model

Confidence Confusion Matrix and Meta-Accuracy

The confusion matrix is a useful tool for evaluating the performance of participants in a multiple-choice task or the performance of a classifier algorithm in solving supervised multi-class problems. Various criteria like accuracy, precision, and recall can be determined from the confusion matrix for both 2AFC and multiple choice (multiple class) problems.

For the present purpose, the confusion matrix is extended to second order judgments, that is, self-reported confidence in judgments which is also associated with a given response. When a participant (a classifier) reports a confidence for each of the first-order judgments, this meta information can be used to compute meta-accuracy (type II accuracy). A confidence confusion matrix (CCM) is defined with the same structure as the confusion matrix as follows:

$$CCM = \begin{bmatrix} C_{1,1} & \cdots & C_{1,N} \\ \vdots & \ddots & \vdots \\ C_{N,1} & \cdots & C_{N,N} \end{bmatrix} \quad (1)$$

$C_{i,j}$ indicates the sum of all confidence ratings associated with response $R_{i,j}$, where a ground truth signal i (class i) has been predicted as signal j (class j). Similar to the accuracy, we define meta-accuracy as the sum of the diagonal divided by the sum of all elements:

$$MetaAccuracy = \frac{\sum_{i=j=1}^{N} C_{i,j}}{\sum_{i=1}^{N}\sum_{j=1}^{N} C_{i,j}} \qquad (2)$$

Meta-accuracy represents the performance of the participants not only by considering the decisions but also by weighting the decisions according to the confidence ratings associated with these decisions. High-confidence correct answers thus increase meta-accuracy more than low-confidence correct answers, while the low-confidence wrong answers decrease the meta-accuracy less than the high confidence wrong answers. In fact, CCM is a more general representation of performance than confusion matrix, which is a special case with just one confidence level.

Decision policy

We modeled the multi-modal integration of perceptual decisions considering an ideal observer. This observer has only access to the decisional information, i.e. first- and second-order decisions. In the present case, perceptual decisions are based on two modalities A and B, thus yielding two separate unimodal confidence confusion matrices, CCM^A and CCM^B. CCM can be interpreted as a decision-value table, for which the value of a decision is an internal estimation based on the confidence values. By this interpretation, decision values are considered to be correlated with the confidences, or, in other words, confidences represent the internal estimation of decision values. For example, given an observed signal i, the ith column of the decision value table represents how valuable (and how likely) is each perceptual decision j. One rational way of taking decisions based on two modalities is to decide based on the Max_CCM table, where the highest confidence value is selected for each cell (*i,j*) from the two confidence confusion matrices,

$$Max_CCM_{i,j} = \max\left(CCM_{i,j}^{A}, CCM_{i,j}^{B}\right) . \quad (3)$$

Given the values of all decisions, the probability of selecting a decision when a specific signal is observed can be estimated by various decision making policies. We have investigated two decision making policies in the present study: linear decision making and parametrized softmax decision making.

The linear decision-making policy is defined as follows:

$$p(d|s_i) = \frac{V(s_i,d)}{\sum_{d' \epsilon D} V(s_i,d')} \quad (4)$$

Where $s_i^{'}$ denotes the state in which signal *i* is observed, *D* is the set of all possible decisions, and $V(s_i,d)$ shows the value of decision *d* when the signal *i* is observed. The Max_CCM table is used as the estimation of the decision value table (V). When an agent decides based on the linear policy, the perceptual accuracy of the agent is actually equal to the meta-accuracy of the Max_CCM.

The top row of Figure 3 depicts the result of the ideal linear decision maker. The model successfully predicts the integration behavior for the first blocks. However, it systematically underestimates the accuracy in the last blocks. We have investigated further the accounts for this systematic underestimation by analyzing the calibration of observed confidence ratings. Figure 4 reveals the confidence calibration for all reliability conditions and across all blocks. The orange line indicates the ideal calibration pattern. In an ideal situation, the confidence should represent the actual probability of making correct decisions. As an example, assume that a participant would report the confidence level of 2 (50% probability) for ten trials. If five decisions out of these ten are correct, the confidence would be well calibrated. Lower or higher numbers of correct decisions would show overestimation or underestimation of the confidence.

Figure 4 reveals that the confidence calibration was almost fixed across all blocks in the unimodal visual reliability conditions. This is probably due to none or very little perceptual learning in these conditions. However, confidence calibration varied across blocks in the tactile and all visual-tactile conditions. Specifically, participants tended to report lower confidence for correct decisions in the last three blocks compared to the first three to four blocks. This indicates an underestimation of confidence reports in the last blocks, which also explains the systematic underestimation of the predicted accuracy by the linear decision-making policy. Therefore, we proposed the parametrized softmax model to counterbalance the systematic underestimation error introduced by the linear model. The parametrized softmax model is mathematically equivalent to the temperature scaling [17], which is a well-known approach in deep neural networks to recalibrate the probability estimation.

Moreover, many studies in neuroscience and psychology utilized the softmax policy to model decision making behavior [18-21]. In a softmax approach, usually a decision with a higher value is selected with an exponentially higher probability. The general form of the softmax policy with a linear objective function [18] can be defined as follows:

$$p(d|s_i) = \frac{\exp\left(\theta^T V(s_i,d)\right)}{\sum_{d' \epsilon D} \exp\left(\theta^T V(s_i,d')\right)} \quad (5)$$

In the present study, θ is described by a linear equation with respect to the block number $\theta = wb + \alpha$, where b is the block number and $\{w,\alpha\}$ are free scalar parameters. We considered separate $\{w_r,\alpha_r\}$ for each reliability condition, where r denotes the reliability condition and $r\epsilon\{low, medium, high\}$. The total number of six free parameters has been fitted to the experimental data for all blocks and all three reliability conditions.

Figure 3 depicts results of both linear and softmax ideal observer decision makers. The performance predicted by the linear ideal observer model is close to the observed performance for the visual-tactile stimulation in the first blocks and this is true for all reliability conditions. However, the softmax model depicts a better performance on average in all blocks. It should be noticed that these models do not imply absolute dominance or a weighted combination of the modalities' inputs. Rather the simple decision integration mechanisms outlined above are sufficient to predict the quite complex result pattern. For example, both models are capable to predict the counterintuitive result that multimodal accuracy is partly higher than accuracies for the unimodal conditions (medium reliability condition, blocks 2 and 3). Both models can also predict multimodal accuracy in the high reliability condition, which is between the accuracies of the two unimodal conditions.

Nevertheless, the modeling results also show that the linear ideal observer cannot successfully predict the performance in all cases. Specifically, in the first half of the experiment (blocks 1 to 3/4), the model predicts the performance of the visual-tactile perception very accurately in all reliability conditions. However, the prediction is worse for the second half of the experiment, where this decision integration model systematically underestimates the observed multimodal accuracy. The discrepancy in the linear model can be attributed to a gradual

change in the mechanism that integrates information from the different sensory modalities as the participants become more familiar with the tactile stimulation, from a pure decisional process towards a genuine perceptual integration process. In contrast to the linear model, the softmax model benefits from a linear increase in model parameters across the learning blocks and therefore can capture the gradual change of perceptual mechanism. We will however address the possibility of a genuine perceptual integration process with a Bayesian multisensory integration model in the next section.

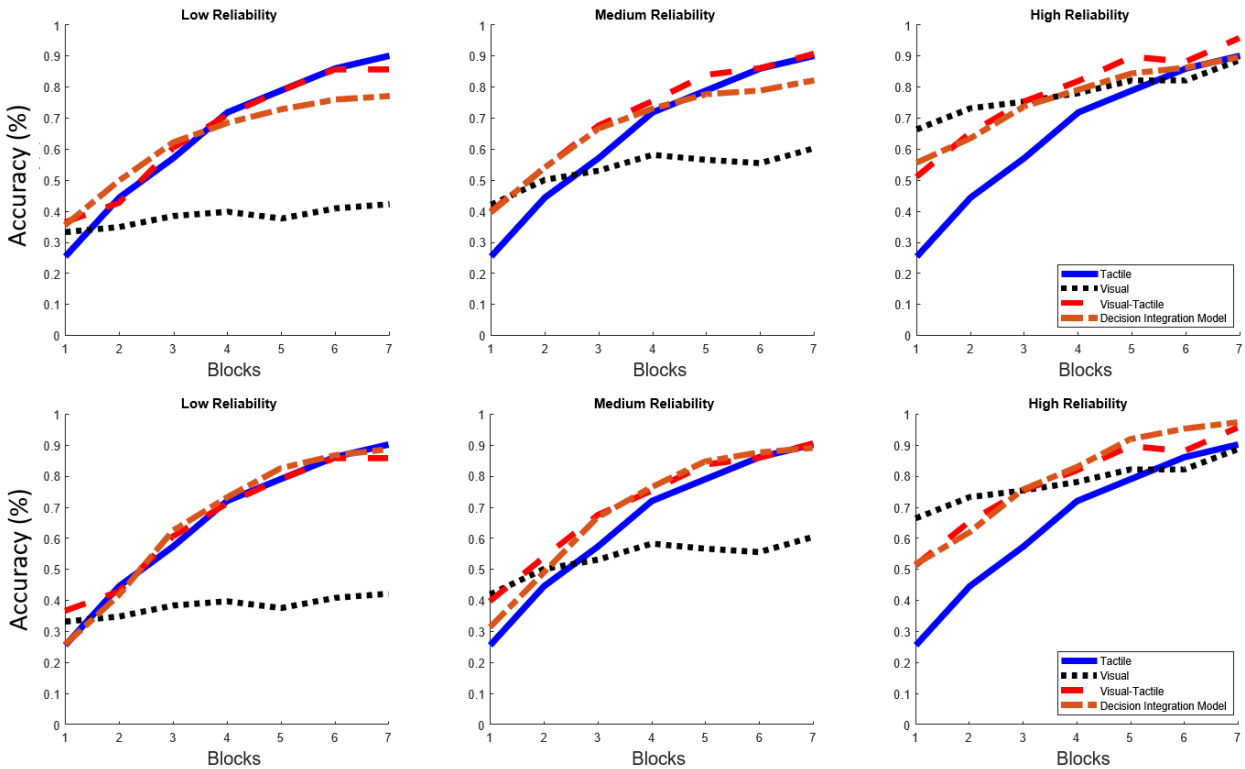

Figure 3. Linear and softmax decision integration models in low, medium and high reliability conditions. The upper row shows the linear model and the lower row depicts the softmax model. Red lines show the experimental visual-tactile integration results while orange lines indicate the modeling result. The model was fitted to the aggregated data from all subjects.

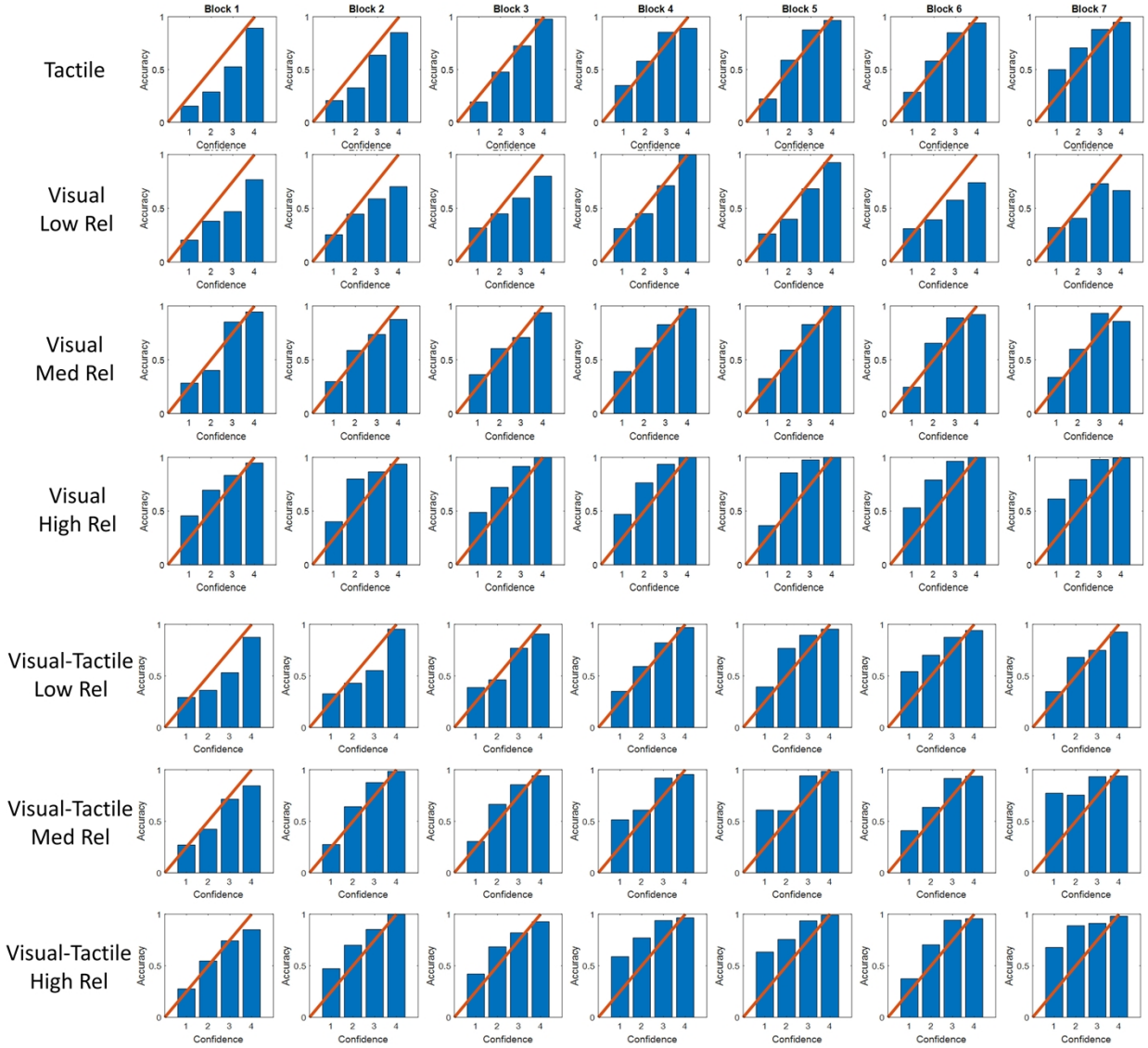

Figure 4. Confidence calibration plots of aggregated data from all participants. The orange line indicates the ideal calibration pattern. The results reveal that unimodal visual confidence calibration changed very little across blocks. However, the change of calibration is obvious in tactile and visual-tactile conditions. Specifically, the participants tended to report lower confidence for correct answers in the last blocks in comparison with the first blocks.

3.2 Multisensory Integration Model

During the previous decade, multisensory integration has become an increasingly prominent research topic and various computational models have been proposed to unravel the mechanism underlying sensory integration across multiple

modalities [10-12,16]. The well-known model of multisensory integration suggests that the sensory inputs are combined according to the principle of maximum-likelihood estimation (MLE)[13]. Even in case of small deviations the core principle of MLE is valid [22] for multisensory integration. Specifically, this model holds that the inputs from different modalities are linearly combined according to the reliability of their respective sensory inputs. This simply means that if there are two modalities, A and B, the weight of each modality in integration, w_A, and $w_B = 1 - w_A$, and the optimal integrated signal S_{AB} are calculated as follows:

$$w_A = \frac{r_A}{r_A + r_B} \qquad (6)$$

$$S_{AB} = w_A S_A + w_B S_B \qquad (7)$$

Where r_A and r_B show the reliability of signal A and reliability of signal B, respectively. We investigated how well the Bayesian model can predict the visual-tactile perception given the unimodal visual and tactile information. The angular distance between the ground truth direction (i.e. dot movement direction) and the chosen (predicted) direction by participants is assumed to represent the reliability of the perception. If in a given trial, the dot display moves in direction α while a participant selects the direction β, then the reliability of perception can be calculated as follows:

$$d_\theta(\alpha,\beta) = cos^{-1}(sin(\alpha) \cdot sin(\beta) + cos\,(\alpha) \cdot cos\,(\beta)) \qquad (8)$$

$$r_{\alpha\beta} = \pi - d_\theta(\alpha,\beta) \qquad (9)$$

$d_\theta(\alpha,\beta)$ is the angular distance between α and β, which has a maximum value of π. Therefore, $r_{\alpha\beta}$ is an estimation of the perceptual reliability that is negatively correlated with the angular distance. A higher angular distance indicates less reliability in the perception.

The confusion matrix of the Bayesian integration model, CM_{ij}^{AB}, is estimated as the weighted sum of unimodal confusion matrices, where the weights are indicated by unimodal perceptual reliabilities:

$$CM_{ij}^{AB} = r_{ij}^{A} CM_{i,j}^{A} + r_{ij}^{B} CM_{i,j}^{B} \qquad (10)$$

r_{ij}^{A} and r_{ij}^{B} are unimodal perceptual reliabilities obtained from Eq (9), CM_{ij}^{A} and CM_{ij}^{B} are unimodal confusion matrices.

Figure 5 depicts the predicted accuracy by the Bayesian integration model. As can be seen in medium and high reliability conditions, the observed accuracy for the visual-tactile condition resembles the predicted accuracy by the Bayesian model especially in the last 3 to 4 blocks. In the low reliability condition, however, the observed accuracy for the visual-tactile condition closely follows the observed unimodal tactile accuracy, suggesting selection behavior rather than integration. Thus, the Bayesian integration model failed to predict the selection behavior in the low reliability condition.

Even though the Bayesian integration model predicted the accuracy of medium and high reliability in the last blocks, it could not fit the experimental data in the first blocks, where the linear decision integration model provided a reasonable data fit. In other words, the linear decision integration model and the multisensory integration model supplement each other.

Table 1 compares BIC statistics of all three models that we mentioned in this paper. These additional results demonstrate that none of three models alone provide a satisfactory description of the experimental data. This additional analysis, however, confirms the conclusion that the linear decision model accounts for the observed accuracy for the first three blocks, whereas the Bayesian integration model accounts for the last three blocks after the initial learning phase.

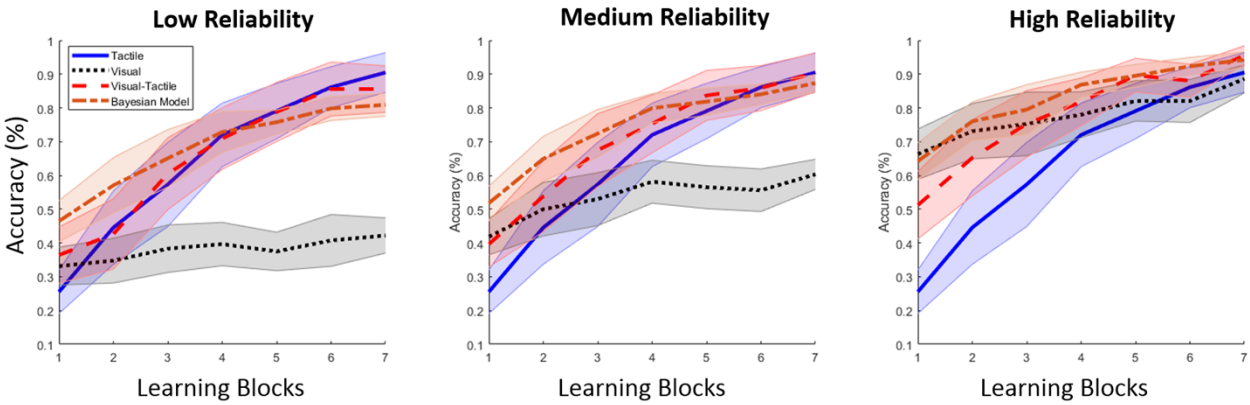

Figure 5. Predicted accuracy by the Bayesian sensory integration model in the low, medium and high reliability conditions. Accuracy is calculated according to Eq (10). Red lines depict the empirical results of the visual-tactile condition, and orange lines show the model fits of the Bayesian integration model. This model was fitted to each individual subject and the confidence intervals were computed by the method explained in [23].

Table 1. Model Comparison

Model	χ^2 Low reliability	χ^2 Medium reliability	χ^2 High Reliability	χ^2 First Three Blocks	χ^2 Last Three Blocks	χ^2 Total	BIC First Three Blocks	BIC Last Three Blocks	BIC Total
Linear Decision Fusion	3.86	1.86	1.29	1.74	5.26	7.00	***1.74***	5.26	7.00
Softmax Decision Fusion	0.86	2.39	3.66	5.82	1.08	6.90	11.54	6.80	14.83
Sensory Integration	8.80	6.86	5.87	19.70	1.83	21.54	19.70	***1.83***	21.54

4. Discussion

With the advance of medical and wearable devices, understanding the mechanisms of perception for novel artificial devices is of great interest to scientists and engineers. This study addresses whether and how symbolic information from a novel non-invasive artificial sensory device is learned and becomes integrated within the human sensory system. We have introduced a custom designed vibrotactile belt that generates novel and unexperienced tactile patterns. Participants were asked to learn the associations between these novel

tactile stimuli and the direction of visual dot motion stimuli across seven consecutive training phases. At the end of each training phase, there was an evaluation phase in which the performance in unimodal visual, unimodal tactile, and visual-tactile conditions was assessed. Thus, perceptual accuracy for the different modality conditions could be tracked over the time-course of the whole experiment.

The results show that accuracy in the tactile condition increased throughout the experiment, and thus, participants could learn the novel tactile patterns. Results also revealed that symbolic information from the novel artificial device can be integrated into the visual-tactile perception from the very beginning of the training. This integration process was evident even though accuracy for the tactile modality was initially lower than for the visual modality. This phenomenon was evident in the medium and high visual reliability conditions. However, in the low visual reliability condition, the tactile information dominated over the visual information, resulting in selection behavior rather than sensory integration.

To better understand the mechanisms underlying the perception and integration of novel symbolic information, two computational models were proposed and evaluated with the data of this study. The decision integration model predicts accuracy in the visual-tactile condition based on first- and second-order perceptual decisions in the two modalities. The multisensory integration model assumes a genuine perceptual information integration, which combines the sensory inputs based on their respective sensory variances on a pre-decisional processing level. This model thus assumes that decision processes operate on the combined sensory multisensory input.

The decision integration model treats confidences as estimates of decision values. Higher confidences are associated with greater values. Assuming that these values are known, the modeling problem is reduced to finding a proper decision-making policy that fits the experimental data. Two decision making policies were evaluated: A linear policy and a parametrized softmax policy. The linear policy explained well the integration behavior in the first half of the experiment. However, it did not fit well the data in the second half of the experiment. Specifically, it underestimated accuracy in the visual-tactile condition. To overcome this problem, we exploited the parametrized softmax policy with a

linear equation parameter (see the modeling section for more details). Since the policy parameters were linearly updated w.r.t. training blocks, the model now fitted reasonably the experimental data in all training blocks. In general, the decision-making model revealed that confidences in decisions properly represent the decision values. This finding together with the notion of confidence confusion matrix provides a new perspective for future research on decision-making models employing self-reported confidence ratings. Moreover, the comparison between the linear policy and the parametrized softmax policy provides evidence for a gradual change in the integration mechanism over the course of training.

We have hypothesized that this gradual change might reflect a switch from a pure decisional integration towards perceptual integration. We addressed this hypothesis by fitting a Bayesian integration model to the experimental data. In contrast to the linear decision-making policy, the Bayesian model fitted well the experimental data only in the second half of the experiment. The complementary fits of the linear policy decision integration model and the Bayesian sensory integration model are consistent with the proposed hypothesis.

In conclusion, the present data show that participants can utilize symbolic tactile information to improve the processing of visual information, thereby improving the accuracy of decision making. We have suggested two models of how this multimodal integration might proceed. In the initial learning phase, participants seem to rely on a linear decision integration process, whereas in a later phase the integration process operates in a Bayesian fashion.

5. Methods

Participants

18 women and 11 men (23.72 ± 3.42 years old) recruited from student population of Tübingen University participated in the experiment. They all reported normal or corrected-to-normal vision and no neurological or psychiatric disorder. The experiment was conducted in accordance with the Helsinki Declaration and the guidelines for scientific work of the University of Tübingen. Written informed consent was obtained from all participants prior to data collection. Participants were compensated with 8 Euros per hour or course credit for their participation.

They were also told that the best performer would get a 50 Euro Amazon gift card. The results of six participants were excluded from all analyses, since their performance in tactile perception was always below 50% in all blocks.

Stimuli and Apparatus

Participants were seated in a sound-attenuated room in front of a 19” screen with 100 Hz refresh rate and 1024x768 pixels resolution, on which the visual stimuli were presented. The distance between participants' eyes and the monitor screen was about 50 cm. The experiment was implemented in C++.

The visual stimuli consisted of a random dot motion display, made up of 100 white dots against a black background. The diameter of each dot was 10 pixels, approximately 0.33° visual angle. The dots were randomly distributed within a circle with a diameter of 400 pixels (approximately 13.3° visual angle) called excircle. The initial x/y position of each dot came from two Gaussian distributions with mean zero (center of the screen) and variances of 0.1 of the width/height of the screen. Each dot started to move with the speed of 11.25°/s in a specific direction. Whenever a dot hit the excircle, it was replaced by a new dot. There were eight possible directions (0, 45, 90, 135, 180, 225, 270, and 315°). The percentage of the dots moving in the same direction, called coherence level, determined the reliability of the visual stimulus. For example, a coherence level of 75% means 75 out of 100 dots moved in the same direction, while 25 dots moved in a random direction. The duration of the visual stimulus was one second.

Tactile stimuli were provided by a custom designed belt fastened around the subjects' waist. The vibrotactile belt was designed in the Cognitive Robotics Lab at University of Tehran. It consists of 12 vibration motors located in a 3x4 matrix formation on a cotton canvas tape. All motors are controlled by a custom embedded device similar to the one employed in a previous study. [16] Eight different tactile stimulation patterns were produced, each consisting of two simultaneous vibrations at pre-defined locations on the belt. These patterns were defined to be maximally distinguishable. The vibration of all tactile stimuli lasted one second, that is, the same duration as the visual stimuli. White-noise auditory stimuli were continuously delivered to the subjects through a pair of Sony MDR-XD200 stereo headphones to override the sound produced by the vibration motors.

Procedure

The experiment consisted of seven consecutive blocks, each of which was divided into a *training phase* and an *evaluation phase (*see Figure 1**Error! Reference source not found.**).

Training phase. Each block started with a training phase in which participants were asked to learn the associations between the tactile patterns and the visual motion directions. A random, subject-specific, and one-by-one mapping associated each of the tactile patterns with one of the visual motion directions. The mappings were randomly assigned prior to the experiment for each participant. The training phase included an associative learning section followed by an associative test section. In the learning section, all visual-tactile association pairs were presented four times in a random order. In the associative test section, participants were asked to decide if the presented pair of visual-tactile stimuli were associated or not, by selecting a green/red circle with a mouse click. Subsequently, they reported their confidence in these decisions on a scale of 1 (lowest confidence level) to 4 (highest confidence level) by selecting one of four adjacent circles by mouse. The cursor position was set to the screen center prior to each judgement to avoid response bias. Participants received feedback about the correctness of their decision at the end of each trial. A splash screen, which showed the overall accuracy achieved during the test section, concluded each training phase. The coherence of the visual stimuli during the whole training phase was 75 %.

Evaluation phase. Following each training phase, performance for visual, tactile, and visual-tactile judgments was assessed. Each evaluation phase included a visual evaluation section, a tactile evaluation section, and a visual-tactile evaluation section, which were presented in random order. In each trial of the visual evaluation section, participants received a unimodal visual stimulus with a coherence level of 10%, 15%, or 25% (low, medium, and high reliability, respectively). Participants were asked to judge the direction of the dot motion by a mouse click on the corresponding sector in a circle. Then, they were asked to report their confidence level on a scale of 1 to 4, similar to the test section of the training blocks. The visual evaluation section consisted of 48 trials (8 directions x 3 coherence levels x 2 repetitions), presented in a random order. The visual-tactile evaluation section was identical to the unimodal visual evaluation section, with

the exception that the visual stimulus in each trial was accompanied by its associated tactile pattern. The tactile evaluation section consisted of 16 trials (8 patterns x 2 repetitions). Participants were asked to report the learned associated motion direction and their confidence in this decision.

References

1 Abboud, S., Hanassy, S., Levy-Tzedek, S., Maidenbaum, S. & Amedi, A. EyeMusic: Introducing a "visual" colorful experience for the blind using auditory sensory substitution. *Restorative neurology and neuroscience* **32**, 247-257 (2014).

2 Maidenbaum, S., Abboud, S. & Amedi, A. Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews* 41, 3-15 (2014).

3 Shull, P. B. & Damian, D. D. Haptic wearables as sensory replacement, sensory augmentation and trainer–a review. *Journal of neuroengineering and rehabilitation* 12, 59 (2015).

4 Spiers, A. J. & Dollar, A. M. in *Haptics Symposium (HAPTICS), 2016 IEEE.* 34-40 (IEEE).

5 Collins, K. L. *et al.* Ownership of an artificial limb induced by electrical brain stimulation. *Proceedings of the National Academy of Sciences* 114, 166-171 (2017).

6 Dadarlat, M. C., O'doherty, J. E. & Sabes, P. N. A learning-based approach to artificial sensory feedback leads to optimal integration. *Nature neuroscience* 18, 138 (2015).

7 Iqbal, M. H., Aydin, A., Brunckhorst, O., Dasgupta, P. & Ahmed, K. A review of wearable technology in medicine. *Journal of the Royal Society of Medicine* 109, 372-380 (2016).

8 Mukhopadhyay, S. C. Wearable sensors for human activity monitoring: A review. *IEEE sensors journal* 15, 1321-1330 (2015).

9 Son, D. *et al.* Multifunctional wearable devices for diagnosis and therapy of movement disorders. *Nature nanotechnology* 9, 397 (2014).

10 Diederich, A. & Colonius, H. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & psychophysics* 66, 1388-1404 (2004).

11 Drugowitsch, J., DeAngelis, G. C., Klier, E. M., Angelaki, D. E. & Pouget, A. Optimal multisensory decision-making in a reaction-time task. *Elife* 3 (2014).

12 Drugowitsch, J., DeAngelis, G. C., Angelaki, D. E. & Pouget, A. Tuning the speed-accuracy trade-off to maximize reward rate in multisensory decision-making. *Elife* 4 (2015).

13 Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429 (2002).

14 Butler, J. S., Smith, S. T., Campos, J. L. & Bülthoff, H. H. Bayesian integration of visual and vestibular signals for heading. *Journal of vision* 10, 23-23 (2010).

15 Pouget, A., Beck, J. M., Ma, W. J. & Latham, P. E. Probabilistic brains: knowns and unknowns. *Nature neuroscience* 16, 1170 (2013).

16 Mahani, M.-A. N., Sheybani, S., Bausenhart, K. M., Ulrich, R. & Ahmadabadi, M. N. Multisensory Perception of Contradictory Information in an Environment of Varying Reliability: Evidence for Conscious Perception and Optimal Causal Inference. *Scientific Reports* 7 (2017).

17 Guo, C., Pleiss, G., Sun, Y. & Weinberger, K. Q. in *Proceedings of the 34th International Conference on Machine Learning-Volume 70.* 1321-1330 (JMLR. org).

18 Reverdy, P. & Leonard, N. E. Parameter estimation in softmax decision-making models with linear objective functions. *IEEE Transactions on automation science and engineering* 13, 54-67 (2016).

19 Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362, 933-942 (2007).

20 Cooper, J. A. *et al.* Training attention improves decision making in individuals with elevated self-reported depressive symptoms. *Cognitive, Affective, & Behavioral Neuroscience* 14, 729-741 (2014).

21 Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* 441, 876 (2006).

22 Meijer, D., Veselič, S., Calafiore, C. & Noppeney, U. J. C. Integration of audiovisual spatial signals is not consistent with maximum likelihood estimation. Cortex 119, 74-88 (2019).

23 Morey, R. D. Confidence intervals from normalized data: A correction to Cousineau (2005). *reason* 4, 61-64 (2008).