2026 11th IEEE RAS/EMBS International Conference for
Biomedical Robotics and Biomechatronics (BioRob)
August 1-4 2026. Edmonton, Canada

Cross-Modal Matching Is Not Bidirectional: Implications for Human Multisensory Feedback

A. Michael West Jr.[1*], Berk Kasimcan[1], Anway S. Pimpalkar[2], Jing Xu[3], and Jeremy D. Brown[1],

2026 11th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob) | 979-8-3315-4813-1/26/$31.00 ©2026 IEEE | DOI: 10.1109/BIOROB66782.2026.11681322

***Abstract*— Cross-modal matching (CMM) is a calibration technique used in multi-modal robotic systems to align the perceptual scales of different sensory feedback channels, such as vision and haptics. In a typical CMM procedure, mappings are learned in both reciprocal directions (e.g., from feedback channel $A \rightarrow B$ and from $B \rightarrow A$). Prior work has reported directional asymmetries in bidirectional matching tasks, but the implications of these asymmetries for calibration remain poorly understood. Here, we examine bidirectionality in two robotics-relevant datasets: a cross-sensory condition involving visual–vibrotactile feedback and a within-sensory, collocated haptic condition involving pressure–vibrotactile feedback. At the individual-subject level, mappings learned in one direction generally failed to describe behavior in the reverse direction, indicating substantial departures from bidirectional consistency. Using parameter-mismatch and cycle-consistency analyses, we demonstrate that these directional asymmetries manifest primarily as systematic perceptual biases rather than intensity-dependent gain distortions or higher-order nonlinearities. As a result, CMM does not simply equate feedback channels but instead introduces direction-dependent perceptual transformations between modalities. Consequently, the choice of mapping direction is not neutral, as it determines how perceptual salience is redistributed across feedback channels. These findings have implications for the design of human-robot interfaces and rehabilitation systems, where calibration choices can influence perception, learning, and sensory reweighting. More broadly, our results reframe CMM not as a neutral calibration procedure, but as a directional design decision that fundamentally shapes multisensory interaction in robotic systems.**

I. INTRODUCTION

Multi-modal feedback (e.g., vision, vibration, force, and/or pressure) is widely used in rehabilitation robotics, assistive devices, human-machine interfaces, and skill training systems to support motor performance and learning [1]–[5]. By distributing task-relevant information across multiple sensory channels, such systems can enhance learning, promote neural plasticity, and compensate for sensory deficits that commonly arise following neurological injury [4], [6], [7]. These approaches implicitly assume that feedback channels can be meaningfully combined; however, sensory modalities differ substantially in scale, salience, and reliability, both across modalities and across individuals [4], [8], [9]. As a result, effective multi-modal feedback requires calibration to ensure that stimuli across modalities are perceptually comparable, referred to here as *perceptual equity* [10], [11]. Without calibration, more reliable modalities tend to dominate perception, biasing multisensory integration toward the most precise sensory channel [9]. Despite its central role in shaping user experience and behavior, calibration of multi-modal feedback is frequently treated as a neutral preprocessing step rather than as an explicit design choice that determines how sensory information is weighted, interpreted, and ultimately acted upon during human-machine interaction.

Cross-modal matching (CMM) is a psychophysical technique for achieving perceptual equity across sensory modalities by estimating individualized mappings between stimulus intensities [12]–[14]. In a typical CMM task, participants are presented with a reference stimulus in one modality (e.g., vision) and adjust the intensity of a second modality (e.g., vibration or pressure) until the two are perceived as equally strong. Repeating this process across stimulus intensities yields a perceptual mapping between modalities that can be used to calibrate multi-modal feedback. CMM has been widely applied in a range of domains including, perceptual assessment [12], sensory substitution and augmentation [15], motor learning [3], and neurorehabilitation [10], [11]. In rehabilitation robotics, it provides a principled means of personalizing feedback to account for individual differences in sensory perception following neurological injury [4], [8].

By construction, CMM yields two perceptual mappings corresponding to the two possible matching directions: from modality A to modality B ($A \rightarrow B$) and from modality B to modality A ($B \rightarrow A$). If these mappings are reciprocal, either can be used interchangeably for calibration. However, if they are not reciprocal, the choice of matching direction directly determines how sensory information is transformed across modalities. Consequently, understanding the extent and consequences of directional asymmetry is critical for the design of multi-modal feedback systems.

Prior work has employed bidirectional CMM procedures and reported directional asymmetries across matching directions [12]–[14]. However, these studies primarily evaluated consistency at the level of individual matching responses, asking whether repeated bidirectional matches produced stable and repeatable judgments. In particular, Pitts et al. [12] demonstrated systematic mismatches that varied across stimulus intensity and highlighted substantial intra-individual variability in repeated matching tasks. Subsequent studies focused on reducing this variability through modifications to the matching procedure and interface design [13], [14]. As a result, existing work has largely interpreted directional

This work was supported by the National Institute of Neurological Disorders and Stroke (NINDS) of the NIH under Award Number R01NS130210 and by the Johns Hopkins University Provost's Postdoctoral Fellowship.

*Corresponding author: A. Michael West Jr. awest36@jhu.edu
[1]Dept. of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA
[2]Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[3]Dept. of Kinesiology, University of Georgia, Athens, GA, USA

asymmetries through the lens of measurement reliability and procedure development. Comparatively little attention has been paid to how these asymmetries influence the perceptual mappings ultimately produced by CMM and, consequently, how they affect the use of CMM as a calibration tool for establishing perceptual equity across sensory modalities.

In robotic and rehabilitative systems, calibration mappings are often embedded directly within closed-loop feedback systems to scale sensory signals, tune assistive devices, and equate training stimuli across feedback channels [3], [4], [11]. Consequently, the direction in which a CMM-based calibration map is applied directly influences how information is delivered and weighted during interaction. If cross-modal mappings are not bidirectional, calibration is no longer a neutral preprocessing step but an explicit design choice that reshapes sensory salience across feedback channels. For human-robot interfaces and rehabilitation systems, incorrect assumptions of bidirectionality may therefore introduce systematic perceptual biases that influence user behavior, learning, and adaptation.

Whether such salience reshaping is desirable depends on the application. In perceptual experiments, reducing the influence of a dominant sensory channel may be beneficial for establishing perceptual equivalence and enabling fair comparisons across modalities [15]. In contrast, rehabilitation and motor learning interventions often depend on preserving rich and informative sensory feedback to promote learning and neural plasticity [3], [4]. Conversely, there may be cases in which intentionally reducing the salience of a dominant modality could be beneficial, for example, to encourage engagement of a weaker or impaired sensory pathway. Taken together, these examples illustrate that understanding how deviations from bidirectionality alter perceptual mappings is critical for determining when cross-modal matching supports, or undermines, the goals of a multi-modal feedback system.

In this work, we investigate the extent and consequences of directional asymmetry in cross-modal matching. We examine two paradigms: vision–vibration and pressure–vibration matching. Unlike much of the prior cross-modal matching literature, which has focused on traditional non-collocated modality pairings, the pressure–vibration paradigm investigates bidirectionality within a collocated tactile–tactile feedback context relevant to haptic interfaces, rehabilitation technologies, and sensory augmentation systems [16]–[18]. Using individual-subject analyses, we evaluate the extent to which mappings learned in one direction generalize to the reverse direction and characterize the resulting perceptual distortions using parameter-mismatch and cycle-consistency analyses.

Here, we move beyond asking whether directional asymmetries exist and instead investigate their perceptual and practical consequences. Our results show that cross-modal mappings fail to exhibit bidirectional consistency at the individual level for both visual-vibrotactile and pressure-vibrotactile modality pairs and that these asymmetries are expressed primarily as systematic perceptual biases. Consequently, mappings obtained in different matching directions produce different calibrations of perceptual intensity, meaning that the choice of matching direction directly influences how sensory information is represented across feedback channels. Together, these findings indicate that cross-modal matching is not a neutral calibration procedure, but one that reshapes how sensory information is distributed across modalities. As a result, the direction of matching becomes an explicit design choice with direct consequences for multi-modal interfaces, haptic feedback, and rehabilitation technologies. This perspective reframes cross-modal matching as a salience-shaping operation that must be considered an integral part of human-robot interaction design, rather than a simple preprocessing step.

II. METHODS

A. Participants

This paper presents two experiments conducted separately (details below). In Experiment 1, N=12 able-bodied adults participated (5 males and 7 females; mean age: 25.2$\pm$6.1 years; 9 right-handed, 1 left-handed, and 2 ambidextrous), while in Experiment 2, N=12 different able-bodied individuals participated (8 males and 4 females; mean age: 23.8$\pm$2.3 years; 12 right-handed). Handedness was assessed using the Laterality Quotient from the Edinburgh Handedness Inventory [19]. Each session lasted $\sim$60 minutes, and participants were compensated $10 per hour. All participants provided informed consent under a protocol approved by the Johns Hopkins School of Medicine Institutional Review Board (#IRB00209185).

B. Cross-Modal Matching Protocol

As in our prior studies [10], [11], we employed a two-phase cross-modal matching procedure consisting of an exploration phase followed by a matching phase (Fig. 1a-b). During exploration, participants sampled the full range of each stimulus with the goal of understanding how changes in one feedback channel corresponded to changes in the other. Stimulus ranges for each modality were defined a priori to span the full safe and usable dynamic range of each device (see details below). Although the experimenter-defined stimulus scaling was linear, participants were not instructed on the specific mapping and were free to form individualized perceptual relationships. Because sensory perception varies between individuals, participants often internalized idiosyncratic mappings between modalities. These individual perceptual mappings were then captured in the subsequent matching phase, in which participants were presented with a reference stimulus in one modality and adjusted the intensity of the other modality until the two were perceived as equally strong. Incorporating these participant-specific mappings enables perceptual equity across modalities in downstream experimental protocols.

Experiment 1 reanalyzed data from [11] and assessed cross-modal matching *across non-collocated* modalities. Specifically, two conditions were tested: (1) *vision-to-vibration*, where participants matched a vibration stimulus to a visual reference stimulus, and (2) *vibration-to-*

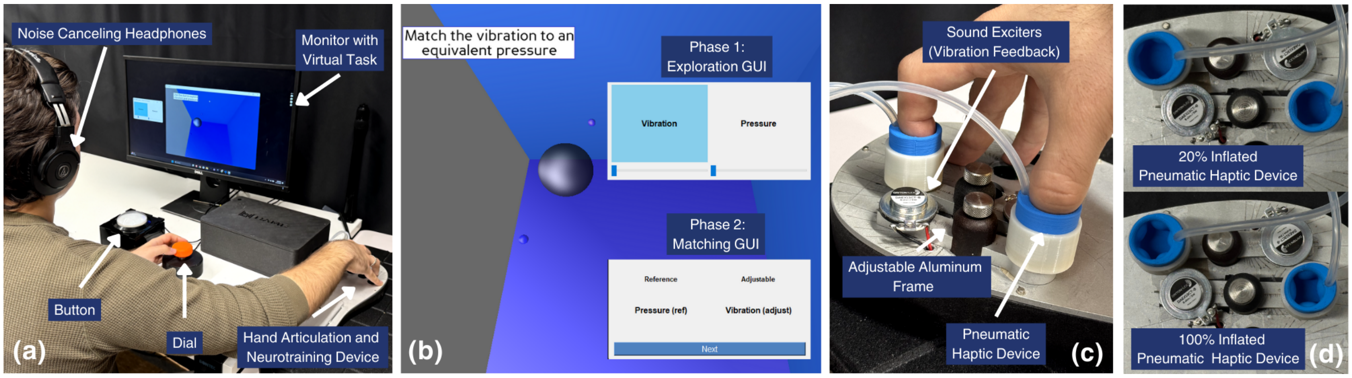

Fig. 1. Experimental setup. (a) Participant workstation. Participants wore noise-canceling headphones and used a continuous rotary dial to adjust stimulus intensity and a button to submit responses. (b) Virtual environment and two-phase graphical user interface. Participants first completed an exploration phase to sample the full stimulus range using a visible slider, followed by a matching phase in which one modality was adjusted to match a reference stimulus in the other modality (example shown: pressure adjusted to match vibration); no slider was shown during matching. The virtual environment was identical across experiments, though in Experiment 2 the visual sphere did not encode stimulus intensity. (c) Pneumatic haptic device used in Experiment 2, providing controllable vibrotactile and pneumatic pressure stimulation at the distal phalanx of the thumb and index fingers. (d) Example pneumatic device states at low and high pressure commands (20% and 100%), illustrating bladder expansion used to modulate pressure intensity.

vision, where a visual stimulus was adjusted to match a vibration reference stimulus. Each condition comprised 32 trials, with 8 reference intensities randomly repeated across 4 blocks. Experiment 2 collected new data and assessed *within-domain, collocated* cross-modal matching. Again, two conditions were tested: (1) *pressure-to-vibration*, where participants matched a vibration stimulus to a pressure (i.e., squeezing) reference stimulus, and (2) *vibration-to-pressure*, where a pressure stimulus was adjusted to match a vibration reference stimulus. Each condition consisted of 24 trials, with 8 reference intensities randomly repeated across 3 blocks. For both experiments, the order of matching conditions was counterbalanced across participants.

C. Hand Articulation and Neurotraining Device (HAND)

As described in [11], Experiment 1 used a modified Hand Articulation and Neurotraining Device (HAND), adapted from prior designs [20], [21] and customized for cross-modal matching (Fig. 1). The device features two compliant silicone contact cups interfacing with the thumb and index finger, each housing a voice-coil actuator (DAEX13CT-8, Dayton Audio, USA) to deliver controlled vibrotactile stimuli to the fingertip pads (see [11] Fig. 1a). Actuators were driven by a custom microcontroller-based amplifier to ensure low-latency and repeatable vibration output across trials. Visual feedback was provided via a virtual display, where stimulus intensity was encoded as the brightness of a gray sphere. Participants adjusted stimulus intensity using a limitless rotary dial.

Participants' forearms were supported by an adjustable armrest, while a mechanically isolated finger rest (not shown) supported non-task digits to prevent transmission of unintended vibratory cues (Fig. 1b). Additionally, participants wore noise-canceling headphones playing broadband white noise to minimize auditory cues from the actuators.

D. Pneumatic Haptic Interface

In Experiment 2, we implemented a collocated pneumatic vibrotactile fingertip interface on the HAND by replacing the silicone contact cups with compliant, 3D-printed thermoplastic polyurethane (TPU) pneumatic bladders (Chinchilla Shore 75A Filament, NinjaTek, USA) mounted at the thumb and index fingertip contact sites (Fig. 1c). Sustained pressure was generated by a piezoelectric disc pump (Lee Ventus, USA; model UBLC5400200A; max 16 kPa) whose flow rate was controlled via pulse-width modulation. Once pressurization stops, the TPU bladders exhibit gradual passive deflation due to leakage/compliance in the pneumatic system, producing a low-pass relaxation behavior consistent with fluidic resistor–capacitor analogs reported in wearable pneumatic haptics [16] (Fig. 1d). Vibrotactile stimulation was delivered independently at the same skin location using the same embedded voice-coil actuators as Experiment 1 with matched command resolution (Fig. 1c).

To accommodate inter-subject variability in finger size, custom-fitted pneumatic finger cups were prepared. A set of cups with inner diameters ranging from 13 to 24 mm in 1 mm increments was available. Participants individually tested cups for the thumb and index finger, beginning with the smallest size and progressing upward. A cup was considered properly fitted if it was the smallest size that comfortably encompassed the distal phalanx and could be lifted without slipping off.

E. Data Analysis

To recap, participants completed a cross-modal matching task in which they matched the perceived intensity of a stimulus presented in one sensory modality to a stimulus presented in another. Two experiments were analyzed: vision–vibration (Experiment 1) and pressure–vibration (Experiment 2). In each experiment, participants performed the task in both directions (e.g., vision reference $\rightarrow$ vibration match and vibration reference $\rightarrow$ vision match). Eight discrete stimulus intensities were presented in a block randomized order, with each intensity repeated four times per direction in Experiment 1 and three times per direction in Experiment 2, yielding 32 and 24 trials per direction (64 and 48 trials total per participant), respectively. All stimulus intensities and matched responses were normalized to the range [0, 1] prior to analysis. Analyses were conducted at the individual-subject level to avoid conflating inter-subject variability with mapping asymmetries.

1) Model Fitting: For each participant and modality pair, cross-modal matching data were modeled using both a linear ($y = mx + b$) model and an exponential ($y = \alpha e^{\beta x}$) model. The linear model captured affine relationships between reference and matched intensities, while the exponential model captured potential nonlinear scaling effects consistent with prior cross-modal matching work [12]–[15]. Models were fit at the individual-subject level using least-squares regression. Model performance was summarized using the ordinary coefficient of determination (R^2). These fits were used both to characterize the shape of the cross-modal mappings and to support subsequent analyses of bidirectionality.

2) Testing Bidirectionality Using F-Tests: To test whether cross-modal matching was bidirectional, we performed model-comparison F-tests at the individual participant level. For each participant, a model was first fit in one mapping direction (e.g., vibration $\rightarrow$ vision). The resulting parameter estimates were then used, without refitting, to predict data collected in the reverse direction (e.g., vision $\rightarrow$ vibration). This cross-direction prediction represents the hypothesis that a single mapping can adequately describe behavior in both directions. The fit of this model was then compared with that of a model fit directly to the reverse-direction data. This procedure was repeated for both linear and exponential models and for both mapping directions, resulting in four F-tests per participant. A significant F-test ($p < 0.05$) indicated that parameters obtained in one direction failed to adequately describe behavior in the reverse direction, demonstrating that the two mappings were not statistically interchangeable and providing evidence against bidirectionality.

3) Parameter Mismatch Analysis: To further investigate the source of non-bidirectionality, we directly compared parameters obtained from forward and reverse cross-modal mappings. For each participant, linear ($y = mx + b$) and exponential ($y = \alpha e^{\beta x}$) models were fit independently in both matching directions. To enable direct comparison of model parameters, mappings obtained in the reverse direction were analytically inverted and rewritten in the same coordinate frame as the forward mapping. Thus, both forward and reverse mappings were expressed as functions relating the same dependent and independent variables. Parameter mismatch was then quantified as the difference between corresponding model coefficients (e.g., Δm, Δb, $\Delta\alpha$, and $\Delta\beta$). Under perfect bidirectionality, corresponding parameters would be identical after this transformation, yielding parameter mismatches of zero. One-sample t-tests were used to determine whether parameter mismatches differed significantly from zero. Significant differences indicate systematic disagreement between forward and reverse mappings and identify the model components contributing to directional asymmetry.

4) Cycle Consistency Analysis: While parameter comparisons identify the coefficients responsible for directional asymmetry, they do not directly describe the perceptual consequences of these differences. We therefore complemented the parameter mismatch analysis with a cycle-consistency framework that expresses directional asymmetry in terms of effective perceptual gain and bias. Specifically, forward and reverse cross-modal mappings were first estimated for each participant. The reverse mapping was then analytically inverted and composed with the forward mapping. For example, the cycle error was defined as

$$\Delta V = f_{X\rightarrow V}^{-1}\left(f_{V\rightarrow X}(V)\right) - V, \tag{1}$$

where $f_{V\rightarrow X}$ denotes the forward mapping from vibration intensity V to the corresponding matched intensity X in the comparison modality (vision in Experiment 1 and pressure in Experiment 2), and $f_{X\rightarrow V}^{-1}$ denotes the inverse of the reverse-direction mapping. Under perfect bidirectionality, composing the two mappings returns the original stimulus intensity and $\Delta V = 0$ for all values of V.

For linear mappings ($y = mV + b$), composition yields

$$\Delta V = \left(\frac{m_1}{m_2} - 1\right)V + \frac{b_1 - b_2}{m_2}, \tag{2}$$

while for exponential mappings ($y = \alpha e^{\beta V}$), composition yields

$$\Delta V = \left(\frac{\beta_1}{\beta_2} - 1\right)V + \frac{1}{\beta_2}\ln\left(\frac{\alpha_1}{\alpha_2}\right) \tag{3}$$

Thus, for both model classes, cycle error reduces to an affine distortion consisting of a gain term (slope) and bias term (offset). Conceptually, the slope captures intensity-dependent distortions in perceptual scaling, whereas the offset captures a constant shift in perceived intensity that is independent of stimulus magnitude. For linear mappings, these distortions are directly apparent from differences in the fitted parameters. For exponential mappings, however, the perceptual consequences of differences in α and β are less intuitive. Cycle consistency therefore provides a common perceptual-space representation for interpreting directional asymmetries across model classes.

Cycle error was evaluated over a dense grid of vibration intensities spanning the overlapping domain of the forward and reverse fits, yielding a continuous cycle-error curve $\Delta V(V)$. Because vibration served as both the input and output of the cycle in both experiments, cycle errors could be compared directly across modality pairs. Cycle error was summarized using the slope and intercept of a linear fit to $\Delta V(V)$, corresponding to effective gain and bias distortions, respectively. One-sample t-tests were used to determine whether cycle-error slopes and offsets differed significantly from zero, while two-sample t-tests were used to compare cycle-error metrics between modality pairs.

All analyses were conducted in MATLAB 2024b. Prior to statistical testing, data were assessed for normality. When normality was violated, nonparametric alternatives were

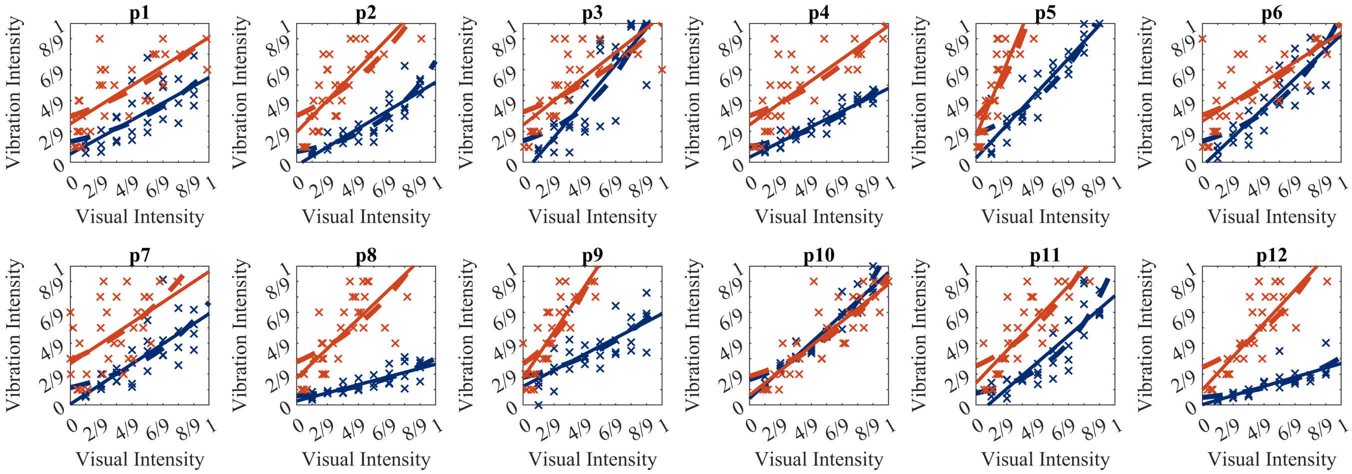

Fig. 2. Individual-subject cross-modal matching data for Experiment 1 (vision–vibration). Each panel shows trial-level responses ('x') and model fits for a single participant. Trials with vibration as the reference (vibration $\rightarrow$ vision) are shown in orange, and trials with vibration as the matched modality (vision $\rightarrow$ vibration) in blue. Linear fits are shown with solid lines, and exponential fits with dashed lines.

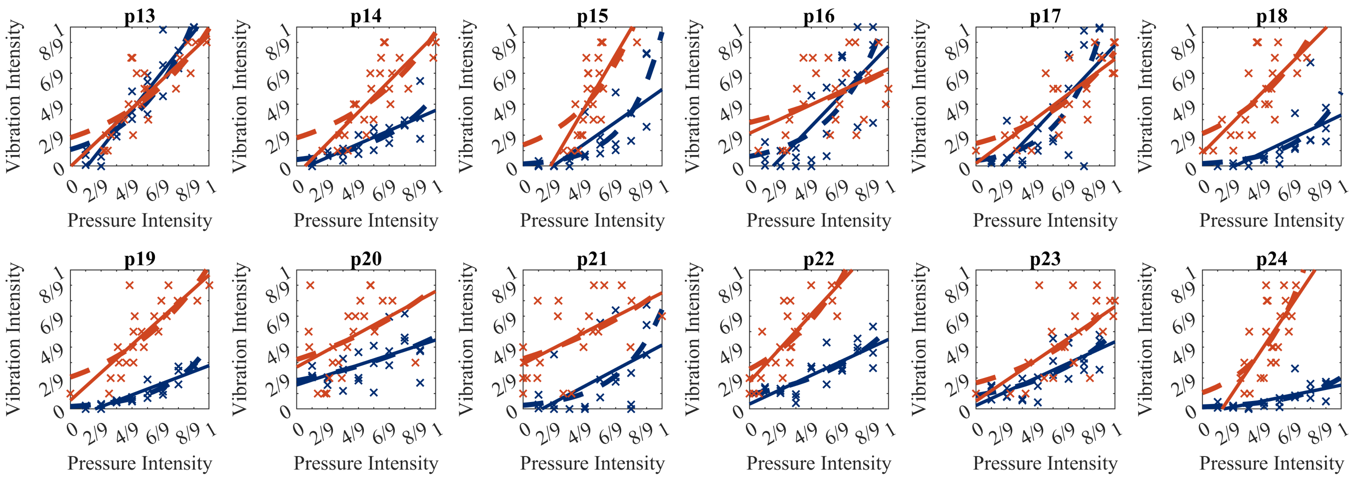

Fig. 3. Individual-subject cross-modal matching data for Experiment 2 (pressure–vibration). Each panel shows trial-level responses ('x') and model fits for a single participant. Trials with vibration as the reference (vibration $\rightarrow$ pressure) are shown in orange, and trials with vibration as the matched modality (pressure $\rightarrow$ vibration) in blue. Linear fits are shown with solid lines, and exponential fits with dashed lines.

used: the Wilcoxon signed-rank test for one-sample comparisons, and the Wilcoxon rank-sum test for independent two-sample comparisons. All statistical tests were conducted using a significance level of $\alpha = 0.05$.

III. RESULTS

A. Model Fits Provide Adequate Descriptions of Cross-Modal Matching Behavior

Figs. 2 and 3 show individual-subject cross-modal matching data for Experiment 1 (vision–vibration) and Experiment 2 (pressure–vibration), respectively. Each panel displays an individual's trial-level responses ('x') along with model fits. Trials in which vibration served as the reference modality (vibration $\rightarrow$ vision or vibration $\rightarrow$ pressure) are in orange, whereas trials in which vibration was the matched modality (vision $\rightarrow$ vibration or pressure $\rightarrow$ vibration) are in blue.

Both linear and exponential models were fit to each participant's data in each direction (solid and dashed lines, respectively). Model fit quality was summarized using ordinary R^2. For Experiment 1, mean R^2 values were 0.65 ± 0.15 (linear) and 0.60 ± 0.16 (exponential); for Experiment 2, they were 0.54 ± 0.18 and 0.52 ± 0.18, respectively. All fits were significant ($p < 10^{-3}$). Together, these results indicate that the two model classes provide comparable descriptive performance, justifying subsequent analyses that directly assess bidirectionality rather than relying on goodness-of-fit alone.

B. F-Tests Demonstrate Directional Asymmetry

To directly test whether mappings were bidirectionally consistent, we conducted F-tests comparing whether a model fit in one direction could adequately explain data collected in the opposite direction. For each participant, this analysis was performed for both linear and exponential models in both mapping directions, resulting in four tests per participant. As summarized in Tables I and II, most tests yielded significant results ($p < 0.05$), indicating that models derived in one

TABLE I

EXPERIMENT 1 F-TESTS. COLUMNS SHOW F-STATISTICS ($df_1 = 2$, $df_2 = 30$) WITH ASSOCIATED P-VALUES (IN PARENTHESES) FOR VIBRATION $\rightarrow$ VISION AND VISION $\rightarrow$ VIBRATION CONDITIONS UNDER LINEAR (LIN) AND EXPONENTIAL (EXP) MODELS. P-VALUES BELOW 10^{-5} ARE OMITTED. ENTRIES IN BOLD INDICATE $p > 0.05$.

	Vib. $\rightarrow$ Vis.		Vis. $\rightarrow$ Vib.	
Participant	**Lin**	**Exp**	**Lin**	**Exp**
p1	51.8	40.8	31.4	23.1
p2	499.5	492.2	72.4	53.6
p3	16.3	11.2 (2.3×10^{-4})	31.0	17.0
p4	925.2	685.2	46.8	37.8
p5	1161.5	4549.9	89.1	36.6
p6	40.1	24.5	24.2	13.1 (8.0×10^{-5})
p7	80.2	85.2	29.0	25.2
p8	1789.5	1728.0	86.8	67.7
p9	322.2	675.3	68.5	48.3
p10	5.8 (7.3×10^{-3})	10.9 (2.8×10^{-4})	**1.6 (0.217)**	**3.3 (0.051)**
p11	106.8	130.9	56.4	45.8
p12	1291.1	1160.2	136.7	93.1

TABLE II

EXPERIMENT 2 F-TESTS. COLUMNS SHOW F-STATISTICS ($df_1 = 2$, $df_2 = 22$) WITH ASSOCIATED P-VALUES (IN PARENTHESES) FOR VIBRATION $\rightarrow$ PRESSURE AND PRESSURE $\rightarrow$ VIBRATION CONDITIONS UNDER LINEAR (LIN) AND EXPONENTIAL (EXP) MODELS. P-VALUES BELOW 10^{-5} ARE OMITTED. ENTRIES IN BOLD INDICATE $p > 0.05$.

	Vib. $\rightarrow$ Pres.		Pres. $\rightarrow$ Vib.	
Participant	**Lin**	**Exp**	**Lin**	**Exp**
p13	5.7 (9.9×10^{-3})	7.2 (3.9×10^{-3})	**2.2 (0.140)**	5.8 (9.3×10^{-3})
p14	173.5	182.3	45.6	36.1
p15	79.3	109.2	39.4	44.9
p16	8.4 (2.0×10^{-3})	9.3 (1.2×10^{-3})	6.0 (8.4×10^{-3})	10.9 (5.2×10^{-4})
p17	**2.0 (0.162)**	5.5 (0.0115)	3.5 (0.0480)	32.7
p18	172.3	196.4	66.9	61.4
p19	1024.4	1413.4	89.9	73.1
p20	41.8	40.3	9.4 (1.1×10^{-3})	9.1 (1.3×10^{-3})
p21	39.4	40.9	38.6	35.3
p22	372.3	469.9	64.4	52.0
p23	24.2	24.3	10.0 (8.1×10^{-4})	9.0 (1.4×10^{-3})
p24	1270.1	2038.5	67.2	67.6

direction failed to explain data collected in the reverse direction. Only one participant in Experiment 1 (p10) and two participants in Experiment 2 (p13, p17) satisfied bidirectionality under the linear model, and only one participant (p10, Experiment 1) did so under the exponential model.

Together, these results demonstrate that cross-modal mappings are generally not bidirectional at the individual-subject level, regardless of whether linear or exponential models are used. Importantly, this asymmetry persists across both visual-vibrotactile and vibrotactile-vibrotactile domains, motivating further analysis of its underlying source.

C. *Parameter Mismatch Reveals the Source of Asymmetry*

To identify the source of the bidirectionality violations observed in the F-tests, we compared the parameters obtained from forward and reverse cross-modal mappings (Fig. 4).

For the linear model, differences in slope (Δm; Fig. 4A) were not significantly different from zero in either the vision-vibration experiment (-0.318 ± 0.518, $p = 0.0569$) or the pressure-vibration experiment (-0.315 ± 0.555, $p = 0.0753$), and did not differ between experiments ($p = 0.9881$). In contrast, differences in intercept (Δb; Fig. 4B) were significantly different from zero in the vision-vibration experiment (-0.201 ± 0.101, $p = 2.66 \times 10^{-5}$), but not in the pressure-vibration experiment (-0.119 ± 0.189, $p = 0.0516$). Intercept

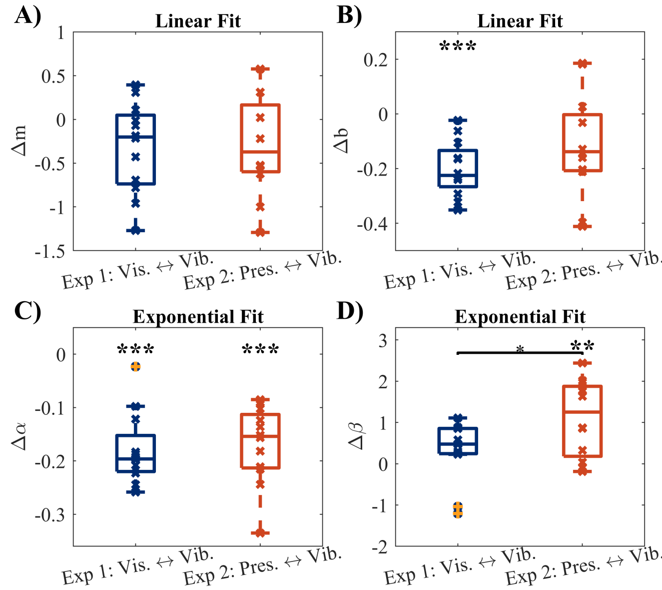

Fig. 4. Differences between model coefficients obtained from forward and reverse matching directions are shown for linear slope mismatch (Δm; A), linear intercept mismatch (Δb; B), exponential scale parameter mismatch ($\Delta \alpha$; C), and exponential exponent mismatch ($\Delta \beta$; D). A value of zero indicates perfect agreement between forward and reverse mappings. Boxplots show group distributions with individual participants overlaid; outliers are highlighted with a yellow '+'. Asterisks indicate one-sample t-test significance against zero (* $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$), while horizontal bars indicate significant between-experiment differences.

differences did not differ between experiments ($p = 0.2041$).

For the exponential model, differences in the scale parameter ($\Delta \alpha$; Fig. 4C) were significantly different from zero in both the vision-vibration experiment (-0.180 ± 0.067, $p = 1.58 \times 10^{-6}$) and the pressure-vibration experiment (-0.169 ± 0.073, $p = 6.19 \times 10^{-6}$), with no significant difference between experiments ($p = 0.7238$). Differences in the exponent parameter ($\Delta \beta$; Fig. 4D) were not significantly different from zero in the vision-vibration experiment (0.314 ± 0.726, $p = 0.1621$), but were significantly different from zero in the pressure-vibration experiment (1.115 ± 0.929, $p = 0.0016$). Moreover, exponent differences were significantly larger in the pressure-vibration experiment than in the vision-vibration experiment ($p = 0.0284$). Together, these results indicate that non-bidirectionality arises from systematic parameter mismatches, particularly the intercept in linear models and the scale and exponent parameters in exponential models.

D. *Cycle Consistency Analysis Reveals Perceptual Biases*

We next expressed the exponential parameter mismatches identified above within a common perceptual coordinate frame using cycle consistency analysis. Whereas differences in the slope and intercept of linear mappings can be directly interpreted as perceptual gain and bias distortions, the perceptual consequences of differences in the exponential parameters α and β are less intuitive. Cycle consistency analysis was therefore applied to the exponential fits to translate these differences into equivalent gain and bias distortions. As

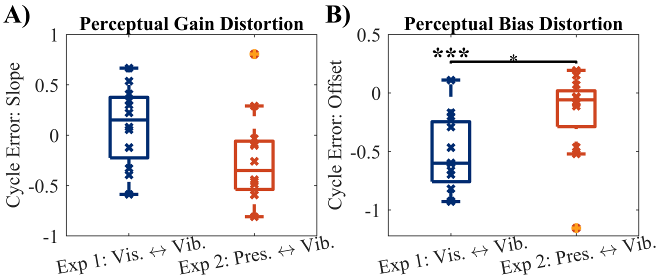

Fig. 5. Cycle-error slope (A) and offset (B) derived from exponential cross-modal matching models in Experiments 1 (vision–vibration) and 2 (pressure–vibration). These quantities were obtained from linear fits to the cycle error as a function of stimulus intensity and provide a common perceptual-space representation of directional asymmetry, capturing effective perceptual gain distortions and constant perceptual biases, respectively. Boxplots show group distributions with individual participants overlaid; outliers are highlighted with a yellow '+'. Asterisks indicate one-sample t-test significance against zero (* $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$), while horizontal bars indicate significant between-experiment differences.

shown in Eq. 3, the cycle error associated with exponential mappings reduces to an affine function of stimulus intensity, where the coefficient of V corresponds to an effective gain distortion and the constant term corresponds to an effective bias distortion. Thus, cycle-error slope and offset were used to quantify perceptual gain and bias, respectively (Fig. 5).

Contrary to what might be expected from the observed parameter mismatches, cycle-error slopes (Fig. 5A) were not significantly different from zero in either the vision-vibration experiment (0.098 ± 0.390, $p = 0.4004$) or the pressure-vibration experiment (-0.230 ± 0.445, $p = 0.1008$). Moreover, cycle-error slopes did not differ significantly between experiments ($p = 0.0677$). Together, these results suggest that directional asymmetries did not manifest as reliable intensity-dependent perceptual gain distortions.

In contrast, cycle-error offsets (Fig. 5B) were significantly different from zero in the vision-vibration experiment (-0.521 ± 0.325, $p = 1.74 \times 10^{-4}$), indicating a systematic perceptual bias. Although offsets were not significantly different from zero in the pressure-vibration experiment (-0.176 ± 0.375, $p = 0.1314$), the offset observed in the vision-vibration experiment was significantly larger than that in the pressure-vibration experiment ($p = 0.0252$). Thus, directional asymmetries manifested primarily as offset-like perceptual biases rather than gain distortions.

IV. DISCUSSION

This study examined the extent and consequences of directional asymmetry in cross-modal matching (CMM). Across two experiments spanning both cross-sensory (vision–vibration) and within-sensory (pressure–vibration) domains, we found that perceptual mappings obtained in one matching direction did not generalize to the reverse direction (Tables I and II). While prior work has reported directional inconsistencies and intra-individual variability in bidirectional matching procedures [12]–[14], our results demonstrate that these asymmetries produce systematic distortions in the perceptual mappings ultimately used for calibration. Parameter mismatch analyses revealed significant differences in intercept (Δb) and exponential scale ($\Delta\alpha$) parameters (Fig. 4), while cycle-consistency analyses showed that these asymmetries manifested primarily as systematic perceptual biases rather than intensity-dependent gain distortions (Fig. 5). Together, these findings indicate that CMM does not yield a single neutral calibration between modalities, but instead produces direction-dependent transformations whose consequences must be explicitly considered when designing multi-modal robotic systems and human-machine interfaces.

Beyond demonstrating directional asymmetry in a visuo–tactile paradigm, this work extends cross-modal matching analysis to a pressure–vibration pairing that is directly relevant to wearable robotics, prosthetics, and soft haptic interfaces. Unlike vision and vibration, pressure and vibration are collocated tactile cues that act on the same effector and are frequently combined in sensory substitution and augmentation systems [15]–[18]. Despite these substantial differences in sensory modality and embodiment, both experiments exhibited remarkably similar forms of directional asymmetry (Fig. 4). In particular, parameter mismatch and cycle-consistency analyses revealed that directional asymmetries were expressed primarily as systematic perceptual biases in both experiments. This convergence suggests that directional asymmetry is not unique to a particular sensory pairing, but instead reflects a more general property of how humans relate and scale sensory feedback across modalities. For designers of haptic and wearable robotic systems, these findings suggest that direction-dependent calibration effects should be expected even when combining closely related, collocated feedback cues.

While cross-modal matching has historically been treated as a neutral preprocessing step for equating sensory modalities [10], [11], our results demonstrate that it functions as an active design operation that reshapes perceptual mappings in a direction-dependent manner. Because CMM yields two distinct mappings (A→B and B→A), the presence of directional asymmetry implies that these mappings cannot be reduced to a single, modality-independent perceptual scale. Applying a CMM-derived mapping is therefore not a passive calibration, but an explicit transformation that redistributes perceptual weighting across feedback channels. In practical terms, selecting a calibration direction applies a gain and offset to one modality relative to the other, compressing the effective perceptual range of one feedback channel and altering the relative influence of the other. In principle, CMM-based calibration could be implemented either by compressing one modality or expanding the other. In practice, however, stimulus ranges are typically fixed prior to calibration by device limits and safety constraints, leaving compression of one feedback channel as the only feasible implementation. As a result, the chosen calibration direction directly determines which modality exerts greater perceptual, and consequently control, authority during interaction. For multi-modal robotic systems, this means that CMM is not merely a technical detail, but a consequential design decision that shapes how users interpret, trust, and act upon multisensory feedback.

A. Impact

In prior studies [10], [11], [15], CMM-based calibration quality has been evaluated primarily using goodness-of-fit metrics (e.g., R^2), implicitly assuming that perceptual mappings are sufficiently consistent across matching directions to support a single calibration. Here, we show that high goodness-of-fit does not guarantee bidirectional consistency. To address this limitation, we introduce a framework that combines parameter mismatch analyses and cycle-consistency analyses to evaluate bidirectionality and characterize its perceptual consequences. Importantly, while parameter mismatches directly identify the source of asymmetry, cycle-consistency analysis expresses these asymmetries in terms of perceptually meaningful distortions, such as bias and gain. Although we apply this framework to linear and exponential mappings, the cycle-consistency approach is model agnostic and can be extended to more complex perceptual mappings where the perceptual implications of parameter differences may be difficult to interpret directly. Our results demonstrate that when mappings are not bidirectional, calibration necessarily privileges one feedback channel over another, making the choice of matching direction consequential. In closed-loop robotic systems, these asymmetries may propagate through the control loop, influencing error correction, exploration, learning, and ultimately user behavior, trust, and control performance.

Critically, the implications of non-bidirectional calibration differ fundamentally between perceptual experiments and rehabilitation systems. In perceptual and psychophysical studies, equalizing saliency across modalities, even at the cost of compressing one channel, is often desirable to enable fair comparisons and isolate sensory contributions [15]. In contrast, rehabilitation robotics and motor learning systems typically aim to maximize informative feedback to support exploration, adaptation, and skill acquisition [3], [5]. Because CMM-based calibration necessarily redistributes perceptual weighting, applying it uncritically may inadvertently suppress task-relevant feedback and hinder learning or sensory integration in clinical populations. At the same time, this directional asymmetry may be advantageous if used intentionally: selectively down-weighting a dominant modality (e.g., vision) could encourage engagement of weaker sensory channels such as tactile or proprioceptive feedback, a strategy that may benefit stroke rehabilitation and sensory reweighting interventions [4], [6], [8], [9]. In this light, CMM is not merely a calibration tool used for equating sensory modalities, but a powerful design lever: a mechanism by which sensory saliency can be intentionally redistributed, and potentially tuned as a control parameter, to shape perception, control, and learning in human-robot interaction and rehabilitation systems.

REFERENCES

[1] B. Lee, I. D. Saragih, and S. O. Batubara, "Robotic arm use for upper limb rehabilitation after stroke: A systematic review and meta-analysis," *The Kaohsiung Journal of Medical Sciences*, vol. 39, no. 5, p. 435–445, Mar. 2023.

[2] J. M. Veerbeek, A. C. Langbroek-Amersfoort, E. E. H. van Wegen, C. G. M. Meskers, and G. Kwakkel, "Effects of robot-assisted therapy for the upper limb after stroke: A systematic review and meta-analysis," *Neurorehabilitation and Neural Repair*, vol. 31, no. 2, p. 107–121, Sep. 2016.

[3] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review," *Psychonomic Bulletin & Review*, vol. 20, no. 1, p. 21–53, Nov. 2012.

[4] J. A. Semrau, T. M. Herter, S. H. Scott, and S. P. Dukelow, "Vision of the upper limb fails to compensate for kinesthetic impairments in subacute stroke," *Cortex*, vol. 109, p. 245–259, Dec. 2018.

[5] A. Ahmed, B. Hugo, S. Lucas, R. Diana, O. Etienne, and G. Pascal, "Distinct and additive effects of visual and vibratory feedback for motor rehabilitation: an eeg study in healthy subjects," *Journal of NeuroEngineering and Rehabilitation*, vol. 21, no. 1, Sep. 2024.

[6] S. M. Schabrun and S. Hillier, "Evidence for the retraining of sensation after stroke: a systematic review," *Clinical rehabilitation*, vol. 23, no. 1, pp. 27–39, 2009.

[7] V. L. Feigin, M. Brainin, B. Norrving, S. Martins, R. L. Sacco, W. Hacke, M. Fisher, J. Pandian, and P. Lindsay, "World stroke organization (wso): Global stroke fact sheet 2022," *International Journal of Stroke*, vol. 17, no. 1, p. 18–29, Jan. 2022.

[8] J. A. Semrau, T. M. Herter, S. H. Scott, and S. P. Dukelow, "Examining differences in patterns of sensory and motor recovery after stroke with robotics," *Stroke*, vol. 46, no. 12, pp. 3459–3469, 2015.

[9] M. O. Ernst and M. S. Banks, "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature*, vol. 415, no. 6870, pp. 429–433, 2002.

[10] A. M. West Jr., A. S. Pimpalkar, J. Xu, and J. D. Brown, "Sensematch: Smartphone-based cross-modal matching for accessible perceptual assessments," in *Proceedings of the International Conference on Neural Engineering (NER)*, 2025.

[11] A. S. Pimpalkar, A. M. West, J. Xu, and J. D. Brown, "Optimizing Cross-Modal Matching for Multimodal Motor Rehabilitation," *IEEE International Conference on Rehabilitation Robotics*, pp. 559–566, 2025.

[12] B. J. Pitts, S. A. Lu, and N. B. Sarter, "Cross-modal matching: The development and evaluation of a new technique," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, p. 1760–1764, Sep. 2013.

[13] B. J. Pitts and N. B. Sarter, "Crossmodal matching: Validation of a more reliable technique," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, p. 1751–1755, Sep. 2014.

[14] B. Pitts, S. L. Riggs, and N. Sarter, "Crossmodal matching: A critical but neglected step in multimodal research," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, p. 445–450, Jun. 2016.

[15] N. Thomas, G. Ung, C. McGarvey, and J. D. Brown, "Comparison of vibrotactile and joint-torque feedback in a myoelectric upper-limb prosthesis," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, p. 70, jun 2019.

[16] B. Jumet, Z. A. Zook, A. Yousaf, A. Rajappan, D. Xu, T. F. Yap, N. Fino, Z. Liu, M. K. O'Malley, and D. J. Preston, "Fluidically programmed wearable haptic textiles," *Device*, vol. 1, no. 3, p. 100059, Sep. 2023.

[17] J. Qi, F. Gao, G. Sun, J. C. Yeo, and C. T. Lim, "Haptglove—untethered pneumatic glove for multimode haptic feedback in reality–virtuality continuum," *Advanced Science*, vol. 10, no. 25, Jun. 2023.

[18] C. du Pasquier, L. Tessmer, I. Scholl, L. Tilton, T. Chen, S. Tibbits, and A. Okamura, "Haptiknit: Distributed stiffness knitting for wearable haptics," *Science Robotics*, vol. 9, no. 97, Dec. 2024.

[19] R. Oldfield, "The assessment and analysis of handedness: The edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, p. 97–113, Mar. 1971.

[20] J. Xu, T. Ma, S. Kumar, K. Olds, J. Brown, J. Carducci, A. Forrence, and J. Krakauer, "Loss of finger control complexity and intrusion of flexor biases are dissociable in finger individuation impairment after stroke," *eLife*, Nov. 2023.

[21] J. Carducci, K. Olds, J. W. Krakauer, J. Xu, and J. D. Brown, "Novel planar strain sensor design for capturing 3-dimensional fingertip forces from patients affected by hand paralysis," *Sensors*, vol. 22, no. 19, p. 7441, Sep. 2022.