

Vibrotactile Preference Learning: Uncertainty-Aware Preference Learning for Personalized Vibration Feedback

Rongtao Zhang*
University of Southern California
Los Angeles, California, USA
rongtaoz@usc.edu

Xin Zhu*
University of Southern California
Los Angeles, California, USA
xinzhu@usc.edu

Masoume Pourebadi
Khotbehsara
University of Southern California
Los Angeles, California, USA
pourebadi@usc.edu

Warren Dao
University of Southern California
Los Angeles, California, USA
wdao@usc.edu

Erdem Biyik
University of Southern California
Los Angeles, California, USA
biyik@usc.edu

Heather Culbertson†
University of Southern California
Los Angeles, California, USA
hculbert@usc.edu

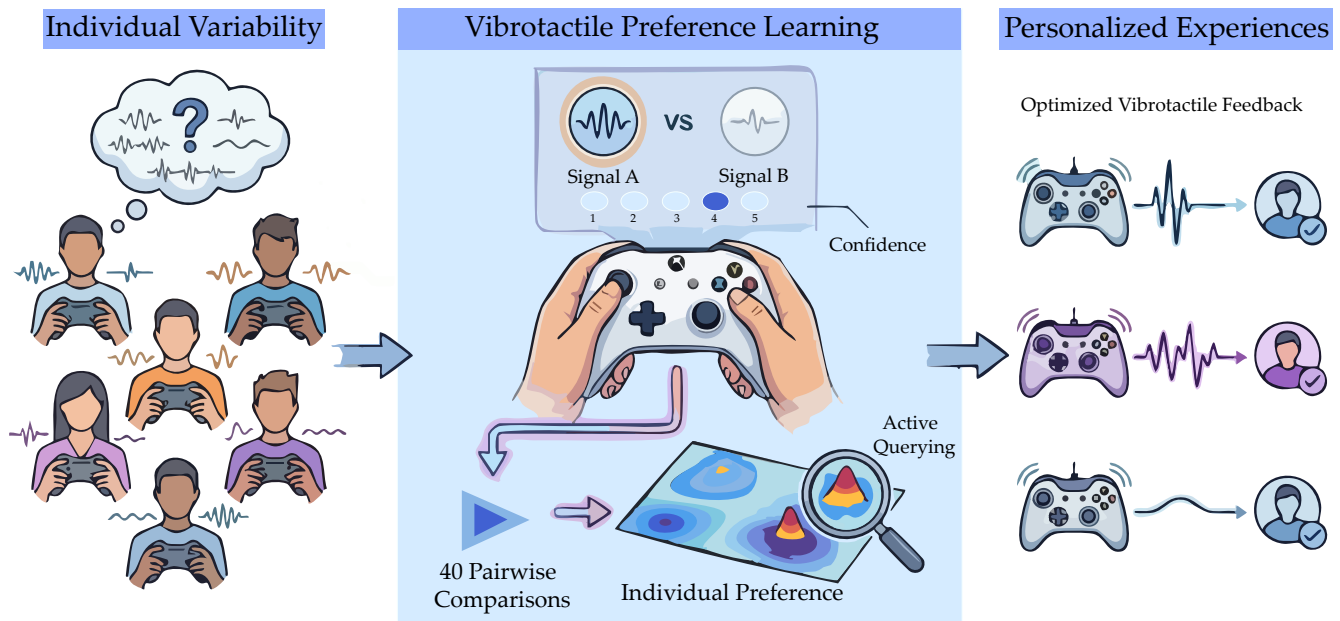


Figure 1: Overview of the Vibrotactile Preference Learning System

Abstract

Individual differences in vibrotactile perception underscore the growing importance of personalization as haptic feedback becomes more prevalent in interactive systems. We propose Vibrotactile Preference Learning (VPL), a system that captures user-specific preference spaces over vibrotactile parameters via Gaussian-process-based uncertainty-aware preference learning. VPL uses an expected information gain-based acquisition strategy to guide query selection over 40 rounds of pairwise comparisons of overall user preference, augmented with user-reported uncertainty, enabling efficient exploration of the parameter space. We evaluate VPL in a user study ($N = 13$) using the vibrotactile feedback from a Microsoft Xbox controller, showing that it efficiently learns individualized

preferences while maintaining comfortable, low-workload user interactions. These results highlight the potential of VPL for scalable personalization of vibrotactile experiences.

CCS Concepts

• Information systems → Personalization; • Human-centered computing → User models; User studies; Haptic devices; • Computing methodologies → Active learning settings.

Keywords

vibrotactile feedback, preference learning, personalization, active learning, gaussian processes, user modeling

1 Introduction

Haptic feedback is a key channel for conveying information and enhancing immersion in virtual and remote environments [3, 24].

*Both authors contributed equally to this research.

† Corresponding author.

Advances in haptic technology have enabled precise feedback that supports perception and manipulation of digital objects. Across mobile, wearable, and VR systems, effective haptics improve realism, reaction times, and emotional engagement [3, 14, 24].

However, a fundamental challenge in haptic design is tactile subjectivity. It relies heavily on individual physiology (e.g., mechanoreceptor density, skin thickness) and psychology (e.g., preferences for subtle versus intense sensations). A signal designed as a “gentle notification” may be imperceptible to elderly users yet uncomfortable for younger ones. Consequently, predefined haptic libraries often fail to deliver the intended user experience.

To accommodate such individual differences, personalization is essential. Traditional approaches often adopt “Method of Adjustment” interfaces [12, 16], where users manually tune sliders for parameters such as frequency and amplitude. While flexible, this approach assumes that users understand these parameters and can effectively navigate a high-dimensional space, imposing extra cognitive load and time burdens on non-expert users. Alternatively, conventional personalization methods may ask users to rate a series of signals (e.g., via a 1–7 Likert scale [18]). However, such methods exhibit fundamental psychometric limitations during prolonged sessions:

- **Score Drift:** Users lack a stable internal reference scale [11, 17]. A stimulus rated as “5” early in a session may be rated as “3” later as the user experiences a broader range of sensations, introducing noise into the dataset.
- **User Fatigue:** Providing absolute quantitative judgments for complex sensory inputs is cognitively demanding. Fatigue leads to data quality degradation and inconsistent responses [17].

Humans are effective comparators, but poor absolute meters [7]. While replacing absolute ratings with pairwise comparisons (A/B testing) addresses the issues of drift and fatigue, binary choice data is significantly sparser than rating scores. Furthermore, selecting which two signals to compare to maximize the informational value of each response poses a challenge in balancing personalization needs with the minimization of user effort.

Inspired by recent advances in preference learning, we seek to overcome these limitations by proposing an interactive **confidence-aware Vibrotactile Preference Learning (VPL)** system. Figure 1 illustrates the framework of the proposed system.

This work makes the following contributions:

- (1) **A vibrotactile personalization preference model:** We use pairwise comparisons and incorporate user-reported confidence to build a Gaussian-process (GP) preference model that captures the utility of vibrotactile signals. We select the next A/B comparison by maximizing the expected information gain over vibrotactile signal parameters, enabling data-efficient learning of the global preference structure under a limited interaction budget.
- (2) **Empirical validation via a user study:** In a user study with an Xbox controller ($N = 13$), we show that VPL achieves high prediction accuracy and produces user-satisfying recommendations within 40 comparisons, aligns closely with manually tuned optima, and maintains low subjective workload.
- (3) **Design insights for uncertainty-aware haptics:** We analyze sampling behaviors (e.g., exploration vs. refinement)

and discuss practical implications for deploying preference-based personalization in vibrotactile settings with perceptual ambiguity and substantial individual variability.

To comprehensively evaluate the effectiveness of the system, we investigate the following three Research Questions (RQs):

- (1) Can the proposed confidence-aware preference learning system recover an individual’s latent vibrotactile preference structure with high predictive accuracy within a bounded interaction budget (40 comparisons)?
- (2) Do users perceive the system’s recommended vibrotactile signal as matching their subjective preferences and expectations?
- (3) Is the workload and time cost associated with the 40-round personalization process acceptable for short-session personalization in practice?

We evaluated our Vibrotactile Preference Learning (VPL) system in a user study ($N = 13$) using vibrotactile feedback from a Microsoft Xbox controller. Additionally, we developed a control system adapted to this controller to validate the learned preferences. The results indicate that our method achieves excellent data efficiency, converging to a stable preference model within 40 rounds. In subjective evaluations, users reported high satisfaction with the recommended signals. Furthermore, NASA-TLX results indicate that the cognitive workload remains low, with users generally finding the interaction intuitive and easy to use. Our code is available as open-source¹.

2 Related Work

Haptic systems are increasingly expected to move beyond generic feedback toward *user-specific* experiences, as tactile perception and comfort vary substantially across individuals, contexts, and body locations. We review prior work along three threads most relevant to our contribution: (1) personalized haptic design, (2) preference learning with uncertainty quantification, and (3) preference-driven learning in interactive systems.

2.1 Personalized Haptic Design

A growing body of HCI work explores personalization to improve the usability, comfort, and effectiveness of haptic systems. One line of research focuses on *applied, user-facing systems*, such as *Calm Commute*, which uses personalized haptic breathing guidance for stress management in drivers [2], and a clippable pneumatic-haptic device by Choi *et al.* that allows users to tailor actuation patterns to individual comfort and body sites [8].

Beyond applications, researchers have proposed higher-level interaction abstractions for reasoning about vibrotactile effects. Examples include Seifi *et al.*’s “affective handles” for tuning vibrations by affective intent [27] and *TacTalk*, which enables conversational customization of haptics via natural language [22]. Collectively, these works underscore both the importance of personalization and the need for interaction paradigms that reduce user burden while supporting exploration of complex haptic design spaces.

¹<https://github.com/HaRVI-Lab/vibrotactile-preference-learning>

2.2 Preference Learning and Uncertainty Quantification

Preference learning replaces absolute ratings with pairwise comparisons, a paradigm rooted in psychometrics and psychophysics [5, 29]. In machine learning, probabilistic models such as GP-based preference learning infer latent utilities from sparse comparisons while quantifying uncertainty [10]. Active preference learning improves data efficiency by selecting informative queries [6, 7], with extensions to interactive optimization and reward learning that emphasize query efficiency under limited human feedback [4].

Crucially, human judgments are noisy and context-dependent. Recent work has begun to explicitly model multiple sources of uncertainty in preference learning, including human response variability [25]. Motivated by this, we incorporate user-reported confidence into GP-based preference learning to modulate trust in each comparison, while using information-gain-driven query selection under a limited interaction budget.

2.3 Preference-Driven Learning in Interactive Systems

Beyond estimating preferences, prior work uses preferences to *drive* learning, generation, and search. In haptics, Lu *et al.* demonstrated preference-driven texture modeling through interactive generation and search, showing how iterative comparisons guide systems toward user-aligned tactile outcomes [20]. Similar interactive loops appear in preference-based reinforcement learning and reward learning, where human comparisons supervise optimization [4, 30].

More broadly, preference-driven training underpins reinforcement learning from human feedback (RLHF). Foundational work learned reward models from human comparisons [9], with later applications to text generation, including summarization [28] and instruction-following language models [23]. Across modalities, these approaches highlight human preferences as flexible supervision when explicit metrics are unavailable. Our work extends this paradigm to vibrotactile feedback, emphasizing efficient, uncertainty-aware preference acquisition for personalization.

3 System & Method

3.1 Vibrotactile Signal Parameterization

We generate vibrotactile signals using a standard Microsoft Xbox controller equipped with two asymmetric actuators: a low-frequency (left) motor and a high-frequency (right) motor. To enable expressive yet manageable personalization, we parameterize the signal space using four dimensions:

$$\phi \triangleq [I, T, R, G]^T \in \Phi \subset \mathbb{R}^4, \quad (1)$$

where the parameters and their hardware-constrained ranges are defined as:

- **Intensity** ($I \in [0.20, 1.00]$): The overall vibration magnitude.
- **Motor Balance** ($T \in [0.00, 1.00]$): The distribution of intensity between the two motors, controlling the tactile “texture.” The left (M_L) and right (M_R) motor strengths are computed as:

$$M_L = I \cdot T, \quad M_R = I \cdot (1 - T). \quad (2)$$

- **Rhythm** ($R \in [0.60, 4.00]$ Hz): The frequency of vibration pulses. The cycle period is given by $P = 1/R$.
- **Grain** ($G \in [0.10, 0.70]$): The duty cycle of each pulse, affecting perceived sharpness.

To ensure signal stability on the hardware, we enforce a minimum inter-pulse gap of 45 ms and a minimum pulse attack of 20 ms. The effective pulse duration d is therefore constrained by:

$$d = \min(P \cdot G, \max(20\text{ms}, P - 45\text{ms})). \quad (3)$$

Each signal is rendered as a sequence of these pulses for a duration of 3 seconds.

3.2 Problem Formulation

We formulate vibrotactile personalization as an interactive preference learning problem over the signal space defined above. To facilitate optimization, we linearly normalize the physical parameter space Φ to a unit hypercube $\mathcal{X} = [0, 1]^4$. For the remainder of this paper, we denote a normalized query point simply as $x \in \mathcal{X}$.

We assume a latent utility function $f : \mathcal{X} \rightarrow \mathbb{R}$ governs the user’s preference. The personalization process proceeds in rounds. At round t , the system presents a pair of signals (x_t^A, x_t^B) . The user provides:

- (1) A binary preference $y_t \in \{+1, -1\}$, where +1 indicates $x_t^A \succ x_t^B$.
- (2) A confidence score $c_t \in \{1, \dots, 5\}$ (Likert scale: 1 = Very Unsure, 5 = Very Sure).

The collected dataset at time t is denoted as $\mathcal{D}_t = \{(x_i^A, x_i^B, y_i, c_i)\}_{i=1}^t$.

Our goal is to learn a user-specific latent utility function f from a small number of pairwise comparisons:

$$\hat{f} = \arg \max_f p(\mathcal{D}_T | f), \quad (4)$$

and use the learned model to guide query selection and make recommendations.

3.3 Uncertainty-Aware Preference Modeling

GP prior. We place a Gaussian Process prior on the latent utility function $f : \mathcal{X} \rightarrow \mathbb{R}$ over normalized parameters $x \in [0, 1]^4$:

$$f \sim \mathcal{GP}(0, k(\cdot, \cdot)), \quad (5)$$

with an isotropic RBF kernel

$$k(x, x') = s^2 \exp\left(-\frac{\|x - x'\|_2^2}{2\ell^2}\right). \quad (6)$$

Confidence-aware probit likelihood. At round i , the user compares (x_i^A, x_i^B) and returns a preference $y_i \in \{+1, -1\}$ (where +1 denotes $x_i^A \succ x_i^B$) and a confidence score $c_i \in \{1, \dots, 5\}$. Let $\Delta_i \triangleq f(x_i^A) - f(x_i^B)$. We map confidence to a human-uncertainty noise scale $\sigma_i \triangleq u(c_i)$ with $u(1) > \dots > u(5) > 0$, so lower confidence corresponds to noisier comparisons. We additionally include a small baseline jitter $\lambda > 0$ to account for residual variability not captured by confidence (e.g., actuator/control noise). The probit likelihood is

$$p(y_i | f, c_i) = \Phi\left(\frac{y_i \Delta_i}{\sqrt{2\lambda^2 + \sigma_i^2}}\right), \quad (7)$$

where $\Phi(\cdot)$ is the standard normal CDF.

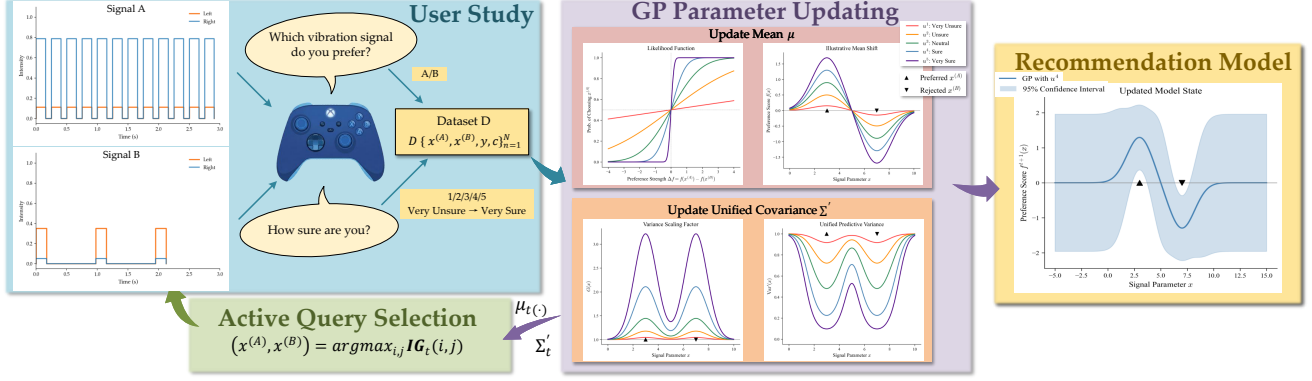


Figure 2: VPL Pipeline: Users make pairwise vibrotactile judgments and report confidence; a confidence-aware GP preference model updates, selects informative queries via information gain, and outputs an uncertainty-calibrated recommendation.

MAP + Laplace posterior. Let \mathbf{f} be the latent utilities evaluated on the unique queried points, with GP prior $\mathbf{f} \sim \mathcal{N}(\mathbf{0}, K)$. We estimate the posterior mode by maximizing the log posterior

$$\hat{\mathbf{f}} = \arg \max_{\mathbf{f}} \left[\sum_{i=1}^t \log \Phi \left(\frac{y_i \Delta_i}{\sqrt{2\lambda^2 + u(c_i)^2}} \right) - \frac{1}{2} \mathbf{f}^\top K^{-1} \mathbf{f} \right]. \quad (8)$$

We then obtain an approximate posterior $(\mu_t(\cdot), \Sigma_t(\cdot, \cdot))$ with a Laplace approximation around $\hat{\mathbf{f}}$, following standard practice [26] and UUPL [25].

3.4 Active Query Selection

To reduce user effort under a fixed interaction budget, we select the next query pair (x^A, x^B) by maximizing expected information gain (IG) [15, 21]. Since the user's future confidence is unknown at selection time, we use a fixed nominal noise scale σ_{nom} for look-ahead evaluation.

Pairwise statistics under the current posterior. For a candidate pair, define $\delta = f(x^A) - f(x^B)$. Under the Laplace GP posterior, δ is approximately Gaussian:

$$\delta \sim \mathcal{N}(m, v), \quad (9)$$

where $m = \mu_t(x^A) - \mu_t(x^B)$ and $v = \text{Var}_t(f(x^A) - f(x^B))$.

Information gain objective. Using the probit observation model with nominal noise σ_{nom} , the predictive probability of choosing A is

$$p \triangleq P(y = +1 \mid x^A, x^B, \mathcal{D}_t) \approx \Phi \left(\frac{m}{\sqrt{v + \sigma_{\text{nom}}^2}} \right). \quad (10)$$

We compute IG in a convenient 1D form (binary entropy $h(q) = -q \log q - (1-q) \log(1-q)$):

$$\text{IG}(x^A, x^B) = h \left(\Phi \left(\frac{m}{\sqrt{v + \sigma_{\text{nom}}^2}} \right) \right) - \mathbb{E}_{z \sim \mathcal{N}(0,1)} \left[h \left(\Phi \left(\frac{m + \sqrt{v} z}{\sigma_{\text{nom}}} \right) \right) \right]. \quad (11)$$

Bıyık et al. [4] derive a closed-form approximation for (11).

We select the next pair by

$$(x_{t+1}^A, x_{t+1}^B) = \arg \max_{x^A, x^B \in \mathcal{X}_{\text{cand}}} \text{IG}(x^A, x^B). \quad (12)$$

Algorithm 1 Confidence-Aware VPL Interaction Loop

Require: GP prior $f \sim \mathcal{GP}(0, k)$, confidence mapping $u(\cdot)$, jitter λ , nominal noise σ_{nom}

- 1: Initialize $\mathcal{D}_0 \leftarrow \emptyset$
- 2: **for** $t = 1$ to T **do**
- 3: Sample candidate set $\mathcal{X}_{\text{cand}} \subset [0, 1]^4$
- 4: Select $(x_t^A, x_t^B) \leftarrow \arg \max_{x, x' \in \mathcal{X}_{\text{cand}}} \text{IG}(x, x')$
- 5: **User:** compare (x_t^A, x_t^B) , return (y_t, c_t)
- 6: $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{(x_t^A, x_t^B, y_t, c_t)\}$
- 7: Update posterior via MAP (Eq. 8) and Laplace to obtain (μ_t, Σ_t)
- 8: **end for**
- 9: **Recommend** $\hat{x}^* = \arg \max_{x \in [0,1]^4} \mu_T(x)$

3.5 Algorithmic Pipeline

Figure 2 illustrates the interaction loop, and Algorithm 1 provides the procedural summary. At each round t , the system (i) samples a candidate set $\mathcal{X}_{\text{cand}}$, (ii) evaluates each candidate pair using the information-gain criterion (Eq. 11), and (iii) queries the most informative pair to the user. The user returns both a preference label y_t and a confidence score c_t , which is mapped to the comparison noise scale via $u(\cdot)$ in the probit likelihood (Eq. 7). The GP posterior is then updated via MAP estimation (Eq. 8) with a Laplace approximation to obtain (μ_t, Σ_t) . After T rounds, we recommend the signal \hat{x}^* that maximizes the posterior mean.

3.6 User Modeling

To clarify the scope of our approach, we explicitly outline assumptions about user behavior, the user attributes captured by our model, and the factors currently excluded.

Assumptions. We rely on two fundamental assumptions regarding the user's psychophysical state:

- **Stationarity:** We assume the user's latent preference function is approximately stationary throughout a short session (40 iterations), i.e., there is no systematic concept drift or fundamental shift in taste within this time window.

- **Local Smoothness:** We assume vibrotactile preferences vary smoothly with respect to the signal parameters. If a user prefers a particular setting, nearby settings in the (normalized) parameter space are expected to have similar utility, rather than exhibiting abrupt discontinuities.

Modeled Aspects. Our system explicitly captures two dimensions of the user:

- **Sensory Preference Structure:** The model learns a latent utility function $f(x)$, representing the user’s subjective mapping from signal parameters to perceived utility.
- **Uncertainty and Calibration:** Beyond estimating *what* the user prefers, we model *how certain* they are. By integrating confidence scores, we treat the user as a variable-noise oracle via the mapping $u(c)$, which modulates the comparison-noise scale in the likelihood. In addition, the GP posterior provides model uncertainty over $f(x)$, supporting uncertainty-aware query selection and conservative recommendations when appropriate.

Unmodeled Aspects. The current framework focuses on within-session sensory evaluation and excludes:

- **Context Dependence:** We do not model how preferences vary with situational factors (e.g., task demands, attentional load, environment, grip style, body location) or the semantic meaning of haptic events.
- **Temporal Adaptation and Non-Stationarity:** The model does not account for sensory adaptation, habituation, learning effects, or systematic drift across sessions.
- **User Fatigue and Strategic Behavior:** We do not explicitly model fatigue-related degradation, careless responding, or strategic behavior. Behavioral signals such as response time are not utilized.

4 Experimental Evaluation

4.1 Simulation

To evaluate the algorithm’s convergence without human variability, we conducted an automated simulation using a synthetic user represented by a fixed Ground-Truth (GT) utility function.

The simulation mirrors the active learning loop used in the user study. In each iteration, the Gaussian Process (GP) selects a query pair (x^A, x^B) that maximizes information gain. The simulator then acts as an oracle: it selects the option with the higher GT utility and assigns a confidence score (1–5) proportional to the normalized utility difference between the two options. This feedback is used to update the GP posterior.

The results, shown in Fig. 3, demonstrate accurate convergence. The Information Gain (Fig. 3a) declines steadily, indicating that the model effectively reduces global uncertainty over time. The Spearman rank correlation between the learned model and the GT function (Fig. 3b) increases and stabilizes, confirming that the algorithm accurately recovers the underlying user preferences.

4.2 Participants

We recruited 13 participants (P01–P13; 8 female, 5 male) aged 18 to 43 years ($M = 28.8, SD = 7.2$). Participants reported varying gaming frequencies (Never: 4, Occasionally: 8, Often: 1) and controller

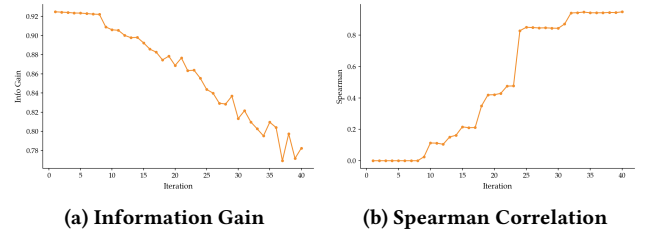


Figure 3: Simulation results

familiarity (Low: 3, Medium: 3, High: 7). The study was approved by the University Institutional Review Board, and all participants gave informed consent.

4.3 Experimental Procedure

The study was conducted in a controlled environment and followed a three-step protocol: Active Learning (Phase 1), Model Validation, and Manual Adjustment (Phase 2).

4.3.1 Phase 1: Active Preference Learning. This phase consisted of 40 iterations aimed at training the Gaussian process (GP) model. The signal generation was initialized with a recorded time-based random seed for each session. In each iteration, participants used the left analog stick of the controller to navigate between two candidate vibrations generated as described in Section 3.1 (Option A and Option B). They could press the ‘X’ button to play the tactile effects as many times as needed. After comparing the signals, participants selected their preferred option and rated their decision confidence on a scale of 1 to 5. The parameters for u in the human-uncertainty noise scale (Section 3.3) were calibrated via pilot studies and simulations: $\{u(1) = 9.0, u(2) = 3.35, u(3) = 1.7, u(4) = 0.66, u(5) = 0.01\}$.

4.3.2 Validation Round. Immediately following the training phase, the system automatically generated a validation set to evaluate the quality of the learned model.

Stimuli Generation. The system first identified the signal with the highest posterior mean (denoted as x_{best}). It then sampled 6 additional points from the parameter space, enforcing a minimum distance constraint (threshold = $0.05 \times$ diagonal length) to ensure diversity. The posterior means of these 7 points (1 best + 6 sampled) were calculated to facilitate specific pairwise comparisons.

Comparison Tasks. To evaluate the accuracy of the learned preference model, participants were asked to perform a set of validation tasks. The system generated 4 specific comparison pairs based on the learned utility function. For each pair, participants compared the two signals (labeled Option A and Option B) and selected their preferred one. The goal was to determine if the participants’ actual choices aligned with the model’s predicted rankings. The pairs were designed to test different aspects of the model:

- (1) **Anchor (Easy):** x_{best} vs. x_{worst} . Sanity check to confirm the participant prefers the optimal signal over the worst one.
- (2) **Anchor (Medium):** x_{best} vs. x_{mid} . Tests whether the model can correctly distinguish the optimal signal from a medium-utility alternative.
- (3) **Global Trade-off:** Two signals with a large Euclidean distance and significant utility difference. Evaluates if the model

correctly predicts preferences between two physically very different signals.

- (4) **Consistency Check:** A repetition of the “Anchor (Medium)” pair with the positions of A and B swapped. Detects potential position bias or random guessing by the participant.

4.3.3 Phase 2: Free-Adjustment Task. To obtain a set of user-defined optima as a high-quality reference, participants entered a free-adjustment phase. They were presented with a graphical interface allowing free control over the four vibration parameters (intensity, motor balance, rhythm, grain). Participants were asked to explore the space and manually tune the parameters to create their ideal vibration. This process was repeated three times, and the top three favorite configurations were recorded.

4.4 Measures

We employed standard psychometric scales, custom questionnaires, and system logs to evaluate task workload and system performance.

System Logs and Objective Data. Throughout the experiment, the system automatically logged detailed interaction data for quantitative analysis:

- **Interaction History:** For all 40 training iterations and 4 validation rounds, we recorded the physical parameters of the candidate pairs (x^A, x^B), the participant’s binary choice ($y \in \{A, B\}$), the self-reported confidence level, and the calculated Information Gain (IG) for each query.
- **Convergence Data:** We logged the model’s final recommended signal (x_{rec}) derived from the posterior mean after the active learning phase.
- **User-Defined Favorites:** We recorded the parameters of the top 3 “favorite” signals manually created by the participants in Phase 2, along with their creation timestamps. These serve as the reference standard for evaluating the model’s recommendation accuracy.

Subjective Workload (NASA-TLX). Participants completed the NASA Task Load Index (NASA-TLX) to assess the workload across six dimensions [13]: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each dimension was rated on a scale from 0 (Low) to 20 (High).

User Experience and System Usability. We designed a custom 7-point Likert scale questionnaire (1 = Strongly Disagree, 7 = Strongly Agree) to evaluate six specific aspects of the system:

- **Ease of Use:** “The system was easy to use”
- **Controllability:** “I understood what each parameter controls”
- **Decision Support:** “The uncertainty rating helped me make decisions”
- **Satisfaction:** “How satisfied are you with the final recommended signal?”
- **Expectation Match:** “Did the recommended signal match your expected preferred haptic feeling?”
- **Comfort:** “Overall comfort was acceptable”

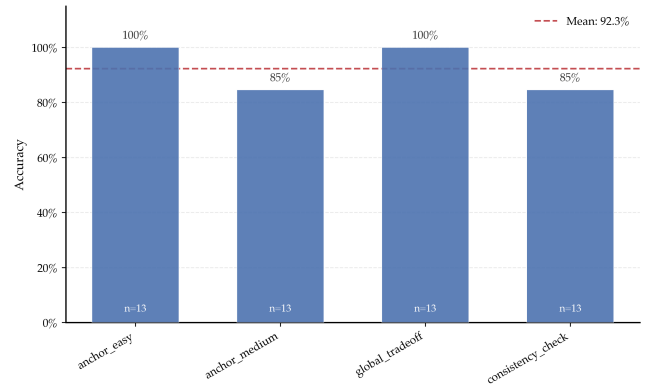


Figure 4: Validation accuracy grouped by item type

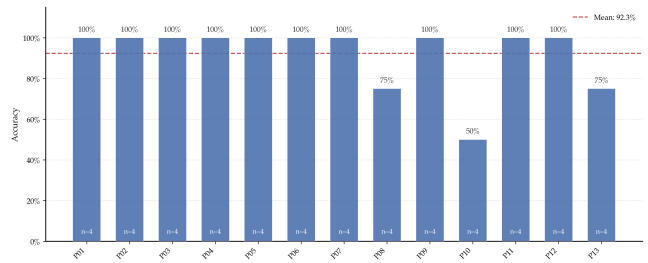


Figure 5: Validation accuracy grouped by participant

5 Results

In this section, we present the evaluation results to answer our three research questions: (1) Can the system learn an accurate, individualized haptic preference distribution within a limited number of interaction rounds? (**RQ1**); (2) Are users satisfied with the signals recommended by the system? (**RQ2**); and (3) Are the interaction workload and usability within an acceptable range? (**RQ3**).

5.1 Objective Metrics (RQ1)

5.1.1 Held-out Validation Accuracy. To assess whether the learned model generalizes to unseen data, we analyzed the prediction accuracy on the held-out validation set.

Figure 4 summarizes the validation accuracy grouped by item type, and Figure 5 shows the results grouped by participant. The overall mean pairwise accuracy reached **92.3%** (range: 85%–100%) with a mean user confidence of **4.38/5**. This high accuracy demonstrates that the GP model successfully captured the latent structure of participants’ haptic preferences and could reliably predict choices in held-out comparisons.

5.1.2 Consistency with User-Defined Optima. To further verify the model’s validity, we examined the alignment between the model’s predictions and the user-defined preferences. We took the “best” signal manually tuned by each participant in Phase 2 and evaluated its standing within their learned preference model. Specifically, we calculated the percentile rank of this user-defined signal within the model’s posterior mean distribution over the entire search space. The results show that the user-defined best signals achieved an average percentile rank of **81.2%**. This indicates a strong alignment:

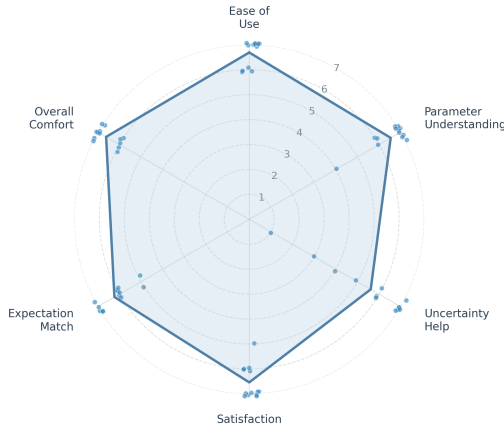


Figure 6: Post-study subjective ratings (N=13)

the signals that users subjectively felt were “best” were also predicted by the model to be in the top tier (top $\approx 19\%$) of the utility space.

Summary for RQ1: Supported by the high held-out validation accuracy (92.3%) and the strong consistency with user-defined optima (81.2% percentile rank), we conclude that the system is capable of learning a stable and accurate individual haptic preference distribution within 40 iterations.

5.2 Subjective Evaluation (RQ2 & RQ3)

5.2.1 System Satisfaction (RQ2). We evaluated user satisfaction through a post-study questionnaire using a 7-point Likert scale (1=Strongly Disagree, 7=Strongly Agree). As shown in Figure 6, participants reported high ratings for **Usability** ($M = 6.69, SD = 0.48$) and **Comfort** ($M = 6.62, SD = 0.51$). **Controllability** ($M = 6.54, SD = 0.88$) and **Overall Satisfaction** ($M = 6.54, SD = 0.66$) were also highly rated. Notably, the **Expectation Match** score ($M = 6.23, SD = 0.73$) suggests that the algorithmically recommended signals closely matched users’ mental models of their preferred haptics. The **Decision Support** rating showed higher variability ($M = 5.62, SD = 1.89$). This is because the uncertainty help operates on the final recommendation generation, without being explicitly perceived by users as an improvement during the interaction.

5.2.2 Workload Assessment (RQ3). The active-learning phase lasted 763.31 ± 188.79 s ($12m43s \pm 3m09s$), including repeated playback, showing that the 40-comparison budget remained a short, bounded session. We assessed interaction workload using the NASA-TLX scale (0–20). Figure 7 presents the scores across six dimensions. The overall workload was low ($M = 3.71 \pm 1.80$). Detailed analysis reveals that **Effort** ($M = 7.69 \pm 6.49$) and **Mental Demand** ($M = 5.15 \pm 3.72$) were the highest among dimensions, reflecting the cognitive attention required to distinguish subtle haptic differences. Conversely, **Physical Demand** ($M = 2.54 \pm 2.82$), **Temporal Demand** ($M = 3.00 \pm 3.39$), and **Frustration** ($M = 1.62 \pm 2.18$) remained very low. Crucially, the **Performance** score was low ($M = 2.23 \pm 2.98$). Note that on the NASA-TLX scale, a lower Performance score indicates higher self-perceived success (i.e., less failure). This confirms that participants felt successful in guiding the system to their preferred outcomes.

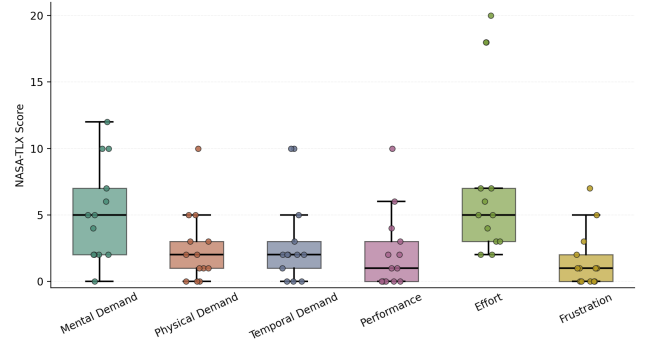


Figure 7: NASA-TLX workload ratings.

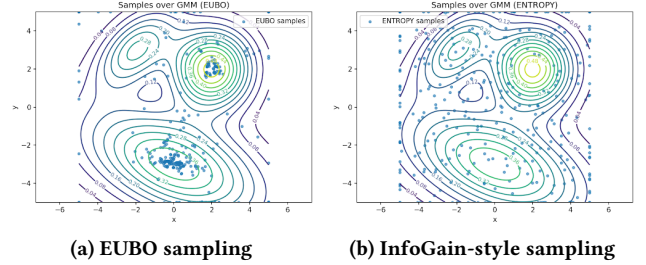


Figure 8: Comparison of sampling behaviors on GMM

Summary for RQ2 & RQ3: The high satisfaction scores confirm that users were satisfied with the recommended signals (RQ2). The NASA-TLX results demonstrate that despite the inherent mental effort required for psychophysical comparisons, the overall workload and usability remained within an acceptable range (RQ3).

6 Discussion

6.1 Sampling Strategy in Haptics

A key design choice in preference-based interactive haptics is the *sampling strategy*: which pair of signals to query next under a limited interaction budget. In this work, we adopt an **information-gain (InfoGain)** strategy that is *model-driven*: it prioritizes queries expected to maximally reduce uncertainty in the learned preference model, leading to broader exploration especially in regions where the model is uncertain or user judgments are inconsistent.

In contrast, **Expected Utility of the Best Option (EUBO)** is a commonly used acquisition function in preferential Bayesian optimization that is *solution-driven* [1, 19]. Rather than optimizing overall model fidelity, EUBO prioritizes query pairs that are expected to most improve the quality of the *final recommended signal* (i.e., the best candidate under the current posterior). Consequently, EUBO tends to more aggressively allocate samples densely around the currently promising region to refine a local neighborhood and reach a high-utility optimum quickly.

Figure 8 contrasts sampling behaviors on a multimodal landscape: EUBO concentrates near a single high-utility mode, while InfoGain explores broadly to resolve global structure. This leads to different trade-offs for haptic applications. For *short, goal-oriented sessions* (e.g., quickly tuning a single “best-feeling” vibration style for a game event), EUBO can quickly converge to a satisfying signal. However, due to perceptual indistinguishability and hardware-limited resolution, it may waste budget on low-information comparisons near

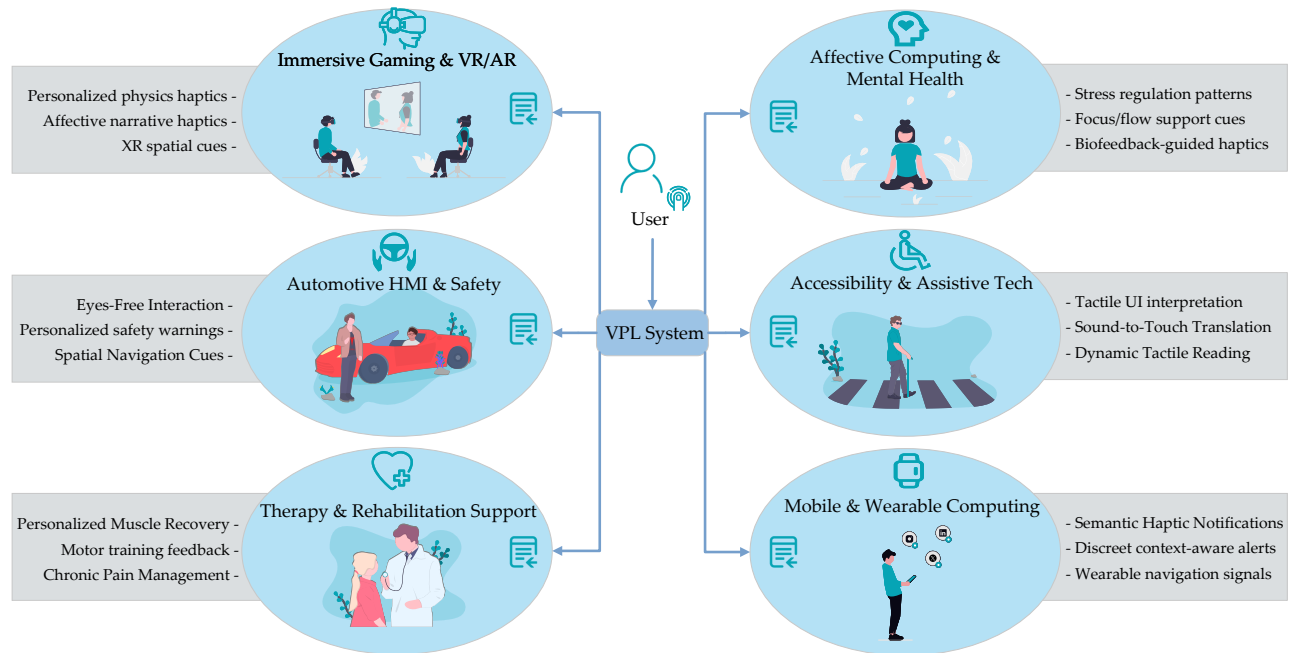


Figure 9: Illustrative application domains where VPL could support personalized vibrotactile feedback.

an apparent optimum, increasing user fatigue. In contrast, InfoGain better supports *robust and reliable* personalization by learning a calibrated preference landscape and identifying uncertain or indistinguishable regions, enabling stable recommendations and cross-context reuse (multiple haptic events) at the cost of slower convergence to a single optimum under a fixed budget.

Overall, EUBO and InfoGain represent complementary priorities: *fast best-signal optimization* vs. *faithful preference learning under uncertainty*. A promising direction is a hybrid schedule: use InfoGain early to map the preference structure and detect indistinguishable regions, then switch to EUBO refinement to polish the final recommendation.

6.2 Applications

Fig. 9 outlines domains where VPL can support personalized vibrotactile feedback. We emphasize two deployment-relevant advantages: **(i) robustness under hard constraints** and **(ii) uncertainty awareness that avoids ambiguous regions**.

Advantage 1: Robustness under hard constraints. Instead of optimizing only a single local optimum, VPL learns an uncertainty-calibrated *global* preference landscape. When deployments impose hard limits (e.g., maximum intensity, restricted frequency ranges, bandwidth limits, comfort/safety caps, or distinctiveness requirements across alerts), VPL can *directly* choose a strong recommendation **within the feasible set**, without re-running optimization or preference elicitation.

Automotive HMI & safety. Vehicle haptics are safety-critical and constrained by intensity caps, timing limits, and *separability* across warning types (e.g., lane departure, collision, navigation), which may vary by context. When certain designs are infeasible, VPL can identify a strong feasible alternative from its global model.

Mobile & wearables. Haptics often encode multiple alerts under intensity/duration caps, limited actuator bandwidth/OS scheduling, and explicit *distinctiveness* requirements. The goal is a usable set of preferred, discriminable signals; VPL can propose alternatives when a preferred design violates constraints or conflicts with other alerts, maintaining quality under changing limits (e.g., silent mode, workouts, long-wear comfort).

Therapy & rehabilitation. Wellbeing applications impose conservative safety bounds and user tolerance may change over time. VPL supports constraint-aware personalization within a safe envelope and provides stable fallbacks when constraints tighten, enabling graded programs instead of over-focusing near a potentially unreliable local optimum.

Immersive gaming and VR/AR. These domains benefit from fast personalization of event-specific haptic styles, especially when multiple events must be pleasant yet mutually distinguishable.

Advantage 2: Uncertainty awareness to avoid ambiguous regions. VPL explicitly models uncertainty in user judgments, which is common due to perceptual indistinguishability and context effects. It learns not only where utility is high but also where confidence is high, allowing it to avoid low-confidence regions. This is useful in affective computing and mental health support, where a conservative, lower-risk pattern is preferable when judgments are ambiguous or preferences fluctuate.

6.3 Limitations

Limited consistency under a finite query budget. With a fixed budget of 40 interactions, the learned preference landscape should be interpreted as a *low-resolution* approximation. Very similar vibrotactile signals may be perceptually indistinguishable, and comparisons between nearby points in the parameter space are dominated by response noise, limiting fine-grained consistency.

Restricted parameterization and scalability. We model vibrations using four global parameters and a fixed 3 s duration, which supports efficient interaction but excludes richer spatiotemporal patterns (e.g., complex envelopes or multi-actuator effects). Extending VPL to higher-dimensional or compositional signal spaces is promising but would increase both computational cost and required user interactions.

Subjective confidence and fixed noise mapping. Our method depends on self-reported confidence and a fixed confidence-to-noise mapping. Because confidence is subjective and may vary across users or over time, a single mapping may be suboptimal. Future work could adapt or personalize this mapping through calibration or online learning.

7 Conclusion

We presented Vibrotactile Preference Learning (VPL), an interactive method that personalizes vibrotactile feedback using simple A/B comparisons, together with users' self-reported confidence. VPL uses the confidence ratings to decide how much to trust each response, and it chooses the next comparison to ask so that each interaction is as informative as possible. In a user study ($N = 13$) with an Xbox controller, VPL reached **92.3%** accuracy on the validation set after **40** rounds, with an average confidence of **4.38/5**, and participants reported high satisfaction and low workload. Future work will extend VPL to richer spatiotemporal haptic patterns and make the approach easier to scale and deploy in real-world systems.

References

- [1] Raul Astudillo and Peter I Frazier. 2023. qEUBO: A Decision-Theoretic Acquisition Function for Preferential Bayesian Optimization. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS) (Proceedings of Machine Learning Research)*.
- [2] Stephanie Balters, Matthew L. Mauriello, So Yeon Park, James A. Landay, and Pablo E. Paredes. 2020. Calm Commute: Guided Slow Breathing for Daily Stress Management in Drivers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–19.
- [3] Premankur Banerjee, Jiakuan Wang, Lauren Tomita, Mia P. Montiel, and Heather Culbertson. 2025. Virtual Encounters of the Haptic Kind: Towards a Multi-User VR System for Real-Time Social Touch. arXiv:2502.13421 [cs.HC] <https://arxiv.org/abs/2502.13421>
- [4] Erdem Bıyık, Nicolas Huynh, Mykel J. Kochenderfer, and Dorsa Sadigh. 2024. Active Preference-Based Gaussian Process Regression for Reward Learning and Optimization. *The International Journal of Robotics Research* 43, 5 (2024), 665–684. doi:10.1177/02783649231208729
- [5] Ralph Allan Bradley and Milton E. Terry. 1952. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 39, 3–4 (1952), 324–345.
- [6] Eric Brochu, Vlad M. Cora, and Nando de Freitas. 2010. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. arXiv:1012.2599.
- [7] Eric Brochu, Nando de Freitas, and Abhijeet Ghosh. 2007. Active Preference Learning with Discrete Choice Data. In *Advances in Neural Information Processing Systems*.
- [8] Kyung Yun Choi, Nizer ElHaouij, Jeon Lee, Rosalind W. Picard, and Hiroshi Ishii. 2022. Design and Evaluation of a Clippable and Personalizable Pneumatic-Haptic Feedback Device for Breathing Guidance. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–36.
- [9] Paul F. Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems* 30 (2017).
- [10] Wei Chu and Zoubin Ghahramani. 2005. Preference Learning with Gaussian Processes. In *Proceedings of the 22nd International Conference on Machine Learning (ICML)*. 137–144.
- [11] Hugh J. Foley, David V. Cross, and James A. O'Reilly. 1990. Pervasiveness and Magnitude of Context Effects: Evidence for the Relativity of Absolute Magnitude Estimation. *Perception & Psychophysics* 48, 6 (1990), 551–558. doi:10.3758/BF03211601
- [12] George A. Gescheider. 1997. *Psychophysics: The Fundamentals*. Lawrence Erlbaum Associates.
- [13] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183. doi:10.1016/S0166-4115(08)62386-9
- [14] Eve Hoggan, Stephen Brewster, and Jody Johnston. 2008. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. Association for Computing Machinery, 1573–1582. doi:10.1145/1357054.1357300
- [15] Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. 2011. Bayesian Active Learning for Classification and Preference Learning. arXiv:1112.5745 [stat.ML] <https://arxiv.org/abs/1112.5745>
- [16] Frederick A. A. Kingdom and Nicolaas Prins. 2016. *Psychophysics: A Practical Introduction*. Academic Press.
- [17] Svetlana Kiritchenko and Saif M. Mohammad. 2017. Best-Worst Scaling More Reliable than Rating Scales: A Case Study on Sentiment Intensity Annotation. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL): Volume 2, Short Papers*. <https://aclanthology.org/E17-2074/>
- [18] Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of psychology* (1932).
- [19] Michael R. Lin, Elena Makarova, Eytan Bakshy, and Peter I Frazier. 2022. Preference Exploration for Efficient Bayesian Optimization with Multiple Outcomes. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS) (Proceedings of Machine Learning Research, Vol. 151)*.
- [20] Shihan Lu, Mianlun Zheng, Matthew C. Fontaine, Stefanos Nikolaidis, and Heather Culbertson. 2022. Preference-Driven Texture Modeling Through Interactive Generation and Search. *IEEE Transactions on Haptics* 15, 3 (2022), 508–520.
- [21] David J. C. MacKay. 1992. Information-Based Objective Functions for Active Data Selection. *Neural Computation* 4, 4 (1992), 590–604. doi:10.1162/neco.1992.4.4.590
- [22] Anchit Mishra and Oliver Schneider. 2025. TacTalk: Personalizing Haptics Through Conversation. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces (CUI '25)*. Association for Computing Machinery. doi:10.1145/3719160.3736638
- [23] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training Language Models to Follow Instructions with Human Feedback. arXiv:2203.02155.
- [24] Claudio Pacchierotti, Stephen Sinclair, Massimiliano Solazzi, Antonio Frisoli, Vincent Hayward, and Domenico Prattichizzo. 2017. Wearable Haptic Systems for the Fingertip and the Hand: Taxonomy, Review, and Perspectives. *IEEE Transactions on Haptics* 10, 4 (2017), 580–600. doi:10.1109/TOH.2017.2689006
- [25] Shaoting Peng, Haonan Chen, and Katherine Driggs-Campbell. 2025. Towards Uncertainty Unification: A Case Study for Preference Learning. arXiv preprint arXiv:2503.19317.
- [26] Carl Edward Rasmussen and Christopher K. I. Williams. 2006. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, USA.
- [27] Hasti Seifi, Matthew Chun, and Karon E. MacLean. 2018. Toward Affective Handles for Tuning Vibrations. *ACM Transactions on Applied Perception (TAP)* 15, 3 (2018), 1–23.
- [28] Nisan Stiennon, Long Ouyang, Jeffrey Wu, Daniel Ziegler, Ryan Lowe, Chelsea Voss, Alec Radford, Dario Amodei, and Paul F. Christiano. 2020. Learning to Summarize from Human Feedback. In *Advances in Neural Information Processing Systems*.
- [29] L. L. Thurstone. 1927. A Law of Comparative Judgment. *Psychological Review* 34, 4 (1927), 273–286.
- [30] Christian Wirth, Riad Akrou, Gerhard Neumann, and Johannes Fürnkranz. 2017. A Survey of Preference-Based Reinforcement Learning Methods. *Journal of Machine Learning Research* 18, 136 (2017), 1–46.