

ESTIMATING HAPTIC PARAMETERS FROM HUMAN PREFERENCES: A BAYESIAN  
ACTIVE LEARNING APPROACH

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF INFORMATICS OF  
THE MIDDLE EAST TECHNICAL UNIVERSITY  
BY

ROJDA TORAMAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF SCIENCE  
IN  
THE DEPARTMENT OF COGNITIVE SCIENCE

JANUARY 2026



**ESTIMATING HAPTIC PARAMETERS FROM HUMAN PREFERENCES: A  
BAYESIAN ACTIVE LEARNING APPROACH**

submitted by **Rojda Toraman** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science Department, Middle East Technical University** by,

Prof. Dr. Banu Günel Kılıç  
Dean, **Graduate School of Informatics**

\_\_\_\_\_

Assoc. Prof. Dr. Murat Perit Çakır  
Head of Department, **Cognitive Science**

\_\_\_\_\_

Assoc. Prof. Dr. Barbaros YET  
Supervisor, **Cognitive Science, Middle East Technical University**

\_\_\_\_\_

**Examining Committee Members:**

Assist. Prof. Dr. Umut Özge  
Cognitive Science, Middle East Technical University

\_\_\_\_\_

Assoc. Prof. Dr. Barbaros Yet  
Cognitive Science, Middle East Technical University

\_\_\_\_\_

Assoc. Prof. Dr. Ceren Tuncer Şakar  
Industrial Engineering, Hacettepe University

\_\_\_\_\_

**Date:**



**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

**Name, Surname: Rojda Toraman**

**Signature :**

# ABSTRACT

## ESTIMATING HAPTIC PARAMETERS FROM HUMAN PREFERENCES: A BAYESIAN ACTIVE LEARNING APPROACH

Toraman, Rojda

M.S., Department of Cognitive Science

Supervisor: Assoc. Prof. Dr. Barbaros YET

JANUARY 2026, 78 pages

Reproducing natural haptic sensations through artificial mechanical systems is a fundamental challenge in robotics. While haptic devices are controlled by precise numerical parameters, human perception of natural objects relies on subjective physical concepts that do not map directly to machine variables. Consequently, users cannot easily tune a device to match a target sensation by manually adjusting values. To solve this challenge, a Bayesian preference learning method is adopted to learn haptic parameters corresponding to real sensations based on human feedback. This approach utilizes pairwise comparisons, as humans make relative judgments with lower noise compared to absolute estimation. The human preference function is modeled using Gaussian Processes (GP), as this method is effective when the latent function's structure is unknown. An Active Learning framework is also adopted to intelligently select comparison pairs in order to speed up the model's convergence by maximizing information gain. The primary objective of this research is to evaluate the model's efficiency in converging to the true mechanical parameters of a reference sensation through pairwise comparisons. The study examines how the model's convergence changes as the number of variables increases, and evaluates whether coactive feedback speeds up the process. Overall, this study demonstrates that integrating human preferences into a Bayesian framework facilitates accurate replication of physical sensations while maintaining cognitively comfortable interactions.

Keywords: preference learning, bayesian models, active learning, haptic perception, coactive feedback

# ÖZ

## İNSAN TERCİHLERİNDEN HAPTİK PARAMETRELERİN TAHMİNİ: BAYESÇİ AKTİF ÖĞRENME YAKLAŞIMI

Toraman, Rojda

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Doç. Dr. Barbaros YET

Ocak 2026, 78 sayfa

Yapay mekanik sistemler aracılığıyla doğal haptik duyuların yeniden üretilmesi robotikte temel bir zorluktur. Haptik cihazlar hassas sayısal parametrelerle kontrol edilirken, insanların doğal nesnelere dair algısı makine değişkenleriyle doğrudan örtüşmeyen öznel fiziksel kavramlara dayanır. Sonuç olarak kullanıcılar, değerleri manuel olarak değiştirerek bir cihazı hedeflenen duyumla eşleştirecek şekilde kolayca ayarlayamazlar. Bu zorluğu çözmek için, gerçek duyumlara karşılık gelen haptik parametreleri insan geri bildirimine dayalı olarak öğrenmek amacıyla Bayesçi tercih öğrenme yöntemi benimsenmiştir. Bu yaklaşım, insanların mutlak tahminlere kıyasla daha düşük gürültü oranıyla görece yargılarda bulunabilmesi nedeniyle ikili karşılaştırmaları kullanmaktadır. Altta yatan fonksiyon yapısının bilinmediği durumlarda etkili bir yöntem olması nedeniyle insanların tercih fonksiyonu Gauss Süreçleri kullanılarak modellenmiştir. Ayrıca, bilgi kazancını maksimize ederek modelin yakınsamasını hızlandırmak ve karşılaştırma çiftlerini akıllıca seçmek için bir Aktif Öğrenme çerçevesi kullanılmıştır. Bu araştırmanın temel amacı, modelin ikili karşılaştırmalar yoluyla referans bir duyumun gerçek mekanik parametrelerine yakınsama verimliliğini değerlendirmektir. Çalışma, değişken sayısı arttıkça modelin yakınsamasının nasıl değiştiğini incelemekte ve koaktif geri bildirim süreci hızlandırıp hızlandırmadığını değerlendirmektedir. Sonuç olarak bu çalışma, insan tercihlerini Bayesçi bir çerçeveye entegre etmenin, bilişsel açıdan rahat etkileşimleri korurken fiziksel duyuların doğru bir şekilde yeniden üretilmesini kolaylaştırdığını göstermektedir.

Anahtar Kelimeler: tercih öğrenme, bayesçi modeller, aktif öğrenme, haptik algı, koaktif geribildirim

## ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Barbaros Yet. I feel truly lucky that our paths crossed. His guidance and support were invaluable.

I would also like to thank Ali Seçkin Kaplan, Yiğit Taşcıoğlu and Umut Özge, for their collaboration and assistance during the design and execution of the experiments. Working alongside such a dedicated research team made this study possible.

I would like to thank my love, Naip, for his limitless support and compassion. His presence gave me strength during the most difficult times.

I am very grateful to my mother Nalan, my father Oktay, and my little sister Hazan. They always supported and encouraged me in my long journey through academia.

Last but not least, a special thanks to my lovely cats, Kuzuş and Zizukuş; they always cheered me up and made the stressful times lighter.

## TABLE OF CONTENTS

ABSTRACT.....	iv
ÖZ.....	v
ACKNOWLEDGMENTS .....	vi
TABLE OF CONTENTS .....	vii
LIST OF TABLES .....	xii
LIST OF FIGURES .....	xiii
LIST OF ABBREVIATIONS .....	xiv
CHAPTERS	
1 INTRODUCTION .....	1
1.1 Problem Statement .....	1
1.2 Proposed Approach .....	2
1.3 Research Questions .....	3
1.4 Contributions of the Study .....	3
1.5 Organization of the Thesis .....	3
2 LITERATURE REVIEW .....	5
2.1 Preference Learning .....	6
2.1.1 Theories of Human Decision-Making .....	6

2.1.2	Preference Learning Types .....	7
2.1.2.1	Label Ranking .....	7
2.1.2.2	Instance Ranking .....	8
2.1.2.3	Object Ranking .....	8
2.1.3	Preference Learning Methods .....	9
2.1.3.1	Learning with Utility Functions .....	9
2.1.3.2	Learning with Preference Relations .....	10
2.1.3.3	Model-Based Preference Learning .....	11
2.1.4	Preference Data Types .....	12
2.1.5	Psychophysics of Pairwise Comparisons .....	14
2.2	Bayesian Inference .....	15
2.2.1	Advantages of Bayesian Inference over Frequentist Methods .....	16
2.2.2	Foundations of Bayesian Inference .....	17
2.2.2.1	Inference Methods .....	19
2.2.2.2	Posterior Predictive Distribution .....	19
2.2.3	Broader Applications of Bayesian Inference .....	19
2.2.4	Gaussian Processes .....	20
2.2.5	Active Learning .....	22
2.2.6	Bayesian Optimization and Acquisition Functions .....	22
2.2.7	Coactive Feedback .....	24
2.3	Haptic Perception .....	25
2.3.1	Perception of Stiffness (K), Damping (B) and Distance (D) .....	25

2.4	Summary .....	26
3	METHOD .....	29
3.1	Computational Method .....	29
3.1.1	Preference Query Type .....	29
3.1.2	Gaussian Process .....	30
3.1.2.1	Prior .....	30
3.1.2.2	Likelihood .....	31
3.1.2.3	Posterior .....	31
3.1.2.4	Optimal Estimate .....	32
3.1.3	Active Learning Framework .....	32
3.1.3.1	Acquisition Function: Penalized Expected Improvement .....	32
3.1.4	Experiment Method .....	33
3.1.4.1	Haptic Device .....	33
3.1.4.2	Haptic Models and Stimuli .....	34
3.1.4.3	Pilot Study: Selection of Coactive Feedback Vocabulary .....	36
3.1.4.4	Experimental Design .....	37
3.1.4.5	Trial Generation Procedure .....	40
3.1.4.6	Participants .....	41
4	RESULTS .....	43
4.1	Learning Performance and Convergence .....	43
4.1.1	Model's Performance .....	43
4.1.1.1	K Condition .....	43

4.1.1.2	KB Condition .....	44
4.1.1.3	KBD Condition .....	44
4.1.2	Effect of Coactive Feedback on Learning Efficiency .....	45
4.1.2.1	K Condition .....	45
4.1.2.2	KB Condition .....	46
4.1.2.3	KBD Condition .....	46
4.2	Participant Behavior Analysis .....	47
4.2.1	Sensitivity to Torque Cues .....	47
4.2.2	Physical Determinants of Coactive Feedback .....	48
4.2.2.1	Influence of Stiffness, Damping, and Distance .....	48
4.2.2.2	Torque Distance Effect .....	50
4.2.3	Final Estimate Distributions .....	50
4.2.3.1	K Condition .....	51
4.2.3.2	KB Condition .....	52
4.2.3.3	KBD Condition .....	52
4.2.3.4	Summary of Final Estimates .....	53
5	DISCUSSION .....	55
5.1	Validation of the Computational Framework .....	55
5.2	Perceptual Sensitivity and Dimensionality .....	55
5.2.1	The Stiffness (K) Condition .....	55
5.2.2	The Stiffness-Damping (KB) Condition .....	56
5.2.3	The Stiffness-Damping-Distance (KBD) Condition .....	56

5.3	Sensory Basis of Decision Making .....	57
5.4	The Impact of Coactive Feedback .....	57
5.4.1	Convergence Rate in Low Dimensions (K) .....	58
5.4.2	Efficacy in Higher Dimensions and Feedback Ambiguity .....	58
6	CONCLUSION AND FUTURE WORK .....	61
6.1	Limitations .....	61
6.1.1	Artificiality of the Haptic Stimuli .....	61
6.1.2	Impact of Device's Inner Friction .....	61
6.1.3	Participant Expertise .....	61
6.1.4	Reference Centering and Initialization Bias .....	62
6.2	Future Work .....	62
6.2.1	Ensuring Perceptual Differences .....	62
6.2.2	Beyond Binary Feedback .....	63
6.2.3	Gamification and Attentional Maintenance .....	63
6.2.4	Evaluation of Higher-Order Query Types .....	63
6.3	Conclusion .....	63
6.4	Summary of Contributions .....	64
APPENDICES APPENDICES		
A	HIERARCHICAL MODELS FOR COACTIVE FEEDBACK .....	73
A.1	Model 1: Determinants of Feedback (Physical Parameters) .....	73
A.2	Model 2: Specificity of Feedback .....	74
B	PILOT STUDY DETAILS .....	77

## LIST OF TABLES

Table 1	Parameter Bounds and Ranges .....	35
Table 2	Warm-up Trials (Reference = 150 mNm) .....	36
Table 3	Counterbalanced Permutations of Experimental Blocks .....	37
Table 4	Fixed Trials for K Condition (Reference $K = 200$ ) .....	37
Table 5	Fixed Trials for KB Condition (Reference $[K, B] = [175, 15]$ ).....	38
Table 6	Fixed Trials for KBD Condition (Reference $[K, B, D] = [150, 20, 0.2]$ ).....	38
Table 7	Parameter Adjustments based on Coactive Feedback .....	38
Table 8	Posterior estimates for the influence of torque distance on choice probability.	48
Table 9	Posterior estimates for the influence of haptic parameters on “Hard” feedback.	49
Table 10	Posterior estimates for feedback specificity: influence of selected vs. rejected option. ....	51
Table 11	Summary statistics of final parameter estimates at Trial 13 ( $n = 72$ ).....	53
Table 12	Frequency of Feeling Words Reported in Pilot Study (N=16) .....	77
Table 13	Pilot Study Parameter Configurations - Group A .....	78
Table 14	Pilot Study Parameter Configurations - Group B .....	78

## LIST OF FIGURES

Figure 1	Illustration of Bayesian Inference. ....	16
Figure 2	The custom 1-DOF haptic device used in the experiment. ....	34
Figure 3	User Interfaces used in the experiment. ....	35
Figure 4	Overview of the Experimental Procedure. ....	39
Figure 5	Detailed Trial Generation Loop (Trials 4-13). ....	41
Figure 6	Demographic information of the participants. ....	42
Figure 7	Evolution of stiffness and damping estimates in the K condition ....	43
Figure 8	Evolution of stiffness and damping estimates in the KB condition ....	44
Figure 9	Evolution of stiffness and damping estimates in the KBD condition ....	45
Figure 10	Stiffness convergence: Coactive vs. Non-Coactive ....	46
Figure 11	Stiffness and damping convergence: Coactive vs. Non-Coactive ....	47
Figure 12	Stiffness, damping and distance convergence: Coactive vs. Non-Coactive	48
Figure 13	Selection probability as a function of relative torque distances ....	49
Figure 14	Influence of haptic parameters on coactive feedback probability ....	50
Figure 15	Distribution of final stiffness estimates for K condition ....	51
Figure 16	Distribution of final stiffness estimates for KB condition ....	52
Figure 17	Distribution of final stiffness estimates for KBD condition ....	53

## LIST OF ABBREVIATIONS

B	Damping (mNm·s/rad)
BO	Bayesian Optimization
CDF	Cumulative Distribution Function
D	Distance / Detent Spacing (rad)
DC	Direct Current
EI	Expected Improvement
GP	Gaussian Process
HDI	Highest Density Interval
HiL	Human-in-the-Loop
HMC	Hamiltonian Monte Carlo
JND	Just-Noticeable Difference
K	Stiffness (mNm/rad)
LED	Light Emitting Diode
MCMC	Markov Chain Monte Carlo
NUTS	No-U-Turn Sampler
PDF	Probability Density Function
RBF	Radial Basis Function
SD	Standard Deviation
UI	User Interface
USB	Universal Serial Bus
1-DOF	One Degree of Freedom

# CHAPTER 1

## INTRODUCTION

Haptic interfaces provide a connection between digital systems and physical sensations. They are essential in fields like teleoperated robotics (where human operator controls a robotic system from a distance), medical simulation, and virtual reality. In these applications, the quality of the haptic feedback directly affects how well a user can perform a task. However, measuring this quality is challenging. Unlike screens, which use standard metrics like resolution to define clarity, haptic perception is subjective. Two people may perceive the same physical stiffness or damping differently due to their own sensitivity and hand mechanics.

This subjectivity leads to a significant design challenge: the “tuning problem.” In order to make a haptic device feel like a specific material, such as a rubber block, engineers must define specific physical parameters. These typically include stiffness ( $K$ ), damping ( $B$ ), and distance ( $D$ ). Traditionally, finding the best parameters is a manual process. It usually requires a domain expert to adjust the values repeatedly based on verbal feedback from the user. This trial-and-error approach is time-consuming and inefficient because it relies heavily on the expert’s presence and intuition (Stecklina et al., 2020).

### 1.1 Problem Statement

This thesis investigates the challenge of mapping subjective human preferences for a target sensation to precise physical control parameters. The objective is to replace manual, trial-and-error tuning with an intelligent algorithm that learns directly from user input. However, this approach faces several difficulties regarding both computation and human perception.

Firstly, user preferences are internal and subjective. Users cannot quantify their desired sensation in numerical terms, for example, by specifying a stiffness of  $150\text{ mNm/rad}$ . However, they can express relative judgments, such as preferring one sensation over another. Capturing these relative judgments requires a model capable of inferring continuous values from binary choices.

Secondly, data efficiency is critical because human attention is a limited resource. Since users can only perform a small number of comparisons before fatigue or habituation affects their sensitivity, the learning algorithm must converge to the optimal parameters within a minimal number of trials (Settles, 2009).

Finally, the problem becomes more difficult when multiple parameters are involved. As the complexity of the physical model increases to include properties like damping and distance (dead zones), the search space grows exponentially. Moreover, human perception tends to integrate these distinct physical parameters into a single holistic sensation, such as “Hardness” (Jones and Hunter, 1990). This integration makes it mathematically difficult to isolate and identify the individual variables based only on subjective feedback.

## 1.2 Proposed Approach

To address the challenges of subjectivity and data efficiency, this thesis proposes a *Bayesian Active Learning* framework. Unlike manual tuning, where a user or expert randomly explores the parameter space, our approach systematically guides the search process to maximize information gain with every interaction.

The core of the proposed solution is to shift the interaction model from “direct tuning” to “preference comparison.” Instead of asking a user to input a numerical value for stiffness, the system presents two distinct haptic renderings (e.g., Setting A vs. Setting B) and asks the user to select the one that feels closer to their desired target. This pairwise comparison method reduces the cognitive load on the user, as humans are significantly more reliable at making relative judgments than absolute estimations (Miller, 1956).

Furthermore, a hybrid feedback mechanism is introduced. While pairwise comparisons provide the baseline data, they are supplemented by *coactive feedback*. This allows the user to provide directional guidance (e.g., “Make it harder”) to the winning option. This additional signal is intended to prevent the algorithm from getting stuck in local optima and to accelerate the search process by explicitly pruning irrelevant regions of the parameter space.

We model the user’s latent preference function using Gaussian Processes (GPs). GPs provide a probabilistic foundation that not only estimates the user’s preferred parameters but also quantifies the system’s uncertainty about those estimates (Rasmussen and Williams, 2006). This uncertainty is crucial for the “active” component of the system.

Using an acquisition function, the system evaluates the GP posterior to autonomously determine the next query. The goal of this active learning process is to select the candidate point that offers the maximum expected improvement over the current best setting. This creates a closed-loop system:

1. The system selects and renders a pair of haptic settings.
2. The user selects the preferred setting.
3. The user optionally provides a directional improvement, such as “Make it harder” (Coactive Feedback).
4. The system updates its Gaussian Process posterior based on this hybrid feedback.
5. The cycle repeats until the system converges to the optimal parameters.

By combining the flexibility of Gaussian Processes with the efficiency of Active Learning and Coactive feedback, this method aims to converge to the user’s subjective ideal parameters with a minimal number of trials, making personalized haptic customization feasible for non-experts.

### 1.3 Research Questions

This research aims to validate the feasibility of this computational framework through human-subject experiments. The specific research questions are as follows:

1. **Feasibility of Pairwise Learning:** Can a Gaussian Process model accurately estimate hidden haptic parameters (Stiffness, Damping, Distance) only from subjective pairwise comparisons?
2. **Impact of Dimensionality:** How is the convergence rate affected by the complexity of the haptic model as it changes from a single parameter (Stiffness) to multi-parameter dynamic environments (Stiffness + Damping + Distance)?
3. **Role of Coactive Feedback:** Does the consideration of directional feedback (“Harder/Softer”) improve the convergence rate when compared to standard pairwise comparisons?

### 1.4 Contributions of the Study

The main contribution of this thesis is the development of a Bayesian optimization framework specifically adapted for haptic parameter tuning, which integrates preference learning with physical control variables. Through this framework, the study provides an empirical analysis of how human perceptual limits, such as Just Noticeable Differences, interact with algorithmic optimization in high-dimensional search spaces. Furthermore, it evaluates a coactive feedback mechanism that enables users to guide the optimization process, offering key insights into how semantic labels like “Hard” map to physical parameters in complex mechanical environments.

### 1.5 Organization of the Thesis

The remainder of this thesis is organized as follows:

**Chapter 2** establishes the theoretical foundations of the study. It provides a comprehensive review of *Preference Learning*, including decision-making theories and ranking methods. It then details the mathematical principles of *Bayesian Inference*, focusing on Gaussian Processes, Active Learning, and the integration of coactive feedback. Finally, it reviews the relevant literature on *Haptic Perception*, specifically regarding stiffness, damping, and spatial parameters.

**Chapter 3** details the methodology, divided into two primary components. First, the *Computational Method* is presented, defining the Gaussian Process priors, likelihood functions, and

the Penalized Expected Improvement acquisition function. Second, the *Experimental Method* describes the haptic hardware, the generation of stimuli, the experimental design, and the trial procedure used to validate the framework.

**Chapter 4** presents the empirical results of the user study. The analysis is split into two sections: *Learning Performance*, which evaluates the convergence of the model across the K, KB, and KBD conditions, and *Participant Behavior Analysis*, which investigates the physical determinants of user choices, sensitivity to torque cues, and the specificity of coactive feedback.

**Chapter 5** provides a critical discussion of the findings, linking algorithmic performance to human perceptual sensitivity. It outlines the limitations of the study regarding ecological validity and device transparency, proposes directions for future work and finally, concludes the thesis.

## CHAPTER 2

### LITERATURE REVIEW

This chapter establishes the theoretical foundation for modeling human preferences in haptic parameter optimization, bridging the gap between the cognitive mechanisms of human decision-making and the computational frameworks required to learn them. To effectively model how an individual perceives and prefers specific haptic stimuli, one must first understand the underlying psychological principles of choice, the perceptual limitations of human haptic sensing, and the mathematical tools available to approximate subjective utility.

The challenge of learning user preferences for haptic feedback lies at the intersection of several research areas. From psychology and behavioral economics, we draw insights into how humans make decisions under uncertainty and why their choices often deviate from rational utility maximization. From psychophysics, we understand the perceptual thresholds that constrain what differences humans can reliably detect. From machine learning and Bayesian statistics, we obtain the computational tools necessary to infer latent preference functions from noisy, sparse human feedback. The synthesis of these perspectives enables the development of sample-efficient algorithms that can learn complex, subjective preferences with minimal user burden.

The chapter is organized as follows: Section 2.1 provides a broad overview of preference learning, beginning with theories of human decision-making that contrast traditional utility theories with modern behavioral perspectives. Later, we examine the taxonomy of preference learning problems, the methods used to solve them, and the specific data structures used to capture preferences. The section concludes with the psychophysics of pairwise comparisons, establishing why comparative judgments provide more reliable data than absolute ratings.

Section 2.2 reviews the statistical frameworks necessary for inference, detailing the shift from Frequentist to Bayesian approaches and motivating the adoption of a probabilistic framework. Consecutively, Gaussian Processes (GP) as a flexible, non-parametric model for representing utility functions introduced, followed by Active Learning strategies that enable efficient data collection in resource-constrained environments. Bayesian Optimization and its acquisition functions are discussed as the mechanism for directing the search toward the user's preferred settings. Moreover, we examine Coactive Feedback as an alternative interaction paradigm that provides richer information per query.

Finally, Section 2.3 reviews the literature on Haptic Perception, focusing on the psychophysics of stiffness, damping, and distance perception. This section establishes the perceptual con-

straints that any haptic preference learning system must consider, and motivates the use of Bayesian methods that can handle the inherent noise in human haptic judgments.

## 2.1 Preference Learning

Preference learning is the construction of a model using empirical data to predict the future preferences of individuals or organizations, or to uncover the less obvious dynamics of their preferences (Fürnkranz and Hüllermeier, 2010). As a subfield of machine learning, preference learning is a powerful tool for modeling and predicting subjective preferences and decisions. It is also used to study group decision-making processes and preference dynamics in different contexts.

### 2.1.1 Theories of Human Decision-Making

Decision-making theories have offered many models of how people make decisions. Expected Utility Theory (EUT) suggests that individuals make decisions by maximizing a utility function. According to this theory, the utility function represents the satisfaction or benefit obtained from each choice, and it is assumed that individuals choose the option with the highest utility (von Neumann and Morgenstern, 1944).

However, research in psychology and behavioral economics has shown that the human decision-making process often deviates from the rational models suggested by utility theory. Kahneman and Tversky (1979) introduced prospect theory, which highlights the systematic biases and heuristics people use under uncertainty. This work demonstrates that decision-making is not always consistent with maximizing expected utility.

Another approach is Simon (1955)'s theory of "Bounded Rationality". Simon argued that humans have limited access to information and limited computational ability to predict outcomes, especially in complex decisions. Instead of optimizing a global utility function, humans search for a solution that meets a minimum acceptability threshold.

Furthermore, Slovic (1995) stated that people's preferences are generally constructed during the elicitation process and are highly context-dependent. This contrasts with traditional views that preferences are fixed and directly accessible. Tversky and Simonson (1993) presented another context-dependent choice model. In their argument, decision-makers evaluate options based on *background context* (the importance of an attribute based on past experiences) and *local context* (the specific set of options currently available). In other words, the value of an option is not stable but is adjusted by what has already been seen (*background*) and what is currently available (*local*).

Finally, Tversky (1969) demonstrated that human preferences generally violate the property of transitivity and can exhibit cyclic preferences (e.g.,  $A \succ B$ ,  $B \succ C$ , and  $C \succ A$ ). This pattern directly contradicts the main assumption of EUT.

Despite its limitations, EUT is seen as the most important and fundamental baseline, since it gives a mathematical expression to human logic and decisions. Subsequent theories were constructed as modifications to account for these cognitive realities.

To sum up, preference learning focuses on modeling and predicting subjective preferences by working with data derived from choices of individuals. This deviates from the rationality assumption of traditional Utility Theory, making it possible to understand individuals' heuristic and biased decisions. By benefiting from this data, preference learning develops systems that represent these decisions better. Therefore, it offers a significant tool for capturing the complexities of real-world decision-making processes based on individuals' past decisions. It is useful in modeling subjective and complex preferences, as well as being able to handle noisy, incomplete, or contradictory data.

## 2.1.2 Preference Learning Types

Preference learning problems can be organized into three main categories based on what is being ranked and how the data is structured: label ranking, instance ranking, and object ranking (Fürnkranz and Hüllermeier, 2010). Although these problem types are different in their formal setup, they share a common goal, which is learning a function that can predict orderings over items based on observed preference information. Distinctions among these types help us to clarify which computational approach is appropriate for a given application.

### 2.1.2.1 Label Ranking

In label ranking, the objective is to learn a mapping from instances to orderings over a fixed set of labels (Fürnkranz and Hüllermeier, 2010). Let  $Y = \{y_1, \dots, y_n\}$  be a set of labels for some  $n \in \mathbb{N}$ , and let  $y_i \succ y_j$  denote the relationship where  $y_i$  is preferred over  $y_j$ . Label ranking is the task of learning a function that maps each element of a given set  $X$  to a preference order of labels in the form  $y_{i_1} \succ \dots \succ y_{i_n}$  for  $(i_1, i_2, \dots, i_n) \in \text{Sym}(n)$  (Fürnkranz and Hüllermeier, 2010). Here,  $y_i \succ y_j$  indicates that the label  $y_i$  is preferred over  $y_j$ . In this context, the types of observed data may appear in different forms depending on how preferences are expressed.

The first type is called pairwise comparison. In this case, for  $x \in X$  and any pair  $i, j \in \{1, \dots, n\}$ , it is observed that  $y_i \succ y_j$ .

The second type is partial set comparison. In this case, for  $x \in X$  and any  $m \in \{1, \dots, n\}$ , a preference ranking is specified over a subset  $\{y_{i_1}, \dots, y_{i_m}\}$  and is expressed as  $y_{i_1} \succ \dots \succ y_{i_m}$ .

The third and final type is full set comparison. In this case, for  $x \in X$ , all labels are expressed as a preference order in the form  $y_{i_1} \succ \dots \succ y_{i_n}$ .

Label ranking has many real-world applications where items must be associated with multiple categories whose relative importance varies for context. As an example, a research paper can be associated with multiple topic labels, and label ranking can be used to rank the topic headings to which the research papers are relevant (Vembu and Gärtner, 2011). This is a valuable tool in fields such as information retrieval and document classification. Similarly,

label ranking can be applied to more technical problems such as the evaluation and selection of machine learning algorithms. For instance, the question of which algorithm to use for a given dataset is an important subject of debate, since many algorithms have been developed. A label ranking model can suggest which machine learning algorithms to use. In this context, studies have been conducted on the usability of label ranking techniques as a structure for recommending algorithms (Aiguzhinov et al., 2010). In this application, “labels” represent machine learning algorithms, and “ranking” represents the order of algorithms according to their relative performance for a specific dataset. One of the significant advantages of this approach is that, instead of suggesting a single best algorithm, it provides a ranked list of algorithms. Therefore, this ranked output provides practitioners with actionable alternatives when the top choice is infeasible.

### 2.1.2.2 Instance Ranking

Instance ranking differs from label ranking in which the goal is to order the instances themselves rather than to assign label orderings to them. Let  $Y = \{y_1, y_2, \dots, y_n\}$  be a set of labels for some  $n \in \mathbb{N}$ , and let  $X$  be a set on which a preference ranking will be revealed. Let a natural preference ranking  $y_{i_1} \succ \dots \succ y_{i_n}$  be defined on the label set for any  $(i_1, i_2, \dots, i_n) \in \text{Sim}(n)$ . The training data consists of mapping each element of any subset of  $X$  to one from the label set. Learning a function that puts any subset of  $X$  into a preference order by learning preferences from this dataset is called instance ranking. This ranking is generally obtained by assigning a score to the subset elements and then sorting these scores. This method is widely used due to the flexibility it provides and ease of applicability. Furthermore, this approach can work effectively on both full rankings and partial rankings.

Web search provides a canonical example of instance ranking. Search engines like Google provide an approach that solves the instance ranking problem when ranking the most appropriate websites for a user’s query. In this process, each website is evaluated as an instance, and these are ranked according to their relevance to the user’s query. Google performed this ranking in the past using the PageRank algorithm (Page et al., 1999). PageRank is a ranking algorithm that represents links between websites with a graph model and an adjacency matrix. It does not use a response variable. In 2013, PageRank became a part of the Hummingbird algorithm, the underlying task still remains as an instance ranking problem (Werner, 2022).

Instance ranking also applies to image retrieval. In the study of Hu et al. (2008), a multiple instance ranking method was developed and all images were ranked according to their relevance based on pairwise image comparisons. It was created to determine which image is more relevant to a keyword phrase provided by the user. This approach offers a highly effective method for producing user-oriented results and improving retrieval performance, especially in visual databases.

### 2.1.2.3 Object Ranking

Learning a function that puts any subset of an object set  $Z$  into a preference order is called object ranking. Unlike instance ranking where each instance has an associated context or

feature vector that predicts relevance to a query, object ranking focuses on objects themselves without assuming a separate input space. This function is generally learned from pairwise comparison data and is obtained by assigning scores to these objects and ranking the objects according to these scores. For example, Recommender Systems are an important application example based on object ranking (Resnick and Varian, 1997). These systems offer personalized recommendations by analyzing data regarding the user’s previous choices and preferences. The ranking constructed from implicit or explicit preference data aims to facilitate decision-making processes by ranking objects suitable for users’ needs.

Similar to instance ranking, object ranking can also be used in the optimization of search engines or information retrieval systems. For example, ranking quality can be optimized using data on which links users prefer to click on the search results page. This click-through data, which is a record of search results that users actually select, can serve as a source of pairwise preferences. Joachims (2002) showed that this approach effectively improved the ranking quality of a meta-search engine and outperformed Google’s performance.

### **2.1.3 Preference Learning Methods**

Preference learning methods can be grouped into three broad categories based on how preferences are presented and learned: learning with utility functions, learning with preference relations, and model-based preference learning (Fürnkranz and Hüllermeier, 2010).

#### **2.1.3.1 Learning with Utility Functions**

Utility functions were first developed by von Neumann and Morgenstern (1944) within the scope of Game Theory. They serve as a tool that assigns a real number to each element of a set for which a preference relationship is to be established, thereby specifying the preference ranking among these elements. These functions can be applied in different ways depending on whether the input set is numerical or categorical. For numerical inputs, additional constraints such as monotonicity can be applied (Dembczyński et al., 2011).

When learning preferences via utility functions, training data can take many forms. For example, from pairwise comparison data, one can infer that one object has higher utility compared to another and learn a utility function based on this data Chin et al. (2018). Alternatively, training data may consist of complete rankings over sets of objects. Pfannschmidt et al. (2018) used deep neural networks to learn context-dependent utility representations from full ranking data, capturing how the value of an object may shift depending on the available alternatives.

The training data in instance ranking consists of examples where the true utility scores are already known. If the utility scores are numerical values, this problem can be treated as a regression problem in machine learning. If the scores are not numerical but contain a specific order, this problem can be treated as ordinal classification. These problems can be solved with existing machine learning algorithms. However, a significant difference from machine learning problems is that in instance ranking, the goal is not to maximize accuracy, but to

correctly predict the ranking of each instance. Therefore, traditional algorithms need to be adapted accordingly to optimize ranking performance (Fürnkranz and Hüllermeier, 2010).

As an application example of learning instance ranking via utility functions, Herbrich et al. (2000) presented a large margin algorithm that maps objects to scalar utility values and determines their rankings by classifying pairs of objects. Experimental results in an information retrieval task reveal that this approach performs better than simple methods such as Support Vector Classification and Support Vector Regression, especially in cases with more than two ranks.

Unlike the instance ranking problem, in object and label ranking problems, training data generally contain constraints such that one object/label should have higher utility than another, rather than exact utility values. In this context, one of the first methods was developed by Tesauro (1988). Tesauro proposed a symmetric neural network that learns from pairwise comparison data. The network learns the preference relationship by generating output for each of the two states, and these outputs are compared. Later, real-valued scores for individual objects are produced using a single network instead of symmetric components. This method both improves ranking performance and offers a more efficient approach in the field of preference learning.

In label ranking problems, the constraint classification method has been developed as a technique (Har-Peled et al., 2002). Essentially, this involves transforming the original problem into a simpler binary classification problem, which is then solved using standard machine learning techniques.

Another method for label ranking is log-linear models (Dekel et al., 2003). This method uses a boosting algorithm to iteratively learn the ranking function. Boosting is an ensemble learning method that combines weak learners (simple models) to create a strong learner. In this case, weak learners iteratively improve the ranking function by focusing on errors from previous iterations. The algorithm minimizes a generalized ranking error in each iteration. This ensures that the learned ranking function performs better over time. Theoretical guarantees are provided for the method by proving a lower bound for the progress made in each boosting iteration.

### **2.1.3.2 Learning with Preference Relations**

Learning with preference relations focuses on learning binary preference relations instead of learning a utility function to rank objects or labels. In this framework, pairwise comparisons of the form  $a \succ b$  are observed to construct a relation that is consistent with the observed data. It offers a simpler approach than the utility function learning method because training data can be used directly and it is closer to real-world problems. However, creating a full ranking from binary preference relations can be complex. In such cases, the goal may be to minimize conflicts between the ranking and binary preferences.

Techniques such as Borda count or algorithms based on weighted degrees offer effective approaches to convert binary preferences into rankings. Coppersmith et al. (2010) made a significant contribution to this field by proposing an algorithm for ranking in tournaments.

The proposed algorithm uses weighted degrees to ensure consistency and offers a proven approximation guarantee. Ukkonen et al. (2009) proposed bucket orders, a type of partial ranking where items are grouped into buckets, with ties within buckets and a total order between them. In this study, they minimized inconsistencies using the  $L_1$  norm and proposed a randomized Quicksort-like algorithm to predict the optimal bucket order that fits pairwise preference probabilities. Despite the NP-hard nature of the problem, this algorithm offers an efficient solution with a constant-factor approximation guarantee.

### 2.1.3.3 Model-Based Preference Learning

In model-based preference learning, a model assumption is made regarding how preferences are given. Among these approaches, probabilistic models offer a robust framework for capturing uncertainty and variability in preferences by representing preferences with probability distributions. These models provide a more detailed understanding of preference compared to deterministic models, with their ability to handle noisy, incomplete, or inconsistent training data. Probabilistic models can predict the probabilities of different preference rankings, combine past information with new information, and dynamically update beliefs in light of this information. These features ensure they are compatible with real-world scenarios where preferences are context-dependent and dynamic. Thus, this approach not only provides more accurate preference predictions but also stands out as a critical tool for real-world decision-making problems by increasing the models' capacity to handle incomplete and inconsistent data (Liu et al., 2019). Examples of probabilistic models in preference learning include the Bradley-Terry model, the Plackett-Luce model, and the Mallows model.

Bradley and Terry (1952) presented a probabilistic model that learns from pairwise comparison data. When two objects  $i$  and  $j$  are selected from the set  $X$  where preference learning will be performed, the probability of  $i$  being preferred over  $j$  is defined as follows:

$$Pr(i \succ j) = \frac{\beta_i}{\beta_i + \beta_j} \quad (1)$$

Here,  $\beta_i$  and  $\beta_j$  are positive parameters representing the preference score (or perceived value) of objects  $i$  and  $j$ , respectively. The Bradley-Terry model offers an effective framework for modeling preference relations probabilistically. For example, a study by Sun et al. (2024) presented a comprehensive analysis on the use of the Bradley-Terry model in reward modeling for large language models where it is used to learn human preferences over model outputs from pairwise annotations.

The Plackett-Luce Model extends the binary rankings of the Bradley-Terry model to full rankings and offers a framework for learning the preference relationship between elements of a set (Luce, 1959; Plackett, 1975). Let  $A = a_1 \succ a_2 \succ \dots \succ a_n$  be a ranking for any  $n \in \mathbb{N}^+$  whose probability we want to learn, and let  $\mu_1, \dots, \mu_n$  be positive scores representing the preference score of the objects in the ranking. Then, the probability of ranking  $A$  becomes:

$$Pr(A) = \prod_{i=1}^n \frac{\mu_{a_i}}{\sum_{j=i}^n \mu_{a_j}} \quad (2)$$

(Liu et al., 2019). Schäfer and Hüllermeier (2018) extended the Plackett-Luce model and adapted it to the dyad ranking problem. This method allows both contexts and alternatives to

be represented by features and aims to increase ranking performance. Two different approaches were presented in this study: BilinPL using the Kronecker product, and PLNet learning high-order non-linear features with artificial neural networks. Their experimental studies show that feature information of alternatives provides significant contributions in ranking problems.

The Mallows model calculates the probability of observing a ranking, just like the Plackett-Luce model. However, unlike the Plackett-Luce model, the Mallows model assumes a central ranking and a dispersion parameter around this ranking (Mallows, 1957). The central ranking represents the “ideal” ranking where rankings are clustered on average. The dispersion parameter controls how much the observed rankings can deviate from the central ranking. The probability of a ranking depends on the distance of the ranking to the central ranking; the further it is from the central ranking, the lower the probability of that ranking occurring. Vitelli et al. (2018) improved the Mallows model and proposed the Bayesian Mallows Model. In this model, the central ranking and dispersion parameter are treated as random variables and prior probabilities are assigned to these variables. In this way, the central ranking and dispersion parameter can also be learned from observed data. In this model, the probability of any observed preference ranking  $\sigma$  occurring, given the central ranking  $\rho_0$ , distance metric  $d$ , and dispersion parameter  $\phi$ , is:

$$Pr(\sigma|\rho_0, \phi) \propto e^{-\phi d(\sigma, \rho_0)} \quad (3)$$

This model can be used in group decision-making problems or problems with missing data. The Bayesian Mallows model makes it possible to model uncertainty and variability in ranking data more flexibly. Sørensen et al. (2024) proposed an SMC<sup>2</sup> algorithm capable of working with sequential data using the Bayesian Mallows model for ranking and preference learning. This method makes working with partial rankings or pairwise preferences efficient, especially in dynamic environments, by sequentially updating posterior distributions as new data arrives.

#### 2.1.4 Preference Data Types

Directly asking decision-makers to define their preference (or utility) functions is often ineffective, as individuals rarely possess explicit knowledge of these internal functions (Payne et al., 1993; Slovic, 1995). They may know which criteria are important for them, or they can even indicate an importance (weight) on the criteria. Instead of this approach, these functions must be elicited. The challenge of modeling utility as a combination of criteria and weights lies at the core of Multi-Criteria Decision Making/Aiding (MCDA) (Keeney and Raiffa, 1976). Although early approaches relied on the direct elicitation of parameters (e.g., asking users to explicitly state weights), this process is often cognitively demanding and unreliable. Consequently, modern MCDA methodologies assume that a user’s latent utility function is better revealed through their choices. MCDA and preference learning share the major goal of constructing practically useful decision models to support humans in tasks like choosing the best alternative, classifying, ranking, or automating decision-making (Hüllermeier and Słowiński, 2024). MCDA has roots in decision sciences and operations research, while PL originated in AI and machine learning (Hüllermeier and Słowiński, 2024).

In order to learn the latent utility function, several type of data can serve as the basis for preference function elicitation. Each of these impose different cognitive demands and provide

different amounts of information per response. These are pairwise comparisons, selecting the best from a subset of the total alternatives, ranking a subset of total alternatives, giving a score for a subset of alternatives or giving a label to the alternatives. Another approach in machine learning is Inverse Reinforcement Learning (IRL) (Ng and Russell, 2000). In IRL, instead of answering questions, the user performs the task. The system assumes the user is maximizing their internal utility and infers the reward function that explains the behavior. A distinct approach involves critiquing, where users provide relative feedback on specific attributes of an alternative (e.g., “this is too high” or “too soft”), often referred to as learning from relative attributes.

Rating scales are widely used because they are easy to administer, although decades of psychophysical research have documented their limitations. Miller (1956) investigated “Absolute Judgment” experiments which test how accurately people can assign numbers to the magnitudes of various aspects of a stimulus. In these experiments, the goal was to increase the amount of input information (stimuli) and also measuring the amount of transmitted information, which is stimulus-response correlation. Miller (1956) demonstrated that the human capacity for Absolute Judgment is limited to roughly 2.5 bits of information, which is equivalent to about 7 distinct categories. When users were asked to map internal sensation onto a fine-grained scale (e.g., 0-100), they exceed this capacity, resulting in high response variability.

A deeper problem is that individuals’ internal scales are unstable and prone to drift during experimentation. According to Parducci (1965), a user’s internal scale is not fixed but stretches and shrinks depending on the range and frequency of the stimuli they have experienced. A stimulus rated “5” in one context might be rated “7” in another, making absolute ratings unreliable for training stable models. Laming (2004) deepened this argument and stated that humans have no capacity for true absolute judgment. What looks like an absolute rating is in fact a comparison against a decaying memory of previous stimuli. Since users must rely on this imperfect memory, absolute ratings become increasingly unreliable as the number of stimuli grows, making them poorly suited for training data. While ordinal data avoids the pitfalls of absolute rating, the method of elicitation remains critical. Alternative methods such as Ranking (sorting a full list) or Selection (picking the best from a set of  $k > 2$ ) theoretically offer higher information density per query than pairwise comparisons ( $\log_2 k!$  and  $\log_2 k$  bits, respectively). However, these methods impose a heavy cognitive load in the context of haptics. Unlike visual tasks, where users can scan multiple items simultaneously, haptic perception is sequential. To rank five different stiffness settings, a user must hold the sensation of the first setting in working memory while exploring the subsequent four. Due to the rapid decay of haptic short-term memory, this often leads to confusion and inconsistent ordering. Pairwise comparisons, therefore, represent the optimal trade-off: they minimize working memory load by requiring the user to compare only the immediate stimulus against the one directly preceding it.

In contrast to ratings, pairwise comparisons require only a relative judgment. Basically, which of two alternatives is preferred. Since no numerical scales involved, comparisons are immune to the drift and anchoring effects that affect absolute ratings. Aggarwal and Tehrani (2019) proposed a method for modeling human decision behavior from pairwise preferences based on the Attitudinal Choquet Integral operator. This operator enhances the conventional Choquet integral by explicitly learning the decision maker’s attitudinal character (tolerance/perfectionism) alongside criteria interaction.

### 2.1.5 Psychophysics of Pairwise Comparisons

Pairwise comparisons derive their reliability from fundamental properties of human perception, which is documented in classical psychophysics. This section reviews theoretical foundations that explain why relative judgments are more stable than absolute ratings.

Just Noticeable Difference (JND) is the smallest change that can produce a difference in the sensation of two stimuli. In other words, it is a difference threshold at which a person can detect a difference in sensory experience. Weber (1834) formulated that the JND is not a constant value but a constant proportion of the original stimuli intensity, known as Weber's Law:

$$\frac{\Delta I}{I} = k \quad (4)$$

where  $\Delta I$  is the difference threshold (JND),  $I$  is the original intensity of the stimulus, and  $k$  is the Weber fraction.

Fechner (1860) expanded this into Fechner's Law, proposing a logarithmic relationship between physical stimulus intensity and internal sensation. He integrated Weber's equation to derive:

$$S = c \cdot \ln(I) \quad (5)$$

where  $S$  is the magnitude of the sensation,  $I$  is the physical intensity, and  $c$  is a constant. This law establishes that as physical intensity increases geometrically, the internal perception of that intensity increases only arithmetically.

Thurstone (1927) developed the mathematical framework that connects Weber's Law to Pairwise Comparisons. According to Thurstone, individual preferences in comparisons are not fixed, rather these judgments are stochastic. In other words, at different times, preferences regarding the same question may change. This is called "The Law of Comparative Judgment". This law treats the internal sensation of a stimulus not as a fixed point but as a Gaussian random variable  $S_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ .

When a user compares two stimuli,  $A$  and  $B$ , they do not assess the absolute value of  $A$  and  $B$  separately. Instead, they perceive the difference in the momentary realizations of these random variables:  $d_{AB} = S_A - S_B$ . The user prefers  $A$  if  $d_{AB} > 0$ .

This model explicitly accounts for the "noise" in human perception ( $\sigma^2$ ). Even if the user's absolute perception drifts (e.g., the means  $\mu_A$  and  $\mu_B$  both shift due to fatigue), the relative difference often remains stable. The law establishes the scale distance between two items based on how consistently one is judged "better" or "more" than the other, therefore accounting for the inherent variability (dispersion) in the human judgment process.

Kingsley and Brown (2010) observed that this consistency improves with experience. Even if early choices may reflect higher "preference uncertainty", respondents become more consistent and better at discriminating among items as they gain experience.

Despite the fact that pairwise comparisons are more accurate and yield less noise than other data collection methods, they suffer from low information density. For a set of  $N$  discrete alternatives, the total space of possible comparisons is  $\binom{N}{2}$ , which grows quadratically. Even

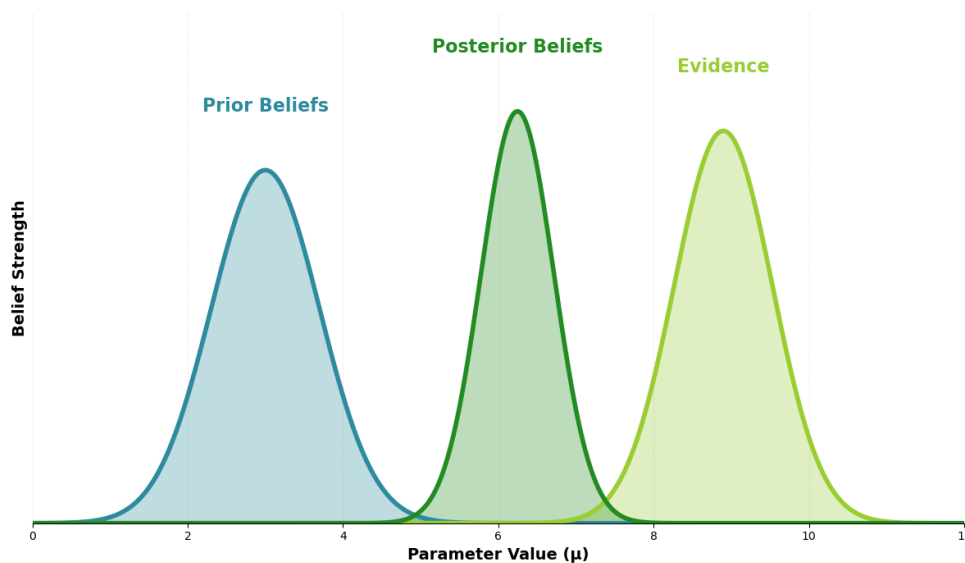
though it is not strictly necessary to query every pair to infer a ranking, pairwise queries provide only limited information per response (1 bit). In contrast, asking a user to select the “best from a subset of  $K$  alternatives” ( $K > 2$ ) yields more information per query (specifically,  $\log_2 K$  bits). However, this task imposes a significantly higher cognitive load.

In order to compensate for the information sparsity of pairwise comparisons, modern approaches integrate Thurstone’s probabilistic framework with machine learning. In addition, it is necessary to employ sample-efficient modeling techniques, such as GPs. Moreover, obtaining reliable preference estimates without exhausting the user requires strategic query selection. GPs are combined with Active Learning strategies to select the most informative queries. For instance, Chu and Ghahramani (2005) demonstrated that GPs can infer utility functions from noisy pairwise data, by leveraging smoothness assumptions encoded in the kernel. Sadigh et al. (2017) demonstrated that Active Learning can reduce the number of queries required to learn a user’s preference. Choosing pairs that maximize expected information gain reduces the number of comparisons required to identify a user’s preferred setting.

## 2.2 Bayesian Inference

Statistics has traditionally been dominated by the frequentist approach, particularly throughout the 20th century, famously formalized by Fisher (1925); Neyman and Pearson (1933). This framework defines probability strictly as the long-run frequency of an event occurring in a hypothetical, infinite series of trials (Venn, 1866; von Mises, 1957). John Venn was one of the first to define probability as the frequency of events in the long run. Central to the frequentist view is the assumption that population parameters are fixed, unknown constants. Consequently, prior knowledge or beliefs are excluded from the analysis. Inference relies merely on the likelihood of the observed data given a null hypothesis ( $H_0$ ). This often involves Null Hypothesis Significance Testing (NHST), where the objective is to determine if there is sufficient evidence to reject the null hypothesis, rather than to accept it. For instance, in a frequentist framework, a fair coin has a fixed parameter  $\theta = 0.5$ . Even if 6 heads in a row are observed, the frequentist assumes the ratio will converge to 0.54 over infinite trials.

Conversely, Bayesian statistics treat probability as a measure of subjective belief or uncertainty (de Finetti, 1937). Unlike Frequentists, the Bayesian approach treats parameters as random variables described by a probability distribution. Fundamentally, it uses prior knowledge (*Prior*) and updates this belief using observed data (*Likelihood*) to form a new belief (*Posterior*) (Gelman et al., 2013). This allows for the assignment of probabilities to unique and non-repeatable events like election results. Furthermore, frequentism cannot easily handle this type of problem. To revisit the coin example: if a Bayesian observes 6 heads in a row, they do not simply wait for the average to converge. Instead, they update their posterior distribution. If their prior allowed for the possibility of a biased coin, the observation of 6 heads would shift their belief, making the probability of a 7th head higher than 0.5, rather than lower. The process of updating prior beliefs is visually illustrated in Figure 1.



**Figure 1: Illustration of Bayesian Inference.** The prior belief (Teal) represents initial assumptions about the parameter. The evidence (Yellow-Green) represents the likelihood of the observed data. The posterior belief (Dark Green) is the updated distribution obtained by combining the prior and the evidence, formalized by the relationship  $p(\theta|y) \propto p(y|\theta)p(\theta)$ .

### 2.2.1 Advantages of Bayesian Inference over Frequentist Methods

The decision to adopt a Bayesian framework rather than a traditional frequentist one is motivated by the methodological limitations inherent in NHST, the noisy nature of human decision-making, and the necessity of learning from sparse data.

First, frequentist methods struggle to mathematically account for hypotheses that span a range of values rather than a single point. As Trafimow (2017) points out, calculating the probability of data given a hypothesis  $P(\text{Data}|H_0)$  is straightforward when the null hypothesis is a precise point, such as a coin being exactly fair ( $\theta = 0.5$ ). However, in many real-world contexts, researchers are interested in a range hypothesis (e.g., the bias is anywhere between 0 and 0.5). In these cases, it becomes mathematically problematic to calculate a single probability for the data without assuming a distribution across that range. To bypass this, the frequentist approach traditionally selects the single point within the range that maximizes the p-value to represent the entire hypothesis. (Trafimow, 2017) critiques this method as arbitrary and, in many cases, “blatantly wrong” because it fails to account for the full spectrum of possibilities. In contrast, the Bayesian approach resolves this naturally by defining a prior probability distribution and integrating across it, allowing for a coherent probability assessment that incorporates uncertainty over the entire range.

Furthermore, the practical utility of NHST has been questioned. Anderson et al. (2000) argue that testing a precise null hypothesis (e.g., “zero effect”) is uninformative in complex systems where effects are rarely exactly zero. Instead of rejecting a trivial null, they advocate for an Information-Theoretic approach that prioritizes model selection and parameter estimation, aligning closely with the predictive goals of Bayesian methods.

A second motivation for adopting Bayesian methods is their alignment with how humans actually process information. Research in cognitive science and computational neuroscience has converged on what is now called the "Bayesian brain" hypothesis: the idea that the nervous system combines prior expectations with incoming sensory data in a manner consistent with probabilistic inference. For instance, Körding and Wolpert (2004) provided compelling evidence for this view by showing that motor adaptation under uncertainty follows predictions derived from Bayesian inference. Knill and Pouget (2004) extended this argument, proposing that the brain's neural architecture is specifically designed to represent probability distributions rather than single, fixed values. Therefore, using a probabilistic model is not only a statistical convenience. It allows capturing and modeling the stochastic nature of the human decision making behavior.

Finally, the Bayesian approach carries a practical advantage for handling incomplete datasets, which is a critical constraint in preference learning. Preference elicitation is an expensive effort. Each query demands a user's time and attention, and fatigue degrades response quality. In a pairwise comparison setting with  $N$  alternatives, collecting a complete set of comparisons requires  $N(N - 1)/2$  queries, a number that grows rapidly and quickly exceeds what any participant can tolerate. Bayesian methods address this constraint directly. They can infer preferences from sparse data by leveraging the prior distribution to fill in gaps where data is missing (Tenenbaum et al., 2011). This allows the model to converge on the user's preference function with significantly fewer queries than would be required by frequentist methods. Frequentist methods, which typically demand larger sample sizes to achieve statistical significance, fails to cover this constraint.

### 2.2.2 Foundations of Bayesian Inference

The origins of the Bayesian approach to statistics date back to the 18th century. Bayes (1763) first introduced the formula

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (6)$$

for two events  $A$  and  $B$ , and this equation is now known as "Bayes' Theorem". Laplace (1774) discovered this principle independently and expanded its applications. Even though this concept initially dominated the field, interest in it decreased over time because it became computationally intractable due to high-dimensional integrals. It experienced a resurgence in the late 20th century due to the development of efficient sampling algorithms like Markov Chain Monte Carlo (MCMC) and the increase in computational power.

Although Bayes' original theorem dealt with discrete events (e.g., a coin coming up heads), modern statistical inference applies this logic to random variables. In this thesis, the unknown parameters of the model (such as the user's haptic preference) are not treated as fixed constants, but rather as random variables characterized by a probability distribution. This distinction is important: it enables the quantification of uncertainty not just about the outcome of an experiment, but about the structure of the world itself.

In this context, we replace the abstract events  $A$  and  $B$  with the hypothesis (or parameter vector)  $\theta$  and the observed data  $y$ . This transformation yields the fundamental equation of

Bayesian inference for continuous random variables:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \quad (7)$$

Here,  $p(\theta)$  is the prior distribution,  $p(y|\theta)$  is the likelihood,  $p(y)$  is the marginal likelihood (or evidence), and  $p(\theta|y)$  is the posterior distribution.

A Bayesian model consists of three fundamental elements: a prior distribution, a likelihood function, and the resulting posterior distribution.

- **Prior Distribution:** The prior distribution reflects prior beliefs or knowledge before observing the data. Its primary advantage is that former results from the domain can be integrated into the statistical model. Priors can be informative, weakly informative, or non-informative (flat). There are also conjugate priors, which are specific families of distributions. A prior is conjugate to a likelihood if the resulting posterior belongs to the same family as the prior. This allows for closed-form analytical solutions, bypassing the need for complex integration. Common examples include the Beta-Binomial and Gamma-Poisson pairs.
- **Likelihood Function:** The likelihood function describes the probability of observing the specific data  $y$  given a set of parameter values  $\theta$ . The choice of likelihood is determined by the generative process assumed for the data. For instance, the Bernoulli and Binomial likelihoods are used for binary outcomes, while the Poisson likelihood is used for count data. The likelihood embodies the “evidence” provided by the data.
- **Model Evidence:** Also known as the marginal likelihood, this is the denominator of Bayes’ theorem ( $p(y) = \int p(y|\theta)p(\theta)d\theta$ ). It acts as a normalizing constant to ensure the posterior distribution sums (or integrates) to one. Despite being often treated as a constant during parameter estimation, it is crucial for Bayesian Model Selection.
- **Posterior Distribution:** The posterior distribution is the synthesis of the prior and the likelihood. It represents the updated state of knowledge about the parameters after observing the data and serves as the complete solution to the inference problem. Unlike a point estimate, the posterior contains all information about the uncertainty of the parameters.

The selection of the prior and likelihood introduces subjectivity to the model since it depends on the person who designs the Bayesian model. Frequentists criticize Bayesian modeling because of this subjectivity, arguing that science must be “objective”. Moreover, if two different scientists analyze the same data with different models, the results will be different. Bayesians counter these critiques by stating that all models involve some assumptions and generalizations which depend on the creator of the model. As Berger and Berry (1988) argue in their seminal critique, standard frequentist methods are not purely objective; they depend heavily on the intention of the researcher and the experimental design. In contrast, the Bayesian framework forces the researcher to explicitly state their assumptions in the form of a prior. Gelman (2008) responds to this argument by noting that “Subjectivity” is just another word for “information,” and all available information should be used. He further states that this

transparency allows the “subjectivity” to be critiqued and understood, rather than remaining implicit. Furthermore, in the context of modeling human preference, this feature is a strength, not a weakness since it allows the model to formally capture the user’s subjective internal state and expectations.

### 2.2.2.1 Inference Methods

The primary computational challenge in Bayesian analysis is to obtain the posterior distribution. This is mainly because exact inference of the posterior is only possible when a conjugate pair is used for the prior and likelihood. While conjugate priors are mathematically straightforward, they are often restrictive. When exact inference is intractable (i.e., the integral cannot be solved analytically), two main classes of approximation methods are used:

1. **Sampling Methods:** MCMC algorithms such as Gibbs Sampling and Hamiltonian Monte Carlo, attempt to construct a chain of samples that converge to the true posterior distribution Gelman et al. (2013).
2. **Variational Inference (VI):** Unlike sampling, VI treats inference as an optimization problem. It approximates the true posterior with a simpler distribution by minimizing the divergence (usually KL-divergence) between the two (Blei et al., 2017).

### 2.2.2.2 Posterior Predictive Distribution

In many applications of Bayesian inference, predicting future observations ( $y^*$ ) that are unobserved is an important objective. This is achieved through the posterior predictive distribution, which marginalizes over the uncertainty in the parameters:

$$p(y^*|y) = \int p(y^*|\theta)p(\theta|y)d\theta \quad (8)$$

This distribution gives the probability of a new data point  $y^*$ , given the observed data  $y$ , fully accounting for the uncertainty in the model parameters (Bishop, 2006). Unlike parameter estimation, which describes the latent structure of the model, the posterior predictive distribution forecasts actual data values at new input locations. This capability allows for the evaluation of hypothetical scenarios, enabling the model to quantify uncertainty in regions of the design space that have not yet been empirically tested. This distinction is foundational for the active learning strategies employed in this thesis, where the model must proactively identify unobserved areas with high potential information gain.

### 2.2.3 Broader Applications of Bayesian Inference

Bayesian methods are used as a powerful tool in several domains where data must be collected from human participants, such as medical decision-making. Webb and Sidebotham (2020)

state that these methods help clinicians interpret the results of disease testing, especially by calculating the probability that a patient who tests positive for a disease actually has it.

Bayesian methods are also used to optimize treatment decisions. In the field of treatment, Pascual and Jain (2025) applied Bayesian Adaptive Randomization to clinical trial designs. Their model continuously evaluates incoming data to adaptively adjust treatment assignments based on efficacy.

Furthermore, Bayesian models are commonly used in phylogenetics. Zhang and Matsen IV (2024) introduced a VI approach as an efficient alternative to MCMC to estimate the posterior probability of a specific tree structure and set of branch lengths given DNA/RNA sequence data.

Finally, Bayesian inference is used not just for fitting a model but also for choosing the best theory among competing ones. Gigiberia (2025) demonstrated this by comparing two models of Dark Energy using Bayesian Model Selection, arguing that Bayesian modeling is particularly convenient when models are complex or hierarchical.

#### **2.2.4 Gaussian Processes**

Models can be categorized based on their parameters. A model is defined as parametric if it possesses a fixed number of parameters, regardless of the amount of data (e.g., linear or logistic regression). Conversely, if the number of parameters is not fixed or grows with the sample size, it is classified as a non-parametric model. While parametric models are often simpler and computationally faster, non-parametric models are better able to capture the complexity of the real world, though often at the cost of computational tractability (Murphy, 2012).

Several non-parametric approaches exist for different types of learning problems. For classification,  $k$ -Nearest Neighbors ( $k$ -NN) is a foundational non-parametric method. It makes no assumptions about the underlying function, as the model is effectively the labeled training data itself; however, it often suffers from poor performance in high-dimensional spaces due to the curse of dimensionality (Murphy, 2012).

In the domain of clustering, Dirichlet Processes (DP) offer a Bayesian non-parametric alternative to finite mixture models. A DP is technically defined as a distribution over distributions. DPs allow the model to discover clusters dynamically from observed data, theoretically drawing from an infinite number of potential components (Murphy, 2012).

For sequential data, standard Hidden Markov Models assume a time series of states governed by a fixed set of transition and emission parameters, making them parametric models. While effective for tasks like natural language processing, the fixed state space can be limiting. The Infinite Hidden Markov Model expands this by allowing for an unbounded number of hidden states. It uses a Hierarchical Dirichlet Process to share an infinite set of atomic states across transition distributions, enabling the model to infer the number of hidden states directly from the data sequence (Beal et al., 2001).

However, for tasks requiring continuous function approximation and uncertainty quantification, Gaussian Processes are often preferred. A Gaussian Process (GP) is a non-parametric Bayesian

model that does not assume a specific functional form for the data and defines a distribution over functions. Mathematically, a GP is defined as a collection of random variables, any finite number of which have a joint multivariate normal distribution (Rasmussen and Williams, 2006).

A GP is completely specified by its mean function,  $m(x)$ , and its covariance function,  $k(x, x')$ . The mean function represents the expected value of the latent function at input  $x$ , and is often assumed to be zero for simplicity. The covariance function, commonly referred to as kernel, defines the similarity between two input values. The GP is denoted as:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')) \quad (9)$$

There are several kernel functions that are used in practice. However, Radial Basis Function (RBF), also known as Squared Exponential, is a common choice. Samples drawn from the prior will exhibit the structural characteristics encoded in the kernel, and it encodes the prior beliefs about the data structure, similar to linear models. For instance, if a periodic kernel is used as prior, sine-wave-like samples will be generated (Rasmussen and Williams, 2006).

A key advantage of GPs is that the posterior mean and variance are calculated using closed-form expressions, eliminating the need for stochastic approximation or sampling. The predictive distribution for a new point  $x_*$  is Gaussian with mean  $\bar{f}_*$  and variance  $\mathbb{V}[f_*]$  given by:

$$\bar{f}_* = k(x_*, X)[K(X, X) + \sigma_n^2 I]^{-1} y \quad (10)$$

$$\mathbb{V}[f_*] = k(x_*, x_*) - k(x_*, X)^\top [K(X, X) + \sigma_n^2 I]^{-1} k(x_*, X) \quad (11)$$

where  $X$  represents the training inputs,  $K(X, X)$  is the covariance matrix of the training data, and  $\sigma_n^2$  is the noise variance.

This formulation provides exact values for the prediction and uncertainty. Inversion of the covariance matrix can be computationally intensive ( $\mathcal{O}(N^3)$ ). However, it avoids the extensive sampling often required by other non-parametric models, making it highly efficient for applications where quantifying uncertainty is critical. As Rasmussen and Williams (2006) note, GPs represent a rare exception in machine learning where the models are both highly flexible and analytically tractable.

GPs are generally applied to physical measurements, but they work equally well for subjective quantities. Yannakakis et al. (2009) applied GPs as an Instance Preference Learning mechanism to construct cognitive models. The model successfully approximated a utility function to predict children’s entertainment preferences from gameplay features.

Building on the success of GPs in preference learning, Bıyık et al. (2020) proposed an active preference-based framework specifically for reward learning in robotics. Their method utilizes GP regression to learn a non-parametric reward function from pairwise comparisons. Crucially, they demonstrated that by actively selecting queries that maximize information gain (using an acquisition function based on entropy), they could learn complex reward functions with significantly fewer user interactions compared to random query selection. This reinforces the utility of GPs not just for modeling static preferences, but for actively guiding the learning process in high-dimensional continuous domains.

### 2.2.5 Active Learning

In standard machine learning, models typically learn from a fixed dataset of randomly sampled queries. However, in many real-world applications, queries are expensive and bounded by the user’s time and fatigue. Randomly selecting query points in these scenarios is often inefficient, as it may waste the limited query budget on uninformative data.

Active Learning proposes an approach to this problem by intelligently selecting the next data point to query. The primary goal is to maximize the information gained from each query. In this way, it would require fewer samples to learn the latent function. Common strategies for Active Learning include Uncertainty Sampling, Query-By-Committee, and Expected Model Change (Settles, 2009).

Generally, the objective of Active Learning is to learn the entire function  $f(x)$  accurately across the whole search space, reducing global uncertainty. However, in many preference learning tasks, the goal is not to know the utility of every possible setting, but rather to locate the user’s preferred setting as quickly as possible. Bayesian Optimization provides a solution to this problem, as it addresses finding the global maximum or minimum of an unknown function (Shahriari et al., 2016).

Standard active learning often focuses on minimizing uncertainty. In addition to that, more recent approaches prioritize the user’s regret. Wilde et al. (2020) proposed a query selection method for robot behavior that minimizes “Maximum Regret”. Rather than relying only on cost function weights, this approach resulted in faster convergence and generated queries that were cognitively easier for users to answer.

### 2.2.6 Bayesian Optimization and Acquisition Functions

Bayesian Optimization is a strategy used to optimize “black-box” functions where the objective function is expensive to evaluate and lacks a closed-form derivative. In the context of preference learning, if the goal is to find the single best setting for the user efficiently, Bayesian Optimization is the ideal approach. Bayesian Optimization is particularly well-suited for problems where the objective function is expensive to evaluate, such as user studies where data acquisition requires human time and attention (Brochu et al., 2010).

As mentioned earlier, Active Learning seeks to reduce uncertainty everywhere. Unlike Active Learning, Bayesian Optimization directs the sampling process toward regions that are likely to contain the global optimum (the user’s favorite setting). It achieves this by constructing a probabilistic “surrogate model”. A surrogate model is a statistical approximation of the true objective function that is computationally cheap to evaluate.

While various models can serve as surrogates, GPs are the most common choice due to their flexibility and ability to provide calibrated uncertainty estimates. The surrogate model combines a prior distribution with an observation model to predict a “Utility Score”  $f(x)$  over the search space (Shahriari et al., 2016).

The core of the Bayesian Optimization framework involves a sequential process:

1. The surrogate model is updated with all available data.
2. The Acquisition Function is optimized to select the next most promising point to query.
3. The objective function is evaluated at this new point, and the result is added to the dataset.

To decide which point to query next, Bayesian Optimization relies on an Acquisition Function. This function evaluates the utility of querying a specific point based on the model's current posterior distribution (mean and variance). It balances exploration (sampling areas of high uncertainty) and exploitation (sampling areas predicted to have high utility). Common acquisition functions are:

- **Upper Confidence Bound:** A strategy that constructs a statistical upper bound for the function value using the formulation  $\mu(x) + \kappa\sigma(x)$ . It is optimistic, favoring points with high variance or high mean predictions, and has been shown to have strong theoretical guarantees on regret bounds Srinivas et al. (2010).
- **Probability of Improvement:** One of the earliest acquisition strategies, proposed by Kushner (1964). It selects the point most likely to yield a result better than the current best observation ( $f(x^+)$ ), regardless of how much better it might be.
- **Expected Improvement (EI):** Popularized by Jones et al. (1998) in the “Efficient Global Optimization” (EGO) algorithm, this function improves upon PI by considering the magnitude of the potential improvement. It calculates the expected value of the improvement over the current best observation, defined as  $\mathbb{E}[\max(0, f(x) - f(x^+))]$ . Because it naturally balances the probability of improvement with the size of the gain, EI is widely regarded as a robust default choice for global optimization.
- **Thompson Sampling:** A randomized strategy where a function is sampled from the GP posterior, and the point that maximizes this sampled function is selected. It naturally balances exploration and exploitation through the randomness of the sampling (Chapelle and Li, 2011).
- **Entropy Search / Predictive Entropy Search:** Information-theoretic methods that select the point expected to maximally reduce the uncertainty (entropy) regarding the location of the global optimum (Hennig and Schuler, 2012).

Probability of Improvement focuses only on the likelihood of finding a better point, which it often leads to excessive exploitation in local neighborhoods. EI is generally preferred as it incorporates the magnitude of the potential gain, offering a more robust balance between exploration and exploitation Brochu et al. (2010).

Bayesian Optimization has been successfully applied across various domains where data acquisition is costly or relies on human feedback. Brochu et al. (2010) demonstrated the efficacy of Bayesian Optimization in a “Preference Gallery” application for material design. In their study, they utilized a GP with a probit likelihood model to infer the latent utility function from discrete pairwise preferences.

Similarly, in the field of robotics, Martinez-Cantin et al. (2009) applied Bayesian Optimization to the problem of online path planning for a mobile robot. By modeling the cost function with a GP, the authors demonstrated that balancing exploration and exploitation using Bayesian Optimization outperformed purely exploratory heuristics (such as maximizing uncertainty), allowing the robot to achieve its sensing goals with fewer physical movements. Lizotte et al. (2007) applied this framework to automatic gait optimization for a quadruped robot. By modeling the gait velocity with a GP and utilizing the “Most Probable Improvement” acquisition function, they were able to find effective gait parameters with drastically fewer evaluations than local gradient-based methods. In the domain of computer vision, Kapoor et al. (2007) applied a similar GP-based active learning framework to the problem of object categorization. Recognizing that manually labeling images is a labor-intensive process, they developed a method where the model actively selects unlabeled images that maximize the margin of the GP classifier while minimizing uncertainty.

Collectively, the results from these studies show that using an active acquisition strategy significantly reduces the number of user interactions required to find the target compared to random selection. This highlights the effectiveness of Bayesian Optimization in minimizing user fatigue during subjective tasks and validates the use of GPs as a robust surrogate model for guiding efficient data acquisition in human-in-the-loop systems.

### 2.2.7 Coactive Feedback

Coactive Learning was originally developed by Shivaswamy and Joachims (2012) as an interactive online learning algorithm. In this model, for each user query  $x$ , a candidate object  $y$  is presented to the user. The user then provides feedback  $y^*$  by slightly improving the given object  $y$ . This correction implies that the utility of the user’s modified  $y^*$  is strictly greater than the predicted object  $y$ . Therefore, this improvement implies a preference information  $y^* \succ y$ . In addition, this coactive feedback indicates a direction of improvement in the search space, acting as a gradient that allows the learning algorithm to converge more efficiently.

Coactive feedback has been applied in robotics successfully. Jain et al. (2013) applied this method to help robots learn object trajectories. In their study, instead of demonstrating a perfect path from scratch, users slightly corrected the robot’s proposed path in the right direction. Small corrections were easier to provide than full demonstrations and the iterative process converged faster.

More recently, Tucker et al. (2020) integrated coactive feedback into a preference learning algorithm for high-dimensional optimization of lower-body exoskeleton gaits. Since pairwise comparisons are slow in large search spaces, they combined coactive feedback with Bayesian optimization. In this approach, the system generates a set of candidate gaits along a one-dimensional subspace and asks the user to select the best one. The algorithm interprets this selection as a coactive improvement over the baseline. This allows the model to be updated efficiently without requiring complex manual adjustments from the user. This distinct form of coactive feedback significantly improved the convergence rate compared to standard random search or traditional Bayesian Optimization.

Meanwhile Jain et al. (2013) and Tucker et al. (2020) demonstrated the efficacy of these interactive learning frameworks in robotics, recent research has extended these principles to haptic perception. Tolasa et al. (2025) applied a Human-in-the-Loop optimization framework to multi-modal haptic rendering, where researchers aiming to maximize “perceived realism”. Similar to the preference-based approaches in robotics, their method did not rely on ground-truth physics but used iterative qualitative user feedback, in order to tune visual-haptic parameters. This successful application confirms that subjective haptic qualities can be effectively optimized using human-in-the-loop algorithms, validating the core approach of this thesis.

## 2.3 Haptic Perception

Haptic perception differs from vision or audition fundamentally, because it is active. Generating haptic sensory information requires voluntary movement. Lederman and Klatzky (2009) define this as “active touch,” and observe that humans naturally adopt distinct motor strategies to isolate and perceive specific physical properties. The present thesis focuses specifically on kinesthetic perception, where sensory signals originate from mechanoreceptors in muscles, tendons and joints, rather than from receptors in the skin

Kinesthetic perception can be tested through three parameters: stiffness, damping, and distance. From a psychophysical perspective, these three parameters cover the relationship between force, motion, and time which is also called Mechanical Impedance. These three parameters form the fundamental "building blocks" of how humans perceive the physical world through touch.

### 2.3.1 Perception of Stiffness (K), Damping (B) and Distance (D)

Stiffness is defined physically as the relationship between force and displacement ( $F = K \cdot x$ ), or in the case of rotational movement, torque and angular position ( $\tau = K \cdot \theta$ ). It serves as the primary cue for identifying material composition. To feel stiffness, the brain must integrate two pieces of information: how far the hand has moved (proprioception) with how much force it feels (mechanoreception).

Jones and Hunter (1990) investigated the psychophysics of stiffness perception and found that humans are generally good at estimating stiffness. However, they also observed that humans are much less consistent in perceiving changes in stiffness compared to force and position. Participants used both force and displacement cues to perceive stiffness. Furthermore, their results confirmed that stiffness perception follows Weber’s Law, meaning the JND is proportional to the stimulus magnitude. Therefore, as stiffness increases, the absolute change required for a user to perceive a difference also increases. Practically, this implies that users can easily distinguish between compliant (soft) environments, but their ability to differentiate between rigid (hard) environments is significantly reduced.

Damping (or viscosity) is the resistance to motion proportional to velocity ( $\tau = B \cdot \dot{\theta}$ ). Unlike stiffness, which can be felt even when holding still, damping is a dynamic property, therefore,

it can only be perceived during movement. This effect is analogous to walking through water: while standing still generates no resistance, attempting to move faster creates an opposing drag force that increases with speed.

Jones and Hunter (1993) investigated the psychophysics of this dynamic property and found that human sensitivity to damping is significantly lower than for stiffness. They reported a Weber fraction of 0.34, which is approximately 50% larger than the threshold for stiffness (0.23). In addition to this, since the resistive force is velocity-dependent, the perception of damping is tied to exploration speed. Authors noted that consistent perception requires consistent movement. In other words, if the user stops moving, the damping force vanishes.

In control theory, distance is defined as a dead zone where no force is rendered until angular displacement exceeds a threshold. In psychophysics, perceiving this “gap” relies heavily on proprioception, which is the ability to sense the position and movement of limbs without visual feedback. Research conducted by Tan et al. (1994) on the resolution of fingertip displacement found that humans are highly sensitive to position changes, capable of detecting displacements as small as 1-2 mm. However, in a rotational context involving the wrist, this sensitivity is lower. The perception of this parameter is distinct from  $K$  and  $B$  because it is geometric rather than dynamic. The user must register how far the limb traveled freely before force onset.

The psychophysical constraints in stiffness and damping may cause significant challenges for optimization algorithms. Since users cannot perceive small perturbations in haptic parameters, a standard gradient descent approach might fail if the step size is too small. Although, using a Bayesian preference learning framework can handle noisy feedback and explore the search space in steps large enough to be perceptually distinct to the human operator. Knill and Pouget (2004) support this with their argument that the nervous system itself represents sensory information probabilistically rather than as fixed values, a view known as the “Bayesian Coding Hypothesis”.

Probabilistic models are increasingly used in studies mapping physical device parameters to human perception. For instance, Browder et al. (2019) applied a model to haptic rendering. They utilized a non-parametric Bayesian approach to model the perceived stiffness of virtual contact events. They argued that since the relationship between vibrotactile parameters and perceived hardness is complex and lacks a simple *a priori* mathematical model, GPs offer a superior solution. By comparing their Bayesian approach against the traditional “Method of Constant Stimuli,” they showed that GPs could accurately predict the psychometric curve of stiffness perception with significantly greater data efficiency. This validates the use of GPs as a surrogate model for capturing the non-linearities of human haptic sensitivity.

## 2.4 Summary

This chapter reviews the theoretical and methodological foundations underlying preference-based learning for haptic parameter optimization. Several key findings emerge from this literature that directly affect the adopted approach in this thesis.

First, research in behavioral economics and cognitive psychology states that human preferences are inherently noisy, context-dependent, and often violate the axioms of EUT. As a result, this

situation requires probabilistic models, since they can explicitly represent uncertainty, while deterministic approaches assume consistent, noise-free responses.

Second, the psychophysics literature establishes that pairwise comparisons provide more reliable and cognitively tractable data than absolute ratings or direct utility elicitation. Thurstone's Law of Comparative Judgment provides the theoretical bridge connecting these comparative responses to latent utility estimation. On the other hand, Weber's Law characterizes the perceptual thresholds that limit discrimination ability.

Third, GPs emerge as a particularly suitable modeling framework for this domain. Their non-parametric nature allows them to capture utility functions without requiring strong assumptions about functional form. GPs provide principled uncertainty calculations, which is essential for both interpreting results and guiding active query selection.

Fourth, Bayesian Optimization and Active Learning strategies overcome the constraint of limited human attention and capability. With intelligent selection of queries, preference models can learn accurately and require significantly fewer interactions than random or exhaustive sampling.

Finally, the haptic perception literature reveals that stiffness, damping, and distance each have distinct psychophysical characteristics. These perceptual limitations emphasize the importance of using learning algorithms that can handle noisy feedback and inconsistent responses.

Bayesian preference learning has been successfully applied in robotics for trajectory optimization and gait tuning. However, its application to multi-parameter haptic rendering, where the goal is to estimate physical parameters from subjective perception, represents a significant contribution. This thesis builds upon the foundations established in this chapter to develop and validate a preference-based optimization framework specifically tailored to the challenges of haptic parameter tuning.



## CHAPTER 3

### METHOD

In this chapter, the experimental methodology and computational framework that is used are given in detail. First, this chapter begins with an overview of the computational methods which includes Gaussian Process model for pairwise comparisons. Moreover, chapter gives details about active learning framework used for adaptive query selection. Secondly, the experimental method is described. It covers the haptic device specifications, haptic models, stimuli parameters, experimental design, and characteristics of participants who joined the experiment.

#### 3.1 Computational Method

##### 3.1.1 Preference Query Type

The core objective of the computational framework is to learn the user’s hidden preference function efficiently while minimizing cognitive fatigue. As discussed in Section 2.1.4, eliciting haptic preferences is constrained by the rapid decay of sensory memory.

Alternative methods such as absolute ratings, full ranking, or selecting the best from a set were excluded, as they impose a heavy memory burden. These tasks require the user to retain and compare the sensations of multiple sequential stimuli simultaneously to make a judgment. Due to the transient nature of haptic memory, this reliance on retention leads to high response variability.

To address this, this study adopts a Pairwise Comparison paradigm. By requiring the user to compare only two immediately presented stimuli, this method minimizes the working memory load and eliminates the drift associated with rating scales. However, to compensate for the lower information density of binary choices (1 bit per query), the framework integrates Coactive Feedback (a form of relative critiquing). This hybrid approach allows users to provide directional guidance (e.g., “Make it Harder”) alongside their preference, combining the robustness of pairwise judgments with the efficiency of gradient-based search.

### 3.1.2 Gaussian Process

To estimate the optimal stimuli parameters, the latent utility function driving the pairwise comparisons must be modeled. In this research, Gaussian Processes (GPs) are employed to model this utility function.

As discussed in Section 2.2.4, GPs offer several distinct advantages for this experimental framework. First, as a non-parametric approach, the GP does not impose a fixed functional form (e.g., linear or quadratic) on the utility function. This flexibility allows the model to capture complex, non-linear relationships between the haptic stimuli and user preference without being constrained by a pre-specified number of parameters. Second, GPs provide a fully probabilistic framework that yields closed-form expressions for both the predictive mean and variance. This analytical tractability is critical for the active learning framework, as it allows for the efficient computation of uncertainty estimates required for adaptive query selection.

This study utilized the `prefGP` library, a Python-based implementation for preference learning using GP. Developed by Benavoli and Azzimonti (2024), this framework proposes a comprehensive GP model for learning a latent function, specifically a utility-based preference function. The authors included various models tailored to specific problem types, distinguishing based on the presence of noise, the availability of an “indifference” option versus forced choice, and whether the problem involves classification. While the likelihood function varies according to these problem specifications, the GP prior remains consistent across all models.

The `prefGP` library offers nine distinct models for preference learning, categorized by the type of feedback and the presence of noise (Benavoli and Azzimonti, 2024). In the present study, participants were restricted to a forced-choice paradigm between two options, with no “indifferent” selection permitted. To account for this experimental design and potential inconsistencies in human judgment, Model 3: Probit for Erroneous Preferences was adopted. This model is specifically designed for pairwise comparisons where participant responses may include stochastic errors.

#### 3.1.2.1 Prior

A zero-mean Gaussian Process prior was placed over the latent utility function  $u$ :

$$u(\mathbf{x}) \sim \mathcal{GP}(0, k(\mathbf{x}, \mathbf{x}')) \quad (12)$$

where  $k(\mathbf{x}, \mathbf{x}')$  is the kernel function defining the covariance between different haptic stimuli. A zero-mean assumption implies no prior bias towards specific parameter regions being preferred.

For this experiment, the Squared Exponential kernel, also known as the Radial Basis Function kernel, was employed. This choice enforces a prior assumption that the underlying preference function is smooth—meaning similar haptic parameters should yield similar utility values

(Benavoli and Azzimonti, 2024). The kernel is defined as:

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\ell^2}\right) \quad (13)$$

where  $\sigma_f^2$  is the signal variance, which scales the magnitude of the utility function, and  $\ell$  is the length-scale parameter, which controls smoothness by determining how quickly the correlation between data points decays as the distance in parameter space increases.

### 3.1.2.2 Likelihood

In our experimental design, participants were forced to choose one option from a pair (Left vs. Right), even when the stimuli were perceptually very similar. The adopted “Model 3: Probit for Erroneous Preferences,” from the work of Benavoli and Azzimonti (2024), accounts for potential inconsistencies arising from this forced choice paradigm. This model assumes that the probability of a participant stating a preference for  $x_i$  over  $x_j$  (denoted as  $x_i \succ x_j$ ) is a function of the difference in their utilities, scaled by a noise parameter  $\sigma$ . The likelihood is given by the Probit function:

$$p(x_i \succ x_j | u) = \Phi\left(\frac{u(x_i) - u(x_j)}{\sigma}\right) \quad (14)$$

where  $\Phi(\cdot)$  is the Cumulative Distribution Function of the standard Normal distribution, and  $\sigma > 0$  is a scaling parameter that models the error rate or “fuzziness” in the user’s decision-making process (Benavoli and Azzimonti, 2024). This probabilistic approach makes the model robust against noisy data characteristic of psychophysical experiments, where users may struggle to distinguish between options with similar utility values (see Section 2.1.5; Thurstone, 1927).

### 3.1.2.3 Posterior

Given the dataset of pairwise comparisons  $\mathcal{D} = \{x_{i_s} \succ x_{j_s}\}_{s=1}^m$ , the posterior distribution over the utility function  $u$  is computed. Since the Probit likelihood is non-Gaussian, the posterior is not a standard GP. However, Benavoli and Azzimonti (2024) showed that the posterior for this model has an analytical form: a Skew-Gaussian Process (Skew-GP).

The `prefGP` library computes the Skew-GP posterior parameters exactly and draws samples using Linear Elliptical Slice Sampling (LINESS), an efficient algorithm for sampling from truncated multivariate normal distributions (Gessner et al., 2020). This approach avoids the approximation errors associated with Variational Inference or Laplace approximation, while remaining computationally tractable.

### 3.1.2.4 Optimal Estimate

To identify the algorithm’s current best estimate at each trial, we compute the parameter configuration that maximizes the posterior mean utility:

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{u \sim p(u|\mathcal{D})} [u(\mathbf{x})] \quad (15)$$

This *maximum of posterior mean* approach integrates over posterior uncertainty before committing to a decision, making it robust to noise in individual preference observations. This estimator is standard in Bayesian optimization and preference learning (Brochu et al., 2010).

### 3.1.3 Active Learning Framework

In order to efficiently identify the haptic parameters which maximize perceived similarity, a Bayesian Optimization framework was adopted. Details of the Bayesian Optimization framework are provided in the following sections.

#### 3.1.3.1 Acquisition Function: Penalized Expected Improvement

The Expected Improvement (EI) was employed as the acquisition function. The standard EI is defined as:

$$EI(x) = \begin{cases} (\mu(x) - f(x^+) - \xi)\Phi(Z) + \sigma(x)\phi(Z) & \text{if } \sigma(x) > 0 \\ 0 & \text{if } \sigma(x) = 0 \end{cases} \quad (16)$$

where:

- $Z = \frac{\mu(x) - f(x^+) - \xi}{\sigma(x)}$ ,
- $\mu(x)$  and  $\sigma(x)$  are the predictive mean and standard deviation of the utility at point  $x$ ,
- $f(x^+)$  is the value of the current best (incumbent) latent utility,
- $\xi$  is an exploration parameter (set to 0 in this implementation),
- $\Phi(\cdot)$  and  $\phi(\cdot)$  are the cumulative distribution function and probability density function of the standard normal distribution, respectively.

However, standard optimization approaches often fail to account for the limits of human perception. Recalling the concept of Just Noticeable Difference (JND) discussed in Section 2.1.5 and Section 2.3, humans require a certain magnitude of change to perceive a difference between two stimuli. For instance, if two haptic signals are numerically distinct but fall within the JND (e.g., a stiffness of 100 mNm/rad versus 102 mNm/rad), they will feel identical to the

user. Consequently, querying parameters that are close to the current best option is inefficient, as the user cannot provide reliable preference feedback on indistinguishable stimuli.

To address this limitation, a Penalized Expected Improvement function was implemented. González et al. (2016) proposed the Local Penalization principle, in which the acquisition value is forced toward zero in a specific region to prevent redundant sampling. Adapting this concept to psychophysics, a simplified penalty term is used in this research. This penalty term forces the system to select a candidate point sufficiently distant from the incumbent to ensure the user perceives a difference, thereby providing meaningful feedback. The penalized metric is calculated as:

$$EI_{penalized}(x) = EI(x) \cdot \left(1 - e^{-w \cdot \|x - x_{x^+}\|}\right) \quad (17)$$

where  $\|x - x_{x^+}\|$  is the Euclidean distance between the candidate point  $x$  and the location of the current best  $x_{x^+}$ , and  $w$  is the penalty weight (default  $w = 1.0$ ). This term forces the acquisition value toward zero as the candidate point approaches the incumbent ( $x \rightarrow x_{x^+}$ ). It effectively prevents the algorithm from becoming trapped in the immediate neighborhood of the known optimum.

### 3.1.4 Experiment Method

The primary objective of this experiment was to validate the GP preference learning framework introduced in Section 3.1. Specifically, we aimed to determine if the model could accurately converge to a set of target haptic parameters based purely on pairwise comparisons. To enable quantitative validation, a simulated reference with fixed parameters was established to serve as a ground truth. When using real physical objects, the “true” parameter values may be ambiguous or subjective. Therefore, a simulated reference was chosen to allow for a direct assessment of whether the model successfully recovers the target values.

#### 3.1.4.1 Haptic Device

The experiment utilized a custom 1-DOF haptic device, illustrated in Figure 2. The mechanical and embedded systems of this device were designed by Ali Seçkin Kaplan as a part of his MSc. studies at the Department of Mechanical Engineering, TED University. His work focuses on haptic rendering and the synthesis of realistic tactile sensations, ensuring the hardware provides high-fidelity feedback. The device connects to a host computer via USB, allowing real-time control of haptic parameters through a custom graphical user interface.

The system is powered by a direct-drive DC motor, which delivers controlled torque as a function of position to simulate stiffness and compliance sensations. Users interact with the device via a rotating knob, with the range of motion mechanically restricted to 90° to standardize the exploration workspace across all participants.



**Figure 2:** The custom 1-DOF haptic device used in the experiment.

### 3.1.4.2 Haptic Models and Stimuli

To characterize the fundamental haptic experience, three mechanical parameters were selected for estimation: Stiffness ( $K$ ), Damping ( $B$ ), and Distance ( $D$ ). These parameters were chosen because, collectively, they constitute a sufficient basis to simulate a wide range of realistic mechanical behaviors. In haptic rendering, interactions are commonly decomposed into elastic (stiffness), viscous (damping), and geometric (constraint) components to successfully approximate the feel of physical objects (Salisbury et al., 1995).

Stiffness and Damping constitute the core of the haptic sensation. Stiffness ( $K$ ) provides the primary restoring force, determining the “hardness” of the interaction. Damping ( $B$ ) adds a dynamic resistance proportional to velocity. While Damping in isolation may resemble dry friction, its primary role in this context is to modulate the dynamic stability and “feel” of the Stiffness.

The third parameter, Distance ( $D$ ), represents a geometric constraint rather than a direct haptic sensation. It defines the spatial threshold where the haptic forces (Stiffness and Damping) initiate. By incorporating this parameter, the model can simulate mechanical slack or “dead zones”, which are regions where the mechanism moves freely before engaging the spring force.

Based on these parameters, three distinct experimental conditions were employed: the K condition (stiffness only), the KB condition (stiffness and damping), and the KBD condition (stiffness, damping, and distance). Each condition incorporates a progressively more complex parameter set. The K condition provides only stiffness feedback, while the KB condition combines stiffness with viscous damping. The KBD condition introduces a geometric constraint: haptic feedback initiates only after a specified distance threshold ( $D$ ) is crossed. In the initial “dead zone” before this threshold, no sensation is perceived except for the device’s inherent dry friction. The haptic device receives these parameter values in real-time to generate the corresponding force feedback.

The parameter bounds were fixed to ensure safety and perceivability, as detailed in Table 1. The upper bound for stiffness ( $K$ ) was specifically determined by the experimental constraints; since participants were limited to rotating the knob through  $90^\circ$ , stiffness values exceeding

**Table 1:** Parameter Bounds and Ranges

Parameter	Symbol	Range
Stiffness (mNm/rad)	$K$	0 – 360
Damping (mNm·s/rad)	$B$	0 – 50
Distance (rad)	$D$	0 – 0.5

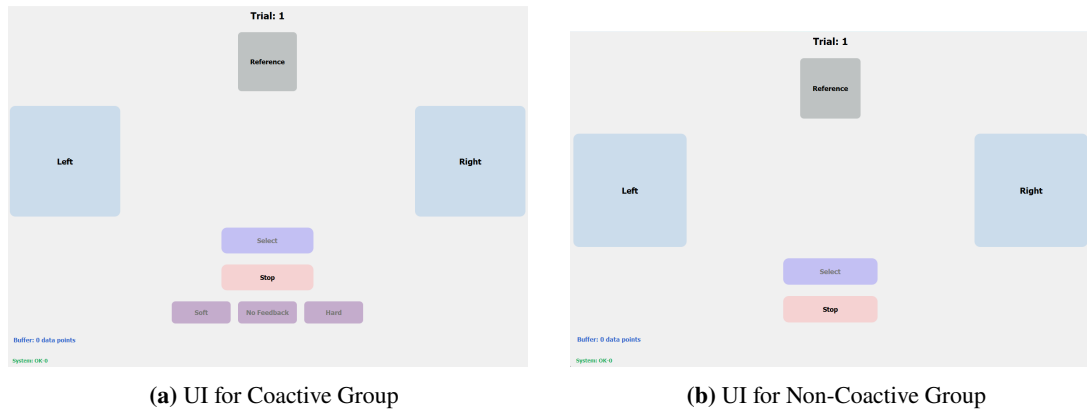
360 mNm/rad would generate torques that might not be fully distinct or reachable within this angular range. Furthermore, excessively high stiffness values were avoided to prevent participant fatigue during the repeated trials.

The required motion was standardized such that participants began at  $0^\circ$  and rotated the knob to  $90^\circ$  before returning to the starting position in a continuous movement. Participants were instructed not to pause at any position or release the knob, and to use their dominant hand.

Three LEDs on the device provided real-time feedback regarding movement speed:

- **Green LED:** Illuminated when participants maintained the desired speed range (0.5–2.5 rad/s).
- **Red LED:** Indicated speeds exceeding 2.5 rad/s.
- **No LED:** Indicated speeds below 0.5 rad/s.
- **Yellow LED:** Indicated positions outside the  $0^\circ - 90^\circ$  boundary.

This feedback mechanism aimed to standardize motion across participants, as haptic sensation varies significantly with speed and position.



**Figure 3:** User Interfaces used in the experiment.

The user interface presented Reference, Left, and Right option buttons, which can be seen in Figure 3. For participants in the coactive groups, additional feedback buttons were displayed:

Soft, No Feedback, and Hard. During each trial, participants experienced the options in the following sequence: Reference, Left, Reference, Right. The main question that was given to the participants was “Which of the Left or Right options is the most similar to the Reference?”. In order to answer this question, participants could manipulate the knob as many times as desired within each option. After completing this sequence, participants could revisit any option until reaching a decision. Participants in the coactive group were additionally asked to provide feedback by answering the following: "Does the chosen option feel softer or harder than the reference, or is there no feedback?". Button presses were logged, and for each press, position, speed, and torque data were recorded over time.

**Table 2:** Warm-up Trials (Reference = 150 mNm)

Trial	Left (mNm)	Right (mNm)
1	30	100
2	190	70
3	150	220
4	320	260

Each participant’s session begins with a warm-up phase consisting of four trials where the dry friction parameter ( $T_{fr}$ ) varies. These trials were predetermined and identical for all participants; the specific values are shown in Table 2. During and after the warm-up, participants were encouraged to ask questions about the procedure. The experiment instructions, including the preference question and the coactive feedback question, were explained during the first warm-up trial. The third trial served as a validation check. In this trial, the “Left” option was set to exactly match the Reference value (150 mNm). This trial was designed to assess the participant’s attention during the warm-up process. Additionally, for participants in the coactive group, the coactive feedback process was conducted during each warm-up trial.

### 3.1.4.3 Pilot Study: Selection of Coactive Feedback Vocabulary

A fundamental challenge in coactive learning is mapping the algorithm’s mathematical parameters (Stiffness and Damping) to the user’s semantic intent. While the algorithm operates on numerical gradients, users express preferences using subjective adjectives. To determine the most intuitive vocabulary for the coactive feedback mechanism, a preliminary pilot study was conducted prior to the main experiment.

The pilot study involved 16 participants who performed a series of 7 pairwise comparison trials with varying Stiffness ( $K$ ) and Damping ( $B$ ) parameters. In each trial, participants were asked two open-ended questions:

1. "Which of the options is the most similar to the reference?"
2. "To get closer to the reference, what needs to be done to the chosen option?"

Participants were free to use any adjectives to describe the required change. The responses were collected and analyzed for frequency. The analysis revealed that descriptors corresponding to

“Hard” and “Soft” were the most dominant and consistent terms used to describe differences in impedance. Synonyms such as “Difficult” or “Tight” were also present but appeared less frequently.

Crucially, participants rarely distinguished between elastic and viscous effects in their verbal descriptions, often conflating both properties under the general sensation of resistance. Based on these results, the binary labels “Harder” and “Softer” were selected as the sole feedback mechanism. Introducing separate vocabulary for distinct parameters (e.g., asking users to distinguish “Stiffer” from “More Viscous”) was avoided, as it would impose excessive cognitive load by requiring users to decouple complex mechanical properties. Instead, this single high-level descriptor pair allows participants to provide holistic feedback that aligns with their natural perception of impedance. Detailed frequency tables and trial configurations for the pilot study are provided in Appendix B.

### 3.1.4.4 Experimental Design

Participants were assigned to either the coactive group or the non-coactive group. In order to control order effects, the sequence of experimental blocks was counterbalanced across participants using six permutations (see Table 3). Assignment was sequential based on participation order (e.g., Participant 1 was assigned Order 1/Coactive, Participant 2 was assigned Order 2/Coactive, and so forth).

**Table 3:** Counterbalanced Permutations of Experimental Blocks

Order	Block 1	Block 2	Block 3
1	K	KB	KBD
2	K	KBD	KB
3	KB	K	KBD
4	KB	KBD	K
5	KBD	K	KB
6	KBD	KB	K

Each participant completed three experiments, with 14 trials per experiment. The first three trials in each experiment used predetermined values. These fixed parameter values are the same for the coactive and non-coactive groups. These are detailed in Tables 4, 5, and 6. The final trial served as a validation trial, in which one option was the reference itself and the other was either the option chosen in the preceding trial or the last option with coactive feedback applied.

**Table 4:** Fixed Trials for K Condition (Reference  $K = 200$ )

Trial	Left ( $K$ )	Right ( $K$ )
1	135	270
2	120	90
3	300	50

**Table 5:** Fixed Trials for KB Condition (Reference  $[K, B] = [175, 15]$ )

<b>Trial</b>	<b>Left <math>[K, B]</math></b>	<b>Right <math>[K, B]</math></b>
1	[210, 40]	[70, 10]
2	[280, 5]	[35, 45]
3	[300, 1]	[100, 35]

**Table 6:** Fixed Trials for KBD Condition (Reference  $[K, B, D] = [150, 20, 0.2]$ )

<b>Trial</b>	<b>Left <math>[K, B, D]</math></b>	<b>Right <math>[K, B, D]</math></b>
1	[45, 30, 0.04]	[300, 25, 0.3]
2	[220, 5, 0.1]	[100, 45, 0.5]
3	[250, 40, 0.4]	[80, 12.5, 0.25]

For Trials 4 through 13, one option was determined by the active learning method rather than being fixed. The other option was the selection from the previous trial for the non-coactive group, or the previous selection plus coactive feedback adjustment for the coactive group.

All preferences were added to the preference dataset  $X$ . For each trial, let the left option be  $L_i$  and the right option be  $R_i$ . For each preference of the participant, the pair is added to the preference data. If the left is chosen,  $(L_i, R_i)$  is added; if the right is chosen,  $(R_i, L_i)$  is added to  $X$ .

**Table 7:** Parameter Adjustments based on Coactive Feedback

<b>Feedback</b>	<b>K</b>	<b>B</b>	<b>D</b>
Soft	+25	+5	0
No Feedback	0	0	0
Hard	-25	-5	0

For participants in the coactive group, an additional pair reflecting the coactive feedback was added to the preference dataset. For any chosen option  $O_i$ , the following adjustments were applied based on feedback type:

**If the feedback is “Soft”:**

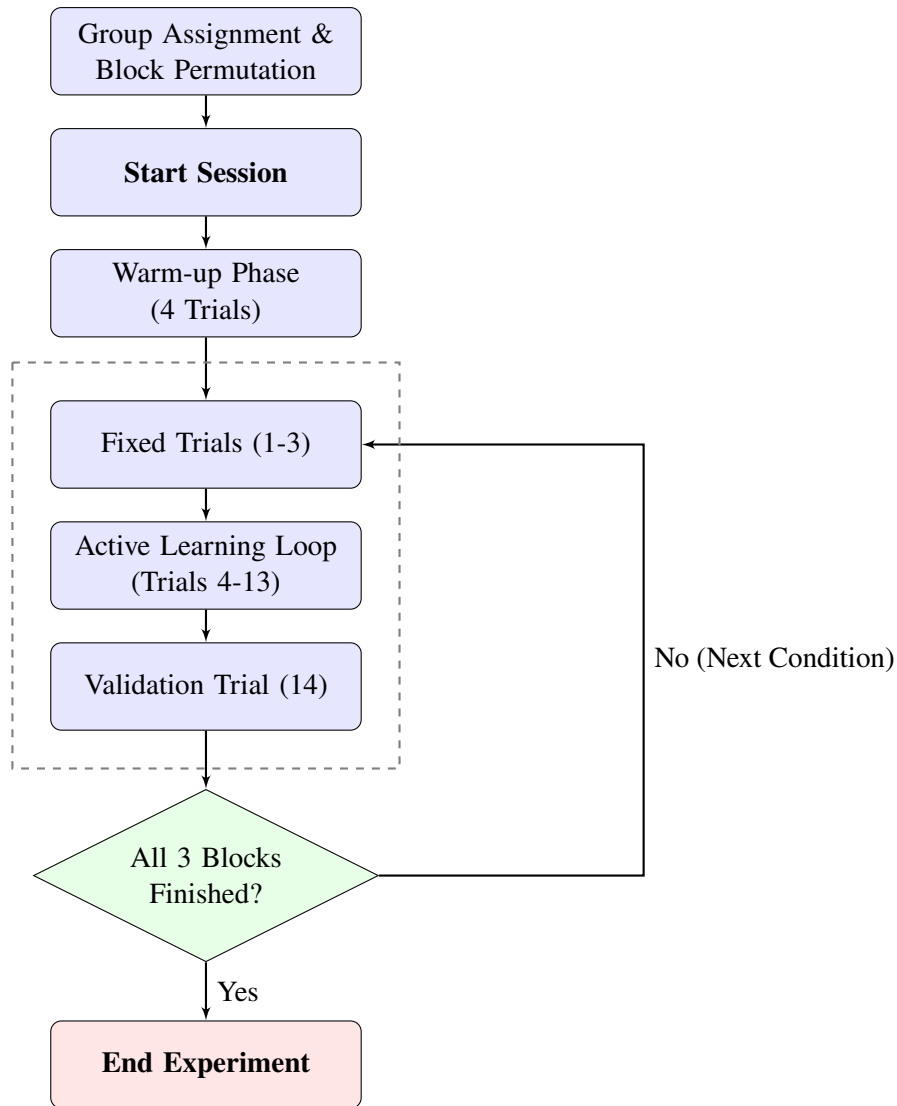
- If K condition  $\rightarrow O_i = [K_i]$ , then  $(K_i + 25, K_i)$  is added to  $X$ .
- If KB condition  $\rightarrow O_i = [K_i, B_i]$ , then  $([K_i + 25, B_i + 5], [K_i, B_i])$  is added to  $X$ .
- If KBD condition  $\rightarrow O_i = [K_i, B_i, D_i]$ , then  $([K_i + 25, B_i + 5, D_i], [K_i, B_i, D_i])$  is added to  $X$ .

**If the feedback is “Hard”:**

- If K condition  $\rightarrow O_i = [K_i]$ , then  $(K_i - 25, K_i)$  is added to  $X$ .

- If KB condition  $\rightarrow O_i = [K_i, B_i]$ , then  $([K_i - 25, B_i - 5], [K_i, B_i])$  is added to  $X$ .
- If KBD condition  $\rightarrow O_i = [K_i, B_i, D_i]$ , then  $([K_i - 25, B_i - 5, D_i], [K_i, B_i, D_i])$  is added to  $X$ .

If no feedback is given, then no other pair is added other than the preference. These parameter adjustments are summarized in Table 7. Each experimental session lasted approximately 30 minutes. The overall flow of the experiment, detailing the sequence from group assignment to the final validation check, is illustrated in Figure 4.



**Figure 4:** Overview of the Experimental Procedure.

### 3.1.4.5 Trial Generation Procedure

For Trials  $4 \leq t \leq 13$ , the comparison pairs were generated dynamically using the Bayesian Optimization loop. The procedure for each trial  $t$  is illustrated in Figure 5 and detailed below:

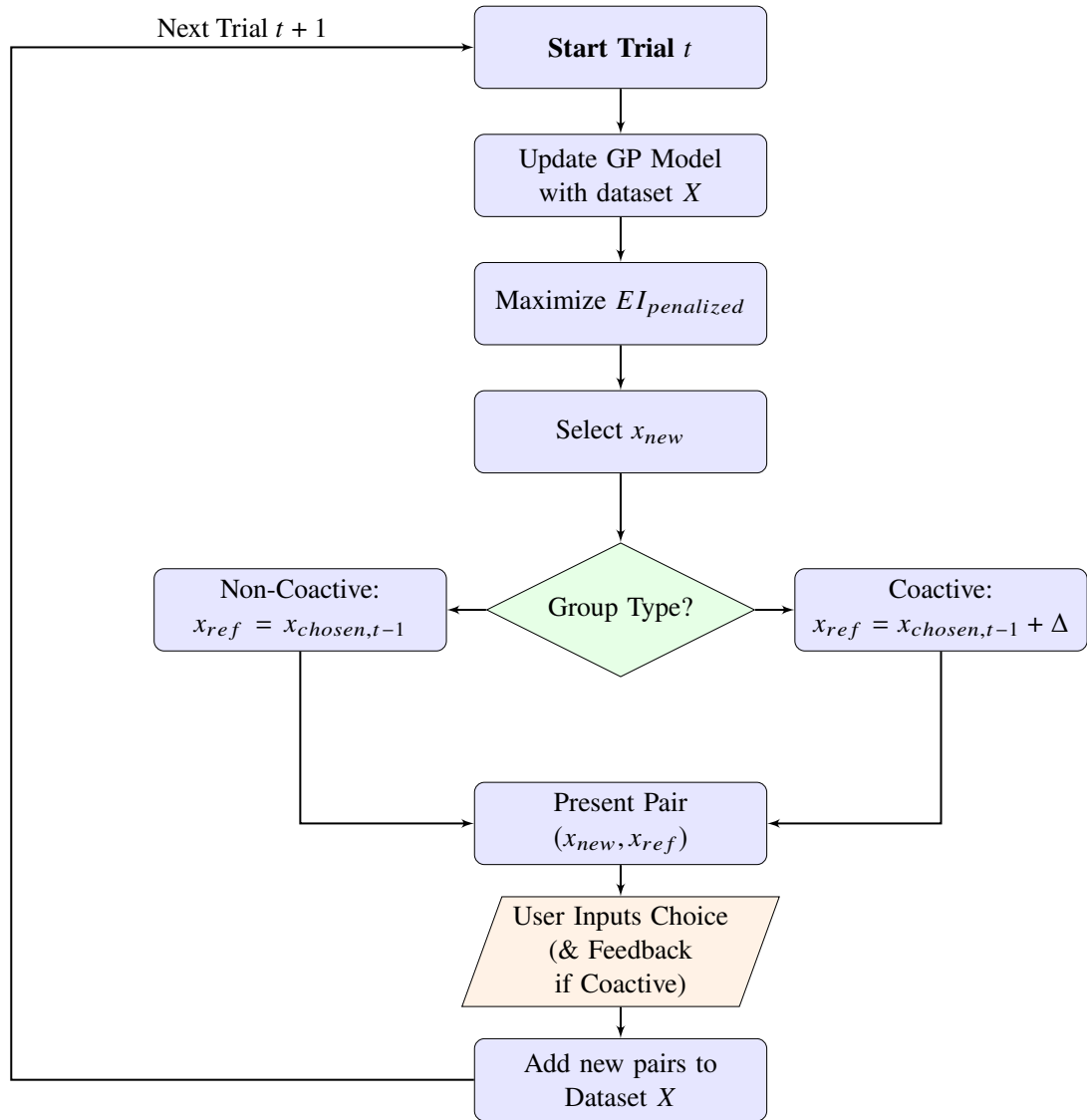
1. **Model Update:** The GP model is updated with all preference pairs observed up to trial  $t - 1$ .
2. **Optimization:** The system calculates the Penalized Expected Improvement ( $EI_{penalized}$ ) over the domain using the updated posterior.
3. **Selection of Option 1 ( $x_{new}$ ):** The first option is selected by maximizing the acquisition function:

$$x_{new} = \underset{x}{\operatorname{argmax}} EI_{penalized}(x) \quad (18)$$

4. **Selection of Option 2 ( $x_{ref}$ ):** The second option is derived from the previous trial ( $t - 1$ ):

- **Non-Coactive Group:**  $x_{ref} = x_{chosen,t-1}$  (the option selected by the user).
- **Coactive Group:**  $x_{ref} = x_{chosen,t-1} + \Delta_{coactive}$  (the chosen option modified by the feedback).

5. **Presentation:** The participant compares the pair  $(x_{new}, x_{ref})$ . The assignment to “Left” and “Right” buttons is randomized.



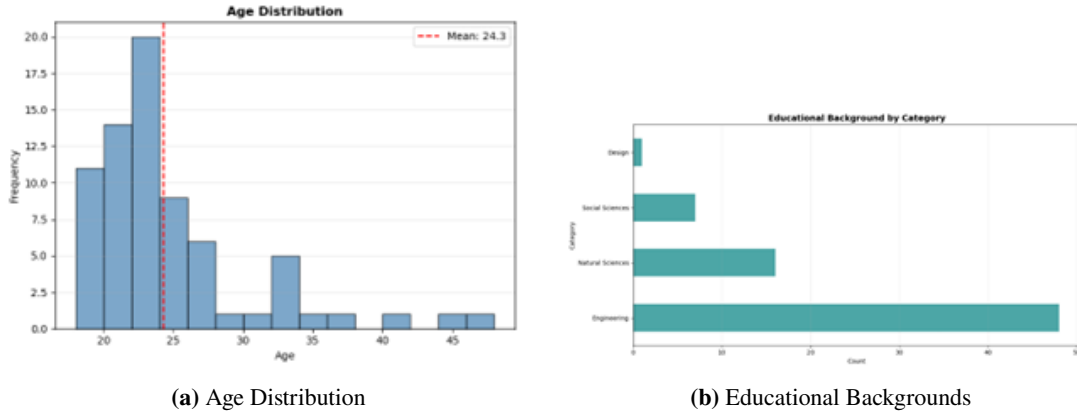
**Figure 5:** Detailed Trial Generation Loop (Trials 4-13).

### 3.1.4.6 Participants

A total of 72 volunteers (38 male, 34 female) participated in the study without compensation. All participants were naive to the experimental hypothesis and had no prior experience with the haptic devices. Participants were recruited from TED University ( $n = 44$ , 61.1%) and Middle East Technical University ( $n = 28$ , 38.9%) which are located in Ankara, Turkey.

The majority of participants were right-handed ( $n = 67$ , 93.1%), while a small minority were left-handed ( $n = 5$ , 6.9%). All participants performed the task using their self-reported dominant hand.

The age distribution of the participants is illustrated in Figure 6a. The mean age was 24.3 years ( $SD = 6.01$ ). As shown in the histogram, the sample consisted predominantly young adults, with ages ranging from 18 to 48 years.



**Figure 6:** Demographic information of the participants.

Participants represented diverse academic disciplines. The largest groups consist of students or graduates from engineering fields such as Industrial Engineering ( $n = 14, 19.4\%$ ), Mechanical Engineering ( $n = 11, 15.3\%$ ), and Computer Engineering ( $n = 10, 13.9\%$ ). Other participants represented fields including Psychology, Physics, Mathematics, Electrical Engineering, and Molecular Biology and Genetics. A detailed breakdown of educational backgrounds is provided in Figure 6b.

All participants reported having normal or corrected-to-normal vision and no history of neuromuscular disorders or injuries affecting the upper extremities that could interfere with the haptic task.

Prior to the experiment, all participants provided written informed consent. The study was conducted in accordance with the Declaration of Helsinki and was approved by the TED University Human Subjects Ethics Committee.

## CHAPTER 4

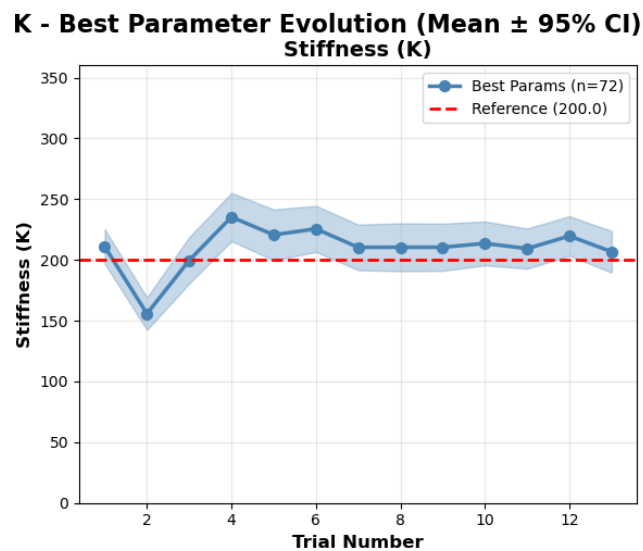
### RESULTS

#### 4.1 Learning Performance and Convergence

##### 4.1.1 Model's Performance

Firstly, the model's best estimate was evaluated for each trial (excluding the validation trial) regardless of the feedback group in order to determine overall model convergence. For each trial, the optimal parameter configuration was determined for every participant. These individual estimates were then averaged across all participants ( $n = 72$ ) to obtain a mean trajectory of convergence.

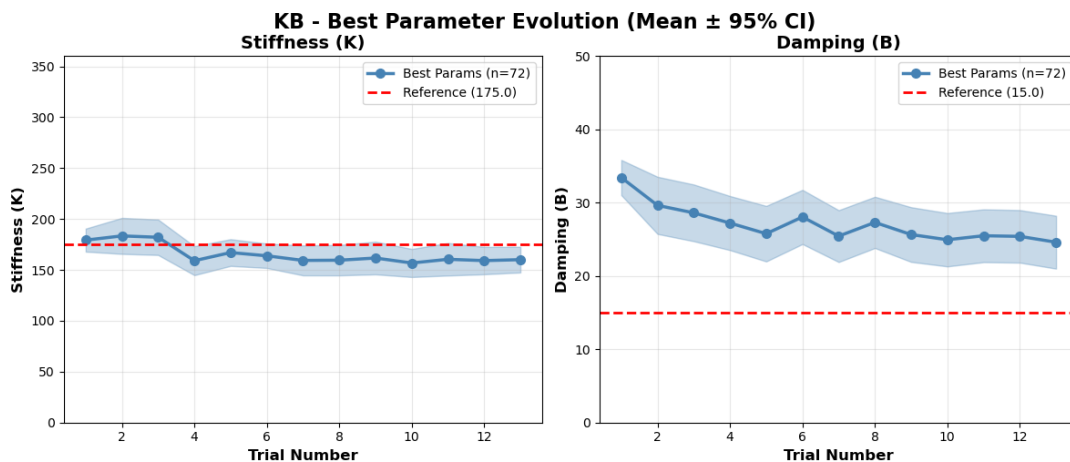
##### 4.1.1.1 K Condition



**Figure 7:** Evolution of the best stiffness estimate for the K condition ( $n = 72$ ). The shaded region represents the 95% confidence interval. After initial exploration, the mean estimate converges to 206.7 mNm/rad, staying within 3.3% of the reference value (200.0 mNm/rad).

In the most basic condition where only stiffness varied, the algorithm achieved high accuracy and precision. As shown in Figure 7, after initial exploration (Trials 1–3), the estimate stabilized near the reference. The final estimate at Trial 13 was 206.7 mNm/rad (95% CI: [189.6, 223.8]), deviating from the reference (200.0 mNm/rad) by only 3.3%. This trajectory demonstrates that the system successfully isolated the target stiffness.

#### 4.1.1.2 KB Condition



**Figure 8:** Evolution of best parameter estimates for the KB condition ( $n = 72$ ). Stiffness (Left) approximates the reference (175.0) with slight underestimation (8.5% error), while Damping (Right) stabilizes at a value significantly higher than the reference (64.1% error).

In the two-parameter condition, the evolution of the best estimates for stiffness and damping revealed distinct behaviors, as shown in Figure 8. The model estimated a final stiffness of 160.0 mNm/rad (95% CI: [147.4, 172.7]), which corresponds to an 8.5% deviation from the reference (175.0 mNm/rad). This estimate remained relatively stable across trials (Figure 8a).

On the other hand, the model overestimated damping at 24.6 mNm-s/rad (95% CI: [21.0, 28.2]), a 64.1% deviation from the reference (15.0 mNm-s/rad). This divergence can be seen in Figure 8b. While the model learned in trials 1 through 5, it did not improve further after trial 6. This pattern shows consistency with perceptual compensation: participants may have perceived higher damping as contributing to “hardness,” leading the algorithm to match the reference sensation through a combination of lower stiffness and higher damping.

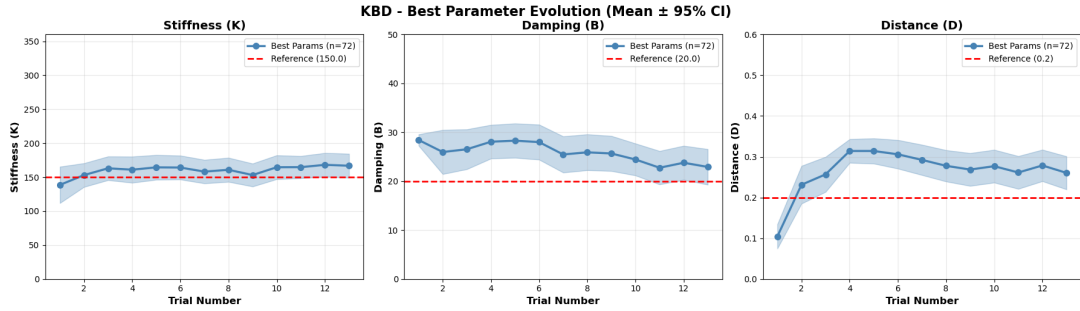
#### 4.1.1.3 KBD Condition

In the most complex condition, the system achieved reasonable convergence across all axes despite the increased dimensionality, as shown in Figure 9.

The Stiffness estimate ( $K$ ) stabilized near the reference (150.0 mNm/rad) early in the process and remained consistent throughout the session (Figure 9a), ending at 166.7 mNm/rad (11.2% deviation; 95% CI: [149.1, 184.4]).

The Damping estimate ( $B$ ) exhibited a continuous learning trend starting from Trial 6 (Figure 9b). The model progressively refined the estimate, ending at 22.9 mNm·s/rad (14.6% deviation; 95% CI: [19.3, 26.6]), which is reasonably close to the reference of 20.0.

The evolution of the Distance parameter ( $D$ ) is illustrated in Figure 9c. Early trials explored higher distance values (peaking near 0.30 rad, a 50% deviation). However, after Trial 4, the accumulating feedback drove the estimate back down toward the reference (0.20 rad). By Trial 13, the distance had converged to 0.26 rad (30.0% deviation; 95% CI: [0.22, 0.30]).



**Figure 9:** Evolution of best parameter estimates for the KBD condition ( $n = 72$ ). Stiffness ( $K$ ) stabilizes early (11.2% error); Damping ( $B$ ) shows continuous refinement (14.6% error); and Distance ( $D$ ) corrects from a large initial deviation down to 30% error.

#### 4.1.2 Effect of Coactive Feedback on Learning Efficiency

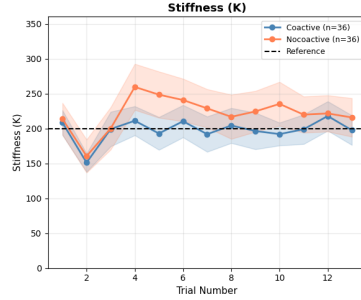
In order to determine whether coactive feedback accelerates learning, the model's best estimate trajectories between the Coactive ( $n = 36$ ) and Non-Coactive ( $n = 36$ ) groups were compared.

##### 4.1.2.1 K Condition

Both groups approached the reference stiffness ( $K_{ref} = 200$  mNm/rad), although with differing levels of accuracy and precision (Figure 10). The Coactive group achieved a final estimate of 197.5 mNm/rad (95% CI: [176.3, 218.7]). This value is within 1.2% of the reference, demonstrating high accuracy. In contrast, the Non-Coactive group showed a learning curve that began at approximately 259.5 mNm/rad and stabilized at 215.8 mNm/rad (95% CI: [188.2, 243.4]), corresponding to a 7.9% deviation from the target.

It is notable that the Coactive group produced narrower confidence intervals compared to the Non-Coactive group (Trial 13 Width: 42.4 vs. 55.2). This suggests that coactive feedback not only improves accuracy but also consistency, reducing the variability between participants by actively guiding them toward the solution space.

Parameter Estimation Across Trials: Coactive vs. Non-Coactive Groups (Mean  $\pm$  95% CI)



**Figure 10:** Evolution of best stiffness estimate by feedback group (K condition). The Coactive group (blue) converges to the reference (200 mNm/rad) with high precision (1.2% error), while the Non-Coactive group (orange) stabilizes at a value further from the reference (215.8 mNm/rad, 7.9% error) and shows wider confidence intervals.

#### 4.1.2.2 KB Condition

At intermediate complexity, the benefits of coactive feedback were not observed. In fact, the Non-Coactive group performed better in the damping dimension (Figure 11).

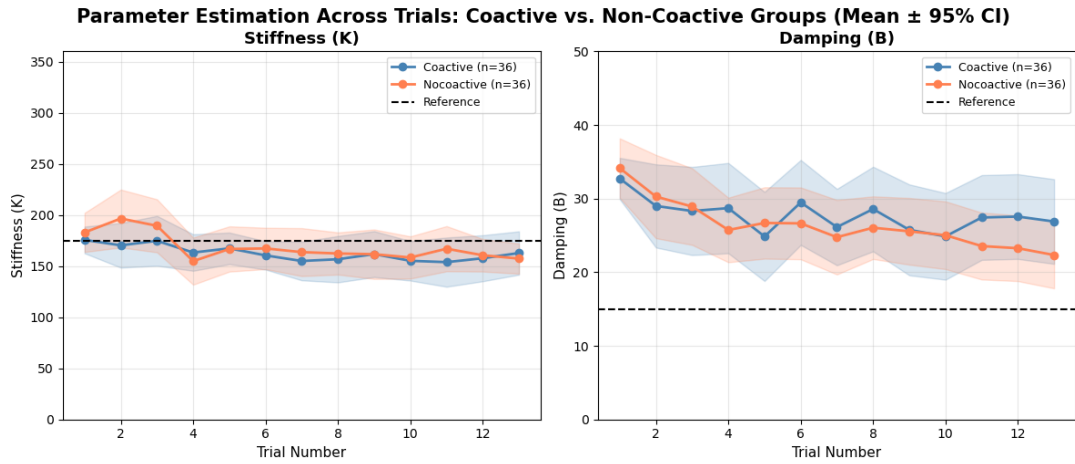
For Stiffness ( $K_{ref} = 175$  mNm/rad), the performance was similar between groups. The Coactive group ended with a mean of 162.8 mNm/rad (7.0% deviation; 95% CI: [141.6, 184.0]), while the Non-Coactive group converged to 157.3 mNm/rad (10.1% deviation; 95% CI: [142.6, 172.1]). Both groups exhibited wide confidence intervals that overlapped significantly throughout the trials.

For Damping ( $B_{ref} = 15$  mNm-s/rad), the groups tracked similarly until Trial 10, after which they diverged. The Non-Coactive group continued to improve, stabilizing closer to the reference at 22.3 mNm-s/rad (48.7% deviation; 95% CI: [17.8, 26.9]). In contrast, the Coactive group converged to a higher value of 26.9 mNm-s/rad (79.3% deviation; 95% CI: [21.1, 32.6]).

These results suggest that evaluating two parameters simultaneously may introduce cognitive load that reduces the effectiveness of directional feedback. Participants likely focused their coactive feedback on Stiffness, as it is the more perceptually dominant dimension.

#### 4.1.2.3 KBD Condition

In the three-parameter condition, the primary benefit of coactive feedback was its ability to correct the Distance estimate, which the Non-Coactive group failed to resolve. As shown in Figure 12, the groups diverged most notably in this parameter ( $D_{ref} = 0.20$  rad). The Non-Coactive group settled at a mean of 0.30 rad (50.0% deviation; 95% CI: [0.24, 0.36]), failing to converge to the target. In contrast, the Coactive group successfully corrected the estimate, ultimately reaching 0.22 rad (10.0% deviation; 95% CI: [0.17, 0.28]), demonstrating a significant improvement in accuracy.



**Figure 11:** Evolution of best parameter estimates for the KB condition. For Stiffness (Left), both groups show similar convergence with wide, overlapping confidence intervals. For Damping (Right), the Non-Coactive group achieves a better final estimate closer to the reference (48.7% vs 79.3% deviation) compared to the Coactive group. Shaded regions represent 95% confidence intervals.

However, this advantage did not extend to the other parameters. For Damping ( $B_{ref} = 20$  mNm·s/rad), there was no significant difference between the groups by the final trial. The Coactive group ended with a mean of 21.8 mNm·s/rad (9.0% deviation; 95% CI: [16.1, 27.5]), while the Non-Coactive group ended at 24.0 mNm·s/rad (20.0% deviation; 95% CI: [19.3, 28.8]). The broad overlap in confidence intervals indicates that coactive feedback did not significantly improve convergence precision for damping in this condition.

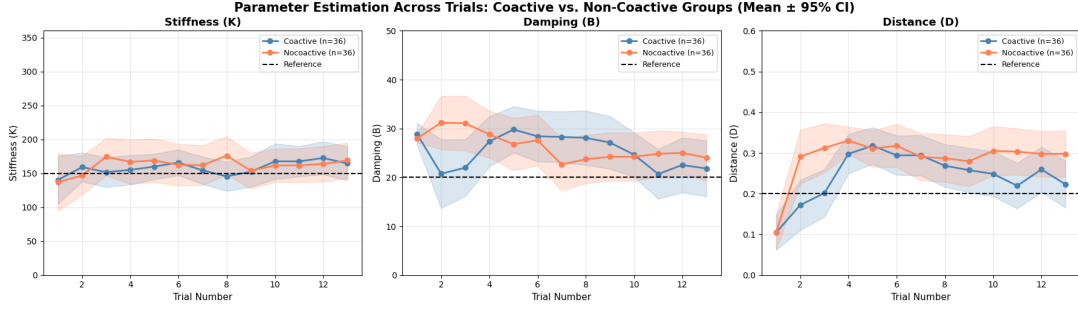
Regarding Stiffness ( $K_{ref} = 150$  mNm/rad), the performance was also similar between the groups. Both groups converged slightly above the reference value (Coactive: 164.6 mNm/rad, 9.7% deviation, 95% CI: [139.5, 189.7]; Non-Coactive: 168.9 mNm/rad, 12.6% deviation, 95% CI: [142.8, 195.0]).

These results indicate that as dimensionality increases, the primary value of coactive feedback is its ability to rescue the algorithm from suboptimal regions in the most difficult dimension (Distance), even if it does not refine the estimates for all parameters simultaneously.

## 4.2 Participant Behavior Analysis

### 4.2.1 Sensitivity to Torque Cues

In order to investigate whether participants rely on torque feedback decisions, peak torque from each option's final rotation was extracted and the Euclidean distance between each option's torque profile and the reference was computed.



**Figure 12:** Evolution of best parameter estimates by feedback group (KBD condition). The Coactive group (Blue) clearly corrects the Distance (Right) estimate (10% error vs 50% error). However, for Stiffness (Left) and Damping (Center), both groups show similar final convergence values with overlapping 95% confidence intervals.

A Bayesian logistic regression, which predicts choice as a function of normalized torque distances is fit:

$$P(\text{Choose Left}) = \text{logit}^{-1}(\alpha_p \cdot \sigma + \beta_L \cdot z_L + \beta_R \cdot z_R) \quad (19)$$

where  $z_L$  and  $z_R$  denote the normalized torque distances of the Left and Right options, respectively, and  $\alpha_p$  captures participant differences.

In Table 8, posterior estimates are reported. Both coefficients were reliably different from zero:  $\beta_L = -0.90$  (95% HDI:  $[-1.00, -0.79]$ ) and  $\beta_R = 1.09$  (95% HDI:  $[0.98, 1.19]$ ).

**Table 8:** Posterior estimates for the influence of torque distance on choice probability.

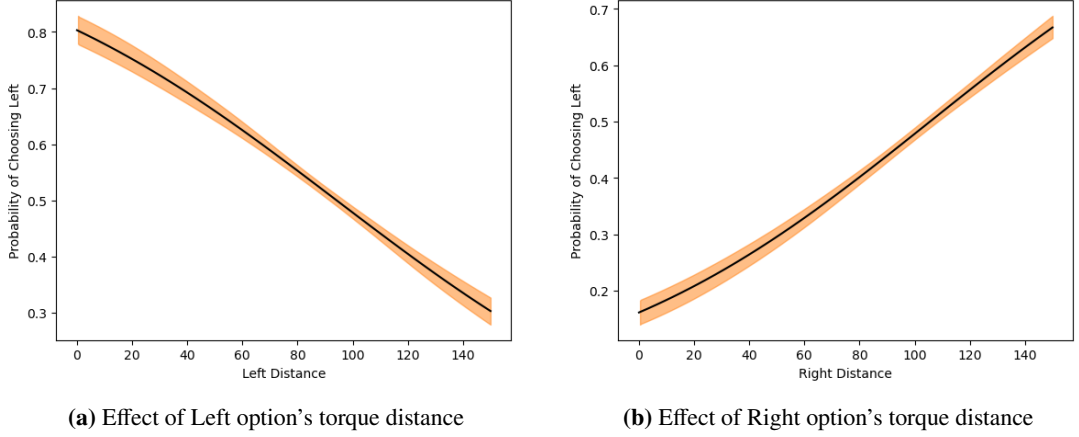
Parameter	Mean	SD	95% HDI
$\beta_{\text{Left}}$	-0.899	0.054	$[-1.00, -0.79]$
$\beta_{\text{Right}}$	1.090	0.054	$[0.98, 1.19]$

As the Left option’s torque deviates from the reference, participants become less likely to choose it. As the Right option deviates, participants reject it in favor of Left. Figure 13 visualizes these effects. Figure 13a indicates the impact of varying the Left option’s distance while the Right option is fixed, whereas Figure 13b shows the impact of varying the Right option’s distance while the Left option remains fixed. Selection probability drops from approximately 80% to 30% as Left Distance increases, and rises symmetrically as Right Distance increases. Therefore, it can be understood that participants’ decisions were grounded in torque similarity rather than arbitrary preferences.

## 4.2.2 Physical Determinants of Coactive Feedback

### 4.2.2.1 Influence of Stiffness, Damping, and Distance

In order to validate that coactive feedback (“Harder”/“Softer”) reflected physical parameters, feedback was modeled as a function of stiffness, damping, and distance using hierarchical



**Figure 13:** Posterior predictive probability of choosing the Left option as a function of torque distance. (a) As the Left option's distance from the reference increases (with Right fixed at 50), selection probability decreases. (b) As the Right option's distance increases (with Left fixed at 50), participants reject it and choose Left. Shaded regions indicate 95% HDI.

Bayesian logistic regression:

$$P(\text{Hard}) = \text{logit}^{-1}(\alpha_p + \beta_K z_K + \beta_B z_B \cdot I_B + \beta_D z_D \cdot I_D) \quad (20)$$

where  $z_K$ ,  $z_B$ , and  $z_D$  are normalized parameter values of selected option, and  $I_B$  and  $I_D$  are indicator variables for conditions that include Damping and Distance, respectively. Detailed specifications of the Bayesian model and its implementation are provided in Appendix ??.

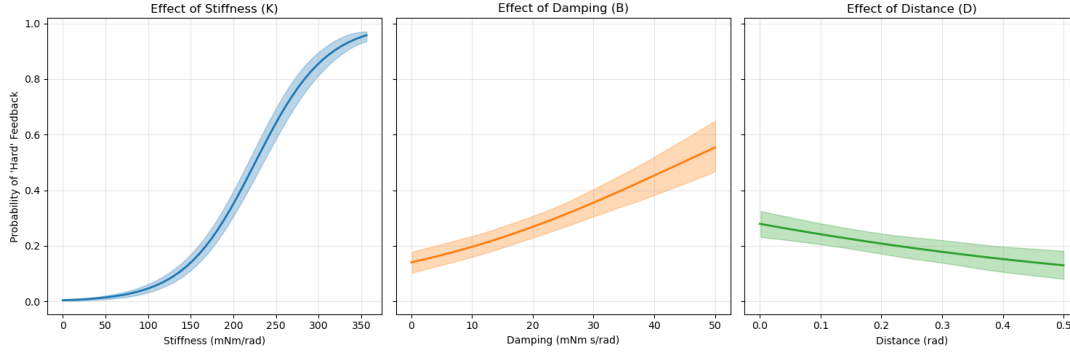
Table 9 summarizes the results. Stiffness was the dominant predictor ( $\beta_K = 1.50$ ), consistent with its physical definition. Damping also showed a positive effect ( $\beta_B = 0.69$ ). Although damping resists velocity rather than displacement, participants perceived high-damping environments as “harder.” This finding suggests that perceived hardness integrates total resistive force, not stiffness alone.

**Table 9:** Posterior estimates for the influence of haptic parameters on “Hard” feedback.

Parameter	Mean	SD	95% HDI
$\beta_K$ (Stiffness)	1.502	0.092	[1.32, 1.67]
$\beta_B$ (Damping)	0.690	0.100	[0.50, 0.88]
$\beta_D$ (Distance)	-0.301	0.089	[-0.47, -0.14]

Distance showed the opposite pattern ( $\beta_D = -0.30$ ). A larger distance threshold, which creates a “dead zone” before force onset, reduced the probability of “Hard” feedback. The presence of free motion apparently creates an impression of softness, even when subsequent stiffness is high.

Figure 14 presents the posterior predictive curves for each parameter. Stiffness exhibits the strongest influence, spanning nearly the full probability range (0% to 95%). As stiffness increases, the probability of providing “Hard” feedback rises sharply. Damping also shows a



**Figure 14:** Posterior predictive probability of “Hard” feedback as a function of each haptic parameter: (a) Stiffness, (b) Damping, and (c) Distance. Stiffness shows a full sigmoid transition across its range, while Distance shows a negative trend. Shaded regions indicate 95% HDI.

positive relationship. However, the effect is more moderate compared to stiffness, reaching a maximum probability of approximately 55%. Conversely, increasing the distance parameter leads to a slight decrease in the probability of reporting “Hard,” suggesting that larger gaps before force onset soften the perceived interaction.

#### 4.2.2.2 Torque Distance Effect

Participants’ feedback might be contaminated by the rejected option. They could be responding to the contrast between options rather than the absolute properties of their chosen option. In order to test this, the “Hard” feedback, as a function of both the Selected and Rejected options’ torque distances, was modeled:

$$P(\text{Hard}) = \text{logit}^{-1}(\alpha_p + \beta_{\text{sel}}z_{\text{sel}} + \beta_{\text{oth}}z_{\text{oth}} + \beta_{\text{sq}}z_{\text{oth}}^2 + \beta_{\text{cond}}) \quad (21)$$

where  $z_{\text{sel}}$  and  $z_{\text{oth}}$  are the normalized torque distances of the selected and rejected options, respectively. The quadratic term checks for non-linear distraction effects. Detailed specifications of the Bayesian model and its implementation are provided in Appendix ??

Table 10 shows that only the Selected option’s distance predicted feedback ( $\beta_{\text{sel}} = 0.76$ , 95% HDI: [0.61, 0.92]). The Rejected option had no credible effect ( $\beta_{\text{oth}} = -0.02$ , 95% HDI: [-0.17, 0.13]).

This confirms that participants successfully isolated the sensation of their chosen option while providing feedback. The coactive signal was specific and not confounded by the comparison context.

#### 4.2.3 Final Estimate Distributions

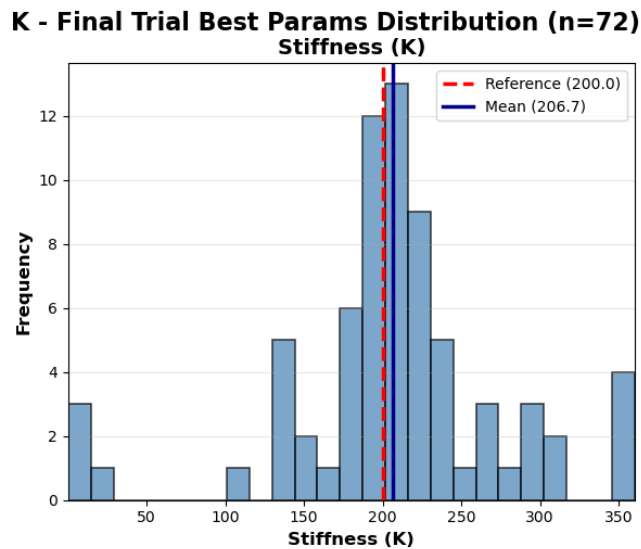
Until this section, the convergence trajectories show how estimates evolved over trials. It can be also examined that the distributions of final estimates across participants, since it will reveal

**Table 10:** Posterior estimates for feedback specificity: influence of selected vs. rejected option.

Parameter	Mean	SD	95% HDI
$\beta_{\text{Selected}}$	0.762	0.084	[0.61, 0.92]
$\beta_{\text{Other}}$	-0.015	0.081	[-0.17, 0.13]
<i>Condition Baselines</i>			
$\beta_K$	-0.533	0.134	[-0.78, -0.28]
$\beta_{KB}$	-0.540	0.133	[-0.79, -0.30]
$\beta_{KBD}$	-0.561	0.130	[-0.80, -0.31]

the precision, accuracy and potential biases of the learning algorithm. Figures 15–17 present histograms of each participant’s best estimate at Trial 13.

#### 4.2.3.1 K Condition

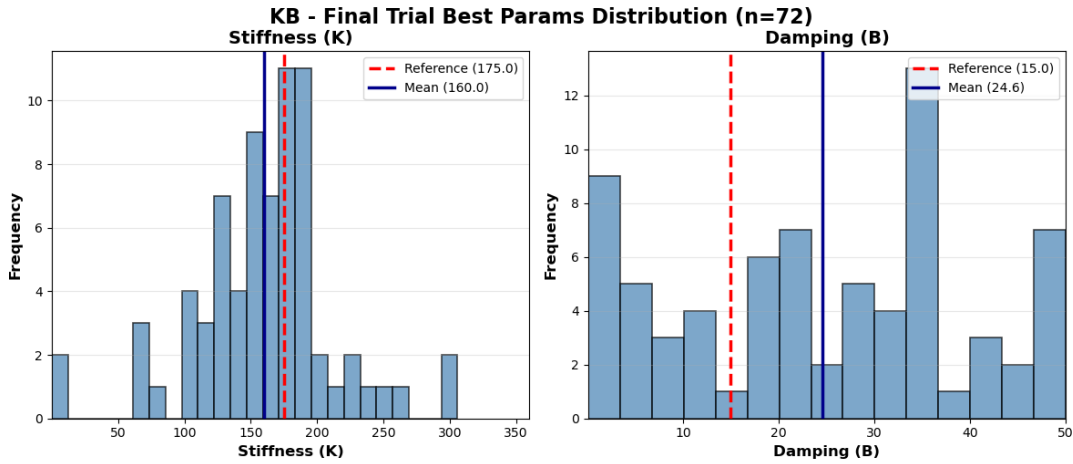


**Figure 15:** Distribution of final stiffness estimates across participants (K condition,  $n = 72$ ). The dashed red line indicates the reference value (200.0 mNm/rad); the solid blue line indicates the sample mean (206.7 mNm/rad).

Figure 15 shows the distribution of final stiffness estimates. The distribution is centered near the reference, with a mean of  $206.66 \pm 72.26$  mNm/rad (reference: 200.0), corresponding to a mean error of only +3.3%. Most participants converged to estimates within the 175–250 mNm/rad range.

However, a small number of outliers remained at extreme values. These deviations may reflect individual differences in perceptual sensitivity or potential artifacts in the models best estimates calculation method, where high predicted utility at the boundaries could arise from inconsistent preference data.

### 4.2.3.2 KB Condition



**Figure 16:** Distribution of final parameter estimates across participants (KB condition,  $n = 72$ ). Left: Stiffness estimates cluster successfully near the reference (175.0 mNm/rad). Right: Damping shows a widely dispersed distribution rather than a clear peak at the reference (15.0 mNm·s/rad), with a notable cluster of participants converging to high values around 34 mNm·s/rad.

In the two-parameter condition (Figure 16), stiffness estimates were distributed with a mean of  $160.04 \pm 53.28$  mNm/rad, which is approximately 8.5% lower than the reference (175.0). As shown in the histogram, the majority of participants successfully estimated the stiffness, with the peak of the distribution aligning closely with the reference value.

In contrast, the damping estimates showed a much wider spread, with a mean of  $24.61 \pm 15.24$  mNm·s/rad (reference: 15.0). Unlike stiffness, the damping estimates were scattered across the entire range rather than converging to a single peak. While there is a notable cluster of participants (approximately 12 individuals) who converged to a high value around 34 mNm·s/rad, the lack of a clear central tendency suggests that participants found it difficult to isolate the damping parameter precisely.

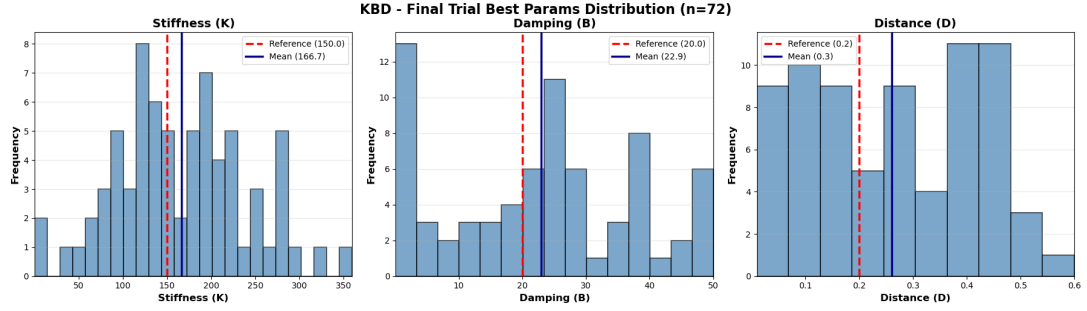
### 4.2.3.3 KBD Condition

The three-parameter condition showed the highest variability, as expected given the expanded search space. Figure 17 presents the distributions for all three parameters.

Stiffness estimates were broadly distributed with a mean of  $166.74 \pm 74.63$  mNm/rad (reference: 150.0), which is 11.2% higher than the reference. The distribution exhibits clustering in two regions (100–150 mNm/rad and 175–225 mNm/rad), suggesting that participants may have settled into distinct perceptual strategies.

Damping estimates showed reasonable precision, with a mean of  $22.92 \pm 15.37$  mNm·s/rad (reference: 20.0), corresponding to a 14.6% error. The distribution peaks sharply in the 20–25 mNm·s/rad range, indicating that most participants successfully identified values near the reference.

Distance estimates showed the highest variability among all parameters, with a mean of  $0.26 \pm 0.17$  rad (reference: 0.20 rad). Rather than converging to a single peak, the estimates were widely dispersed across the entire search space. While a considerable number of participants identified values near the reference (0.1–0.2 rad), others converged to much higher values (0.4–0.5 rad). This lack of consensus suggests that the distance parameter was difficult for participants to perceive consistently.



**Figure 17:** Distribution of final parameter estimates across participants (KBD condition,  $n = 72$ ). Left: Stiffness (reference: 150.0 mNm/rad, mean: 166.7 mNm/rad). Center: Damping (reference: 20.0 mNm-s/rad, mean: 22.9 mNm-s/rad). Right: Distance (reference: 0.20 rad, mean: 0.26 rad). Variability increases with dimensionality, particularly for the Distance parameter.

#### 4.2.3.4 Summary of Final Estimates

Table 11 summarizes the final estimate statistics across all conditions.

**Table 11:** Summary statistics of final parameter estimates at Trial 13 ( $n = 72$ ).

Condition	Parameter	Reference	Mean	SD	Error (%)
K	Stiffness (mNm/rad)	200.0	206.66	72.26	+3.3
KB	Stiffness (mNm/rad)	175.0	160.04	53.28	-8.5
	Damping (mNm-s/rad)	15.0	24.61	15.24	+64.1
KBD	Stiffness (mNm/rad)	150.0	166.74	74.63	+11.2
	Damping (mNm-s/rad)	20.0	22.92	15.37	+14.6
	Distance (rad)	0.20	0.26	0.17	+30.0



## CHAPTER 5

### DISCUSSION

#### 5.1 Validation of the Computational Framework

This study investigated whether the given a Gaussian Process could learn haptic parameter values only from pairwise comparisons. The results indicate that the proposed model successfully estimates the latent preference function. A critical requirement for the application was real-time feasibility. Especially, the GP model needed to update the posterior distribution and select the next query pair without disrupting the user’s experiment flow.

Precision and accuracy were highest in low-dimensional condition (K), where the model reliably approached the target values. By aggregating best estimates across participants, it can be concluded that the system generally converged close to the reference values. This suggests that, for applications requiring one, two or three degrees of freedom, the current GP framework provides a viable method for automating haptic tuning without requiring expert manual intervention.

#### 5.2 Perceptual Sensitivity and Dimensionality

The performance of the model varied across conditions, highlighting the relationship between algorithmic search capacity and human perceptual sensitivity.

##### 5.2.1 The Stiffness (K) Condition

The system achieved its highest precision and accuracy in the single-parameter condition. The final estimated stiffness was 206.66 mNm/rad against a reference of 200.0 mNm/rad. This deviation represents an error of only 3.3%, which falls well below the reported Just Noticeable Difference (JND) thresholds for kinesthetic stiffness. This implies that mathematical convergence was almost exact. Since stiffness is a static, position-dependent property, it presents a stable signal for both the human sensory system and the GP model.

### 5.2.2 The Stiffness-Damping (KB) Condition

"In the two-parameter condition, the convergence was significantly less precise for the dynamic parameter. While the stiffness estimate remained relatively robust with only an 8.5% error (160.0 vs. Ref 175.0 mNm/rad), the damping estimate deviated largely from the target, showing a 64.1% overestimation (24.6 vs. Ref 15.0 mNm·s/rad). Two factors probably account for this pattern. First, damping depends on velocity, meaning the actual force varied with each participant's movement speed. Second, stiffness and damping tend to merge into a single percept of hardness.

Firstly, unlike stiffness, damping is a dynamic property. The physical force generated depends entirely on the participant's rotation speed. Since exploratory speeds vary naturally between participants, the "objective" force differed across users even for identical parameter settings.

Secondly, there is a known perceptual trade-off between stiffness and damping. High damping can mask the sharpness of a stiffness boundary, leading to "perceptual aliasing" where different mathematical combinations (e.g., High  $K$ /Low  $B$  vs. Moderate  $K$ /High  $B$ ) generate similar overall sensations of "hardness." For the optimization algorithm, this perceptual overlap creates difficulty: two parameter combinations that are far apart mathematically can feel nearly identical to the user.

### 5.2.3 The Stiffness-Damping-Distance (KBD) Condition

The KBD condition represented the highest level of complexity, expanding the search space to three dimensions. Despite this increased difficulty, the model successfully recovered the Stiffness and Damping parameters with reasonable precision. The mean stiffness (166.7 mNm/rad) and damping (22.9 mNm·s/rad) estimates remained close to their respective references (150.0 and 20.0), suggesting that the core haptic sensations remained distinguishable even in a high-dimensional context.

"However, the Distance parameter ( $D$ ) showed the highest inconsistency across participants. The estimated mean of 0.26 rad deviated from the reference (0.20 rad) by 30%. More importantly, the standard deviation for distance was 0.17, which is notably high relative to the mean. This indicates a lack of consensus among participants compared to the tighter clustering observed in stiffness. This deviation can be attributed to the specific perceptual nature of the distance parameter, which creates a gap before the force is felt. Accurately detecting exactly where this gap ends requires the user to have a very precise sense of their hand position, which is challenging during rapid movements. In such cases, participants likely moved through the initial free-motion zone too quickly to accurately detect the transition point. Moreover, their attention was likely dominated by the sensation of the virtual wall at the end of the motion. As a result, without force feedback in that initial range, participants had little sensory information to guide their judgment, which likely explains the higher variance relative to ( $K$  and  $B$ ).

### 5.3 Sensory Basis of Decision Making

A fundamental assumption of the proposed human-in-the-loop framework is that participants can perceive and minimize the difference between the rendered haptic stimuli and the reference. Torque sensitivity analysis indicates that participants based their decisions on actual physical differences rather than guessing or random preference.

The Bayesian regression results show a connection between torque distance and selection probability. The negative coefficient for the Left option ( $\beta_{Left} = -0.90$ ) indicates that as the torque of the Left option deviated further from the reference torque, the likelihood of selecting it dropped significantly. In contrast, the positive coefficient for the Right option ( $\beta_{Right} = 1.09$ ) indicates that as the Right option deviated further from the reference, participants were more likely to reject it in favor of the Left option.

The symmetry in these coefficients suggests participants treated torque magnitude as their primary cue for judging similarity. Even though the underlying parameters ( $K, B, D$ ) were hidden from participants, the resulting force feedback provided a sufficient sensation for discrimination. This consistency is critical for the Gaussian Process model since it confirms that the human user effectively approximates a consistent objective function (minimizing torque difference).

### 5.4 The Impact of Coactive Feedback

The introduction of coactive feedback provided an important mechanism for the learning process. A key observation from the experiments was the tendency of the standard pairwise approach (without coactive feedback) to suffer from early convergence to local optima. When a participant repeatedly selects the same option across multiple trials, the Gaussian Process interprets this consistency as high certainty in that area. Thereby, the acquisition function begins to exploit this region and generates new queries clustered around the current “best” option. This creates a loop: the model assumes it has found the target and stops exploring, trapping the user in a local optimum even if the current setting is not actually close to the reference.

Coactive feedback effectively breaks this cycle. In the standard approach, if a user picks an option, the model typically retains it for the next trial. However, with coactive feedback, the user can modify the preferred option by specifying a direction, such as “make it harder.” The system processes this feedback and calculates a new parameter set shifted in that direction. The modified option then appears in the next trial instead of the last chosen option. This ensures that the comparison evolves: instead of repeatedly seeing the same static options, the participant is constantly presented with a potentially better alternative. This forces the algorithm to keep moving toward the reference rather than settling on a value that is simply “good enough.”

To validate the reliability of the coactive feedback, it is necessary to determine whether participants were reacting just to the selected option or if they were influenced by the rejected alternative. A potential concern in pairwise comparisons is the “contrast effect,” where a par-

participant’s judgment is biased by the properties of the rejected option. However, the regression analysis uncovered that coactive feedback was driven only by the physics of the selected option ( $\beta_{selected} = 0.76$ ), with no credible influence from the rejected option ( $\beta_{other} \approx 0$ ). This independence matters because it means that participants were making an absolute judgment about the chosen option, not a relative reaction to the previous comparison. This ensures that the active learning algorithm receives consistent, targeted constraints rather than noisy, context-dependent data.

While coactive feedback provided a mechanism for directional guidance, its impact on the final convergence point varied across conditions. It did not guarantee a higher final accuracy compared to the standard pairwise approach, but rather influenced the *rate* and *robustness* of the learning process.

#### 5.4.1 Convergence Rate in Low Dimensions (K)

In the single-parameter stiffness condition, coactive feedback improved both the convergence rate and the final accuracy. As detailed in the final trial statistics, the coactive group achieved a near-perfect estimation with only 1.2% error (197.5 vs. Target 200.0), whereas the non-coactive group settled at a 7.9% error (215.8 vs. Target 200.0). As illustrated in Figure 10, the coactive group estimated the reference in fewer trials. Visual inspection of convergence trends suggests the coactive group stabilized near the target by approximately Trial 5, compared to Trial 13 for the non-coactive group. Even if standard pairwise comparisons appear sufficient for eventual convergence, directional feedback significantly tightens the final bounds by ruling out large portions of the search space.

#### 5.4.2 Efficacy in Higher Dimensions and Feedback Ambiguity

In the multi-parameter conditions (KB and KBD), the effectiveness of coactive feedback was constrained by the ambiguity of the participant’s signal relative to the dimensionality of the search space (See Figures 11, 12). In a single-parameter setting, a request for “Harder” feedback maps directly to an increase in stiffness. However, in higher dimensions, the source of the sensation becomes ambiguous.

This is most evident in the KB condition, where the coactive group actually performed worse than the non-coactive group on damping (79.2% error vs. 48.9% error). This suggests that when users requested “Harder,” the algorithm aggressively increased damping to satisfy the request, leading to over-shooting. This interpretation is supported by the analysis of feedback determinants in Section 4.2.2.1, which reveals that the probability of providing “Hard” feedback is primarily driven by stiffness ( $K$ ). Although there is a positive relationship between the perception of “Hardness” and damping ( $B$ ), the relationship is significantly stronger for stiffness. However, the coactive feedback algorithm modifies both stiffness (7%) and damping (10%) simultaneously. Consequently, the system effectively “hallucinates” a damping deficiency when the user is likely only responding to insufficient stiffness, creating a preference ambiguity that degrades the damping estimate.

Conversely, in the KBD condition, coactive feedback proved highly effective for the geometric parameter. The non-coactive group struggled significantly with Distance (48.9% error), while the coactive group converged to within 11.5% of the reference (0.22 vs 0.20 rad). This suggests that the presence of the dead-zone ( $D$ ) interacts with the ambiguity seen in stiffness and damping. As noted in Section 4.2.2.1, there is a negative relationship between distance and the “Hard” feedback label; increasing the distance makes the haptic setting feel “Softer.”

As a result, there is an inherent ambiguity arising from the multi-dimensionality of the sensation versus the one-dimensionality of the coactive feedback. When a participant perceives a stimulus as “softer” than the reference, this sensation could arise from insufficient stiffness ( $K$ ), insufficient damping ( $B$ ), or the presence of a large dead zone ( $D$ ). While the algorithm successfully used the “Harder” command to reduce the Distance error, the underlying ambiguity regarding which specific parameter caused the sensation remains a constraint of scalar feedback.

Ultimately, coactive feedback helped with early exploration and prevented the algorithm from getting stuck in local optima, but the ambiguity of scalar feedback constrained how precisely the system could converge in later trials. The system effectively learned to approximate the reference intensity, but the scalar nature of the feedback (“Harder/Softer”) was insufficient to fully resolve the specific contributions of each parameter in the complex KBD space.



## CHAPTER 6

### CONCLUSION AND FUTURE WORK

#### 6.1 Limitations

##### 6.1.1 Artificiality of the Haptic Stimuli

One of the limitations of this study was the use of a simulated “virtual” reference rather than a physical object. Using a mathematical reference guaranteed that the stimulus never changed. On the other hand, physical objects which can degrade over time. However, this consistency came at the cost of realism, making the reference feel artificial. Participants reported that the reference signal did not clearly resemble everyday objects (e.g., a sponge or a spring) even if it is noticeable. Participants had to memorize an abstract force profile rather than drawing on intuition about familiar materials, which may have added cognitive load. Future studies could validate these findings using physical benchmarks to assess how the algorithm performs when matching real-world textures.

##### 6.1.2 Impact of Device’s Inner Friction

Some mechanical friction in the device was unavoidable despite the overall fidelity of the force feedback. While the magnitude of this friction was low, it acts as a constant background force that can mask subtle changes in the damping parameter ( $B$ ). Since damping is a velocity-dependent force, distinguishing low-damping settings from the device’s intrinsic friction requires high sensitivity. Friction may partly explain why damping estimates showed higher variance than stiffness estimates, since stiffness is position-dependent and therefore less susceptible to friction.

##### 6.1.3 Participant Expertise

The study relied on a general population of university students who were largely novices in haptics and mechanics. As a result, coactive feedback had to remain at a semantic level (“Harder” or “Softer”) rather than targeting specific parameters. If the participant pool had been stratified to include domain experts (e.g., mechanical engineering students familiar with system dynamics), the feedback mechanism could have been expanded to targeted parameter

tuning (e.g., “Increase Stiffness,” “Decrease Damping”). Recruiting novices showed that the system generalizes beyond expert users, but it also meant noisier feedback since these participants found it difficult to separate the physical components of what they felt.

#### 6.1.4 Reference Centering and Initialization Bias

The experimental design imposed specific constraints on the parameter space to ensure feasibility within a single session. Specifically, the reference values were deliberately selected from the center of the search space (e.g.,  $K = 200$  in a range of  $0 - 360$ ). This decision was made to avoid “ceiling effects” where extreme parameters might result in sensations that were either imperceptibly soft or uncomfortably hard, thereby degrading the depth of the sensory feedback.

However, this centrality introduces a potential bias when combined with the fixed initialization phase. The first three trials were designed to form a uniform grid across the search space to ensure broad initial coverage. Consequently, the average of these initial inputs naturally gravitates toward the geometric center of the domain. Since the target reference was also located in the center, the Gaussian Process’s initial estimate converged close to the target simply as a statistical artifact of this uniform sampling, rather than through learning.

This effect is observable in the results: the model’s error often appears artificially low during the fixed initialization (Trials 1–3). Once the active learning algorithm takes over (Trial 4+), the error temporarily increases as the acquisition function begins to explore the boundaries of the space, moving the estimate away from the safe “middle” before eventually converging back to the true target. Future studies should test reference values located near the boundaries of the parameter space to verify that the model’s convergence is not reliant on this central tendency.

## 6.2 Future Work

### 6.2.1 Ensuring Perceptual Differences

An analysis of the query logs revealed a sub-optimal behavior in the acquisition function. The current implementation penalized points close to the *current best estimate* to encourage exploration. However, it did not explicitly penalize points close to the immediately preceding query. As a result, the system occasionally generated options that were mathematically distinct but perceptually indistinguishable (falling below the Just Noticeable Difference, or JND) from the previous trial. This resulted in “wasted” queries where the user could not perceive a difference. A future version of the acquisition function could include a repulsion term tied to the JND so that every pair differs enough for participants to perceive a real contrast.

### 6.2.2 Beyond Binary Feedback

The current coactive feedback system was binary (“Make it Harder” / “Make it Softer”). Participant behavior hinted that people often had a sense of how far off the stimulus was, not just which direction. Future work should integrate scalar feedback options (e.g., “Slightly Harder” vs. “Much Harder”). Gradient information of this kind could inform step size: large jumps when the error feels substantial, smaller adjustments when the sensation is close to the target.

### 6.2.3 Gamification and Attentional Maintenance

Sessions lasted approximately 25 minutes which were short enough to avoid physical fatigue. However, some participants still showed signs of wandering attention as trials accumulated, likely due to the repetitive nature of the task. Since there is no “correct” answer in a subjective preference task, participants lacked explicit objective and performance feedback. To maintain high engagement levels, future experiments could introduce gamification elements. For example, providing a monetary incentive proportional to the consistency of their choices (verified against a hidden ground truth) might help participants stay engaged, potentially reducing the noise in the final convergence data.

### 6.2.4 Evaluation of Higher-Order Query Types

The current framework relied strictly on pairwise comparisons to minimize the working memory load on participants. However, this approach is limited by low information density (1 bit per trial), which necessitates a larger number of queries for convergence. Future research should investigate the feasibility of higher-order query types, such as *List Ranking* or *Best-of-K Selection*. While ranking large sets is cognitively prohibitive in haptics due to sensory decay, a “micro-ranking” approach (e.g., ordering  $k = 3$  stimuli) could theoretically increase the information gain per trial (from 1 bit to  $\approx 2.58$  bits) without exceeding short-term memory limits.

## 6.3 Conclusion

This thesis introduced a human-in-the-loop framework for automating haptic interface calibration. Gaussian Processes combined with Active Learning enabled the system to learn user preferences from only subjective pairwise comparisons. This approach effectively removes the need for manual parameter tuning by domain experts. Instead, the framework relies on a probabilistic model to interpret user intent.

The experimental results indicate that the proposed framework is effective across different levels of complexity. In the low-dimensional Stiffness condition, the learned parameters matched the reference values with high precision and accuracy. In the more complex Stiffness-Damping (KB) and Stiffness-Damping-Distance (KBD) conditions, the search space was

significantly wider and more challenging. Despite this increased dimensionality, the model achieved reasonable convergence to the reference regions using only 13 pairwise comparisons. This demonstrates that the algorithm is computationally efficient and capable of handling multi-parameter optimization with a limited budget of user trials.

The inclusion of coactive feedback served as a valuable support mechanism for the model. It was observed that standard pairwise comparisons could occasionally lead to premature convergence, where the model settled on a local optimum. Coactive feedback addressed this by providing directional guidance, which prevented the algorithm from getting stuck in these loops. The pairwise model acted as the core learning engine, and the coactive feedback improved the robustness and speed of the search. It ensured that the system continued to explore relevant areas of the parameter space.

All in all, the findings indicate that subjective haptic experience, despite its inherent variability, can be captured mathematically and translated into device parameters. Automated calibration of this kind could allow haptic devices to adapt to individual users without requiring either technical expertise or tedious manual adjustment.

## 6.4 Summary of Contributions

This thesis presented a Bayesian Active Preference Learning framework to automate haptic tuning, bridging the gap between subjective human perception and objective machine control. The primary contributions of this work are summarized as follows:

Firstly, the proposed Gaussian Processes model can successfully translate qualitative, subjective feelings into precise mathematical control parameters ( $K$ ,  $B$ ,  $D$ ) without requiring users to articulate numerical values. Furthermore, the experiment validated that the active learning framework achieves high convergence accuracy within a limited budget of 13 trials, making it computationally and cognitively suitable for real-time applications.

Secondly, it is shown that Coactive Feedback (“Harder/Softer”) accelerates learning and prevents local optima. Furthermore, the “Hard” feedback is in a strong relationship with Stiffness parameter, has a positive relationship with Damping ( $B$ ) and has a negative relationship with distance ( $D$ ).

Thirdly, it is confirmed that user decisions are driven by physical torque cues rather than arbitrary preference.

Finally, this study developed a workflow that enables non-expert users to intuitively replicate natural haptic sensations, effectively removing the traditional reliance on domain experts for manual parameter adjustment.

## Bibliography

- Aggarwal, M. and Tehrani, A. F. (2019). Modelling human decision behaviour with preference learning. *INFORMS Journal on Computing*, 31(2):318–334.
- Aiguzhinov, A., Soares, C., and Serra, A. P. (2010). A similarity-based adaptation of naive bayes for label ranking: Application to the metalearning problem of algorithm recommendation. In Pfahringer, B., Holmes, G., and Hoffmann, A., editors, *Discovery Science*, pages 16–26. Springer.
- Anderson, D. R., Burnham, K. P., and Thompson, W. L. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *The Journal of Wildlife Management*, 64(4):912–923.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53:370–418.
- Beal, M., Ghahramani, Z., and Rasmussen, C. (2001). The infinite hidden markov model. In Dietterich, T., Becker, S., and Ghahramani, Z., editors, *Advances in Neural Information Processing Systems*, volume 14. MIT Press.
- Benavoli, A. and Azzimonti, D. (2024). A tutorial on learning from preferences and choices with gaussian processes. *arXiv preprint*.
- Berger, J. O. and Berry, D. A. (1988). Statistical analysis and the illusion of objectivity. *American Scientist*, 76(2):159–165.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New York, NY.
- Biyik, E., Huynh, N., Kochenderfer, M. J., and Sadigh, D. (2020). Active preference-based Gaussian process regression for reward learning. In *Proceedings of Robotics: Science and Systems (RSS)*, Corvallis, OR, USA.
- Blei, D. M., Kucukelbir, A., and McAuliffe, J. D. (2017). Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877.
- Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345.
- Brochu, E., Cora, V. M., and de Freitas, N. (2010). A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599*.
- Browder, J., Bochereau, S., van Beek, F., and King, R. (2019). Stiffness in virtual contact events: A non-parametric Bayesian approach. In *2019 IEEE World Haptics Conference (WHC)*, pages 313–318. IEEE.

- Chapelle, O. and Li, L. (2011). An empirical evaluation of Thompson sampling. In *Advances in Neural Information Processing Systems*, pages 2249–2257.
- Chin, R., Manzie, C., Ira, A., Nescic, D., and Shames, I. (2018). Gaussian processes with monotonicity constraints for preference learning from pairwise comparisons. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 1150–1155. IEEE.
- Chu, W. and Ghahramani, Z. (2005). Preference learning with gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning (ICML)*, pages 137–144. ACM.
- Coppersmith, D., Fleischer, L. K., and Rudra, A. (2010). Ordering by weighted number of wins gives a good ranking for weighted tournaments. *ACM Transactions on Algorithms*, 6(3):55:1–55:13.
- de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7(1):1–68.
- Dekel, O., Singer, Y., and Manning, C. D. (2003). Log-linear models for label ranking. In *Advances in Neural Information Processing Systems*, volume 16. MIT Press.
- Dembczyński, K., Kotłowski, W., Słowiński, R., and Szełąg, M. (2011). Learning of rule ensembles for multiple attribute ranking problems. In Fürnkranz, J. and Hüllermeier, E., editors, *Preference Learning*, pages 217–247. Springer, Berlin, Heidelberg.
- Fechner, G. T. (1860). *Elements of Psychophysics*. Breitkopf und Härtel, Leipzig. English translation: Holt, Rinehart and Winston (1966).
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd, Edinburgh.
- Fürnkranz, J. and Hüllermeier, E. (2010). Preference learning: An introduction. In Fürnkranz, J. and Hüllermeier, E., editors, *Preference Learning*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Analysis*, 3(3):445–450.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC Press, Boca Raton, FL, 3rd edition.
- Gessner, A., Kanjilal, O., and Hennig, P. (2020). Integrals over Gaussians under linear domain constraints. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 108 of *Proceedings of Machine Learning Research*, pages 2764–2774. PMLR.
- Gigiberia, N. (2025). Bayesian model selection with an application to cosmology. *arXiv preprint arXiv:2512.09724*.
- González, J., Dai, Z., Hennig, P., and Lawrence, N. (2016). Batch bayesian optimization via local penalization. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 648–657, Cadiz, Spain. PMLR.

- Har-Peled, S., Roth, D., and Zimak, D. (2002). Constraint classification: A new approach to multiclass classification. In Cesa-Bianchi, N., Numao, M., and Reischuk, R., editors, *Algorithmic Learning Theory*, pages 365–379. Springer.
- Hennig, P. and Schuler, C. J. (2012). Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13:1809–1837.
- Herbrich, R., Graepel, T., and Obermayer, K. (2000). Large margin rank boundaries for ordinal regression. In Smola, A. J., Bartlett, P. L., Schölkopf, B., and Schuurmans, D., editors, *Advances in Large Margin Classifiers*, pages 115–132. MIT Press, Cambridge, MA.
- Hu, Y., Li, M., and Yu, N. (2008). Multiple-instance ranking: Learning to rank images for image retrieval. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE.
- Hüllermeier, E. and Słowiński, R. (2024). Preference learning and multiple criteria decision aiding: differences, commonalities, and synergies – part i. *4OR*, 22(2):179–209.
- Jain, A., Wojcik, B., Joachims, T., and Saxena, A. (2013). Learning trajectory preferences for manipulators via iterative improvement. In *Advances in Neural Information Processing Systems*, pages 575–583.
- Joachims, T. (2002). Optimizing search engines using clickthrough data. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 133–142. ACM.
- Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492.
- Jones, L. A. and Hunter, I. W. (1990). A perceptual analysis of stiffness. *Experimental Brain Research*, 79(1):150–156.
- Jones, L. A. and Hunter, I. W. (1993). A perceptual analysis of viscosity. *Experimental Brain Research*, 94(2):343–351.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.
- Kapoor, A., Grauman, K., Urtasun, R., and Darrell, T. (2007). Active learning with Gaussian processes for object categorization. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE.
- Keeney, R. L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*. John Wiley & Sons, New York, NY.
- Kingsley, D. C. and Brown, T. C. (2010). Preference uncertainty, preference learning, and paired comparison experiments. *Land Economics*, 86(3):530–544.
- Knill, D. C. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719.
- Körding, K. P. and Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971):244–247.

- Kushner, H. J. (1964). A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86(1):97–106.
- Laming, D. (2004). Human judgment: The eye of the beholder. *Psychological Review*, 111(2):484–503.
- Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les évènements. In *Mémoires de Mathématique et de Physique, Présentés à l'Académie Royale des Sciences*, volume 6, pages 621–656.
- Lederman, S. J. and Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71(7):1439–1459.
- Liu, Q., Crispino, M., Scheel, I., Vitelli, V., and Frigessi, A. (2019). Model-based learning from preference data. *Annual Review of Statistics and Its Application*, 6(1):329–354.
- Lizotte, D., Wang, T., Bowling, M., and Schuurmans, D. (2007). Automatic gait optimization with Gaussian process regression. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*, volume 7, pages 944–949.
- Luce, R. D. (1959). *Individual Choice Behavior*. John Wiley & Sons, New York, NY.
- Mallows, C. L. (1957). Non-null ranking models. i. *Biometrika*, 44(1–2):114–130.
- Martinez-Cantin, R., de Freitas, N., Brochu, E., Castellanos, J., and Doucet, A. (2009). A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot. *Autonomous Robots*, 27(2):93–103.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2):81–97.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Adaptive Computation and Machine Learning Series. MIT Press, Cambridge, MA.
- Neyman, J. and Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337.
- Ng, A. Y. and Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In *Proceedings of the Seventeenth International Conference on Machine Learning (ICML '00)*, pages 663–670, San Francisco, CA. Morgan Kaufmann.
- Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab. Previous version available as technical report SIDL-WP-1999-0120.
- Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, 72(6):407–418.
- Pascual, C. B. and Jain, S. (2025). Platform-of-1: A bayesian adaptive n-of-1 trial design for identifying an optimal treatment among multiple candidates. *Statistics in Biopharmaceutical Research*, pages 1–12.

- Payne, J. W., Bettman, J. R., and Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press, Cambridge, UK.
- Pfannschmidt, K., Gupta, P., and Hüllermeier, E. (2018). Deep architectures for learning context-dependent ranking functions. *arXiv preprint arXiv:1803.05796*.
- Plackett, R. L. (1975). The analysis of permutations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 24(2):193–202.
- Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning Series. MIT Press, Cambridge, MA.
- Resnick, P. and Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, 40(3):56–58.
- Sadigh, D., Dragan, A. D., Sastry, S., and Seshia, S. A. (2017). Active preference-based learning of reward functions. In *Proceedings of Robotics: Science and Systems (RSS)*, Cambridge, Massachusetts.
- Salisbury, K., Brock, D., Massie, T., Swarup, N., and Zilles, C. (1995). Haptic rendering: programming touch interaction with virtual objects. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, pages 123–130. ACM.
- Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55.
- Schäfer, D. and Hüllermeier, E. (2018). Dyad ranking using Plackett–Luce models based on joint feature representations. *Machine Learning*, 107(5):903–941.
- Settles, B. (2009). Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin-Madison, Department of Computer Sciences.
- Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and De Freitas, N. (2016). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175.
- Shivaswamy, P. and Joachims, T. (2012). Online structured prediction via coactive learning. In *Proceedings of the 29th International Conference on Machine Learning, ICML’12*, pages 59–66, Madison, WI, USA. Omnipress.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1):99–118.
- Slovic, P. (1995). The construction of preference. *American Psychologist*, 50(5):364–371.
- Sørensen, Ø., Stein, A., Netto, W. L., and Leslie, D. S. (2024). Sequential rank and preference learning with the Bayesian Mallows model. *arXiv preprint arXiv:2412.13644*.
- Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. (2010). Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 1015–1022.

- Stecklina, J. et al. (2020). Haptic interaction in virtual reality environments for manual assembly validation. In *Proceedings of the 53rd CIRP Conference on Manufacturing Systems*.
- Sun, H., Shen, Y., and Ton, J.-F. (2024). Rethinking Bradley-Terry models in preference-based reward modeling: Foundations, theory, and alternatives. *arXiv preprint arXiv:2411.04991*.
- Tan, H. Z., Durlach, N. I., Beauregard, G., and Srinivasan, M. A. (1994). Manual resolution of length, force, and compliance. *Advances in Robotics, Mechatronics and Haptic Interfaces*, 49:13–18.
- Tenenbaum, J. B., Kemp, C., Griffiths, T. L., and Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331(6022):1279–1285.
- Tesauro, G. (1988). Connectionist learning of expert preferences by comparison training. In *Advances in Neural Information Processing Systems*, volume 1. Morgan-Kaufmann.
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286.
- Tolasa, H., Catkin, B., and Patoglu, V. (2025). Human-in-the-loop optimization of perceived realism of multi-modal haptic rendering under conflicting sensory cues. *IEEE Transactions on Haptics*, 18(2):295–304.
- Trafimow, D. (2017). The probability of the null hypothesis. *Synthese*, 194(5):1545–1563.
- Tucker, M., Cheng, M., Novoseller, E., Cheng, R., Yue, Y., Burdick, J. W., and Ames, A. D. (2020). Human preference-based learning for high-dimensional optimization of exoskeleton walking gaits. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3423–3430. IEEE.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1):31–48.
- Tversky, A. and Simonson, I. (1993). Context-dependent preferences. *Management Science*, 39(10):1179–1189.
- Ukkonen, A., Puolamäki, K., Gionis, A., and Mannila, H. (2009). A randomized approximation algorithm for computing bucket orders. *Information Processing Letters*, 109(7):356–359.
- Vembu, S. and Gärtner, T. (2011). Label ranking algorithms: A survey. In Fürnkranz, J. and Hüllermeier, E., editors, *Preference Learning*, pages 45–64. Springer, Berlin, Heidelberg.
- Venn, J. (1866). *The Logic of Chance*. Macmillan, London.
- Vitelli, V., Sørensen, Ø., Crispino, M., Frigessi, A., and Arjas, E. (2018). Probabilistic preference learning with the Mallows rank model. *Journal of Machine Learning Research*, 18(158):1–49.
- von Mises, R. (1957). *Probability, Statistics and Truth*. Dover Publications, New York, NY. Second revised English edition.
- von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ. 60th Anniversary Commemorative Edition published in 2004 with introduction by Ariel Rubinstein.

- Webb, M. and Sidebotham, D. (2020). Bayes' formula: a powerful but counterintuitive tool for medical decision-making. *BJA Education*, 20(6):208–213.
- Weber, E. H. (1834). *De tactu: annotationes anatomicae et physiologicae*. Koehler, Leipzig. (Original formulation of Just Noticeable Difference).
- Werner, T. (2022). A review on instance ranking problems in statistical learning. *Machine Learning*, 111(2):415–463.
- Wilde, N., Kulić, D., and Smith, S. L. (2020). Active preference learning using maximum regret. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 10952–10959, Las Vegas, NV, USA. IEEE Press.
- Yannakakis, G. N., Maragoudakis, M., and Hallam, J. (2009). Preference learning for cognitive modeling: A case study on entertainment preferences. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 39(6):1165–1175.
- Zhang, C. and Matsen IV, F. A. (2024). A variational approach to bayesian phylogenetic inference. *Journal of Machine Learning Research*, 25(1):1–56.



## APPENDIX A

### HIERARCHICAL MODELS FOR COACTIVE FEEDBACK

This appendix details the Bayesian hierarchical models constructed to analyze the determinants of participant feedback. Two distinct models were developed: one to quantify the influence of physical parameters on feedback (Section A.1) and another to assess the specificity of feedback relative to the rejected option (Section A.2).

#### A.1 Model 1: Determinants of Feedback (Physical Parameters)

To quantify the influence of physical parameters on the participants' coactive feedback ("Harder" vs. "Softer"), a hierarchical Bayesian logistic regression model was constructed. This model evaluates how the physical properties of a chosen haptic environment predict the probability of a participant providing a specific coactive signal. The results of this analysis and the corresponding posterior distributions are presented in Section 4.2.2.1.

The model accounts for the repeated measures design (each participant provided multiple responses) by including participant-level random intercepts. This structure distinguishes between a participant's baseline tendency to report an environment as "Hard" and the actual effect of the physical parameters (Stiffness, Damping, and Distance).

Let  $y_i$  be the binary response for the  $i$ -th trial. The outcome is defined as  $y_i = 1$  for "Hard" feedback and  $y_i = 0$  for "Soft" or "No Feedback." The probability of reporting "Hard," denoted as  $p_i$ , is modeled as:

$$y_i \sim \text{Bernoulli}(p_i) \quad (22)$$

$$\text{logit}(p_i) = \alpha_{j[i]} + \beta_K z_{K,i} + \beta_B z_{B,i} I_{B,i} + \beta_D z_{D,i} I_{D,i} \quad (23)$$

where:

- $j[i]$  indicates the participant identifier associated with trial  $i$ .
- $\alpha_j$  is the random intercept for participant  $j$ , capturing their baseline bias for reporting "Hard."

- $z_{K,i}, z_{B,i}, z_{D,i}$  are the z-score normalized values of Stiffness, Damping, and Distance for the selected option in trial  $i$ .
- $I_{B,i}$  and  $I_{D,i}$  are indicator variables (0 or 1) that activate the damping and distance terms only for conditions where those parameters were varied (i.e., KB and KBD conditions).

## Priors

Weakly informative priors were assigned to all parameters to regularize estimates while allowing the data to drive the posterior:

$$\alpha_j \sim \mathcal{N}(0, \sigma_\alpha) \quad (24)$$

$$\sigma_\alpha \sim \text{Exponential}(1.0) \quad (25)$$

$$\beta_K, \beta_B, \beta_D \sim \mathcal{N}(0, 0.5) \quad (26)$$

## Computation and Inference

The model was implemented using the probabilistic programming library `PyMC` in Python (Salvatier et al., 2016). Posterior distributions were approximated using the No-U-Turn Sampler (NUTS), a self-tuning variant of Hamiltonian Monte Carlo (HMC). Sampling was performed with a target acceptance probability of 0.95 (`target_accept=0.95`) to ensure robust exploration. Convergence was verified by inspecting trace plots and ensuring that the Gelman-Rubin statistic ( $\hat{R}$ ) was  $\leq 1.01$  for all parameters.

### A.2 Model 2: Specificity of Feedback

To determine whether the coactive feedback was specific to the selected haptic option or influenced by the rejected alternative, a second hierarchical logistic regression model was constructed. This model tests for potential contrast effects or distractions caused by the rejected option’s torque profile. The results of this analysis are presented in Section 4.2.2.2.

The probability  $p_i$  of reporting “Hard” for trial  $i$  is modeled as:

$$y_i \sim \text{Bernoulli}(p_i) \quad (27)$$

$$\text{logit}(p_i) = \alpha_{j[i]} + \beta_{\text{sel}} z_{\text{sel},i} + \beta_{\text{oth}} z_{\text{oth},i} + \beta_{\text{sq}} z_{\text{oth},i}^2 + \beta_{\text{cond}[i]} \quad (28)$$

where:

- $\alpha_{j[i]}$  is the random intercept for participant  $j$ .
- $z_{\text{sel},i}$  is the normalized torque distance of the *selected* option from the reference torque.

- $z_{\text{oth},i}$  is the normalized torque distance of the *rejected* (other) option from the reference torque.
- $\beta_{\text{sq}}$  is the coefficient for the squared term of the rejected option's distance, included to detect potential non-linear distraction effects.
- $\beta_{\text{cond}[i]}$  represents the baseline effect for each experimental condition (K, KB, KBD).

## Priors

As in the first analysis, weakly informative priors were used:

$$\alpha_j \sim \mathcal{N}(0, \sigma_\alpha) \quad (29)$$

$$\sigma_\alpha \sim \text{Exponential}(1.0) \quad (30)$$

$$\beta_{\text{sel}}, \beta_{\text{oth}}, \beta_{\text{sq}}, \beta_{\text{cond}} \sim \mathcal{N}(0, 0.5) \quad (31)$$

## Computation and Inference

Computation was identical to Model 1, utilizing NUTS sampling with `target_accept=0.95` and standard convergence checks ( $\hat{R} \leq 1.01$ ).



## APPENDIX B

### PILOT STUDY DETAILS

A pilot study was conducted to identify the most natural adjectives users assign to haptic impedance variations. Sixteen participants contributed to the study, engaging with distinct comparison sets to validate the vocabulary. Table 12 details the frequency of adjectives reported by participants when describing the difference between haptic stimuli.

**Table 12:** Frequency of Feeling Words Reported in Pilot Study (N=16)

Adjective	Frequency
Hard / Stiff	22
Soft	14
Difficult	9
Easy	11
Light	7
Tight	5
Strong	4
Resistant	5
Heavy	4

Terms appearing only once or twice (e.g., “viscoelastic,” “spring-like”) were considered noise and excluded from the summary. The dominant cluster of “Hard” and “Soft” informed the design of the coactive feedback interface.

#### Pilot Trial Configurations

The parameter sets used to elicit these responses are listed below. Participants were assigned to either Group A or Group B to cover a broader range of parameter contrasts. Group A focused on distinct changes in individual parameters, while Group B examined more complex combinations.

**Table 13:** Pilot Study Parameter Configurations - Group A

<b>Trial</b>	<b>Set 1</b>	<b>Set 2</b>	<b>Description</b>
1	$K = 0$	$K = 300$	High K value
2	$B = 0$	$B = 40$	High B value
3	$K = 300, B = 15$	$K = 100, B = 15$	Much lower K value with constant B
4	$K = 150, B = 10$	$K = 150, B = 30$	Higher B value with constant K
5	$K = 350, B = 30$	$K = 200, B = 20$	Lower K and slightly lower B values
6	$K = 150, D = 0$	$K = 150, D = 0.5$	Distanced K value
7	$K = 300, B = 20, D = 0$	$K = 300, B = 20, D = 0.5$	Distanced K and B values

**Table 14:** Pilot Study Parameter Configurations - Group B

<b>Trial</b>	<b>Set 1</b>	<b>Set 2</b>	<b>Description</b>
1	$K = 200$	$K = 100$	Lower K value
2	$B = 10$	$B = 25$	Slightly higher B value
3	$K = 50, B = 10$	$K = 50, B = 25$	Slightly higher B value with constant small K
4	$K = 200, B = 20$	$K = 100, B = 20$	Low K value with constant B
5	$K = 100, B = 30$	$K = 150, B = 15$	High K and low B value
6	$K = 150, B = 20$	$K = 300, B = 5$	High K value and small B value
7	$K = 100, B = 20, D = 0.3$	$K = 200, B = 25, D = 0.3$	High K and B with constant Distance