Development and Initial Validation of the Haptic Experience Inventory (HXI)

by

Tianzheng Shi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences and Engineering

Waterloo, Ontario, Canada, 2026

© Tianzheng Shi 2026

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of contributions

This thesis documents the development and initial validation of the Haptic Experience Inventory (HXI), a novel scale for measuring haptic experience.

This thesis is based in part on a manuscript published during Tianzheng Shi's master's program [73] authored by Tianzheng Shi and Oliver Schneider. Tianzheng Shi was the primary author of the published manuscript, which was written under the supervision of Prof. Oliver Schneider.

The work presented in this thesis constitutes a comprehensive effort that covers all major stages of questionnaire development and validation. The conceptualization and design of this work were partially informed by prior efforts to develop measures of haptic experience conducted by former students under the supervision of Prof. Oliver Schneider [3, 45, 65]. In addition, Tianzheng Shi contributed to the data analysis reported in one of these prior projects [3].

Chapter 11 documents ongoing work on the development and validation of a French version of the HXI. This work constitutes an independent research project, in which Tianzheng Shi is a co-author.

Prof. Oliver Schneider provided guidance throughout all stages of the research and offered feedback on manuscript drafts. He also contributed to the writing of the published manuscript referenced above [73].

Abstract

Haptic Experience (HX) encompasses distinct quality criteria specific to haptic interactions, yet no standardized instrument exists to measure it. This makes understanding and evaluating HX challenging. This thesis reports on the development and validation of the Haptic Experience Inventory (HXI), a questionnaire measuring HX. An item pool of 50 items is developed through theoretical construction, expert reviews (N=10), and cognitive interviews (N=9). These items are then subjected to exploratory and confirmatory factor analysis using data from 591 participants across in-person and online studies, covering vibrotactile, force-feedback, and mid-air devices. Eventually, a 20-item HXI with five dimensions is established: Autotelics, Realism, Involvement, Harmony, and Discord. The HXI converges with theory and shows strong reliability, validity, and measurement invariance, suggesting it is effective across deployed modalities and contexts. The HXI provides empirical evidence about the structure of HX and offers a robust, standardized tool for assessing haptic feedback in research and practice. Ongoing efforts to develop and validate a French version of the HXI are also reported.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Oliver Schneider, for his invaluable guidance, continuous support, and insightful feedback throughout my program. His expertise and mentorship were essential to both my professional and personal development during my time at the University of Waterloo.

I am deeply grateful to my family for their unconditional love and unwavering support throughout my graduate studies. Their understanding, encouragement, and trust gave me the confidence to pursue this work and sustained me through challenges and uncertainties along the way.

I also want to thank the University of Waterloo HCI community, especially the HXLab and the Games Institute for valuable feedback and support, as well as Tor-Salve Dalsgaard and Margot Racat for their feedback and help in the theoretical construction and item generation stages. I also thank collaborators Vincent Lévesque and Naomi Catwell for their collaboration on the French adaptation of the instrument.

I acknowledge the support of the Canada Foundation for Innovation (CFI), Ontario Research Foundation (ORF), and Mitacs Globalink Graduate Fellowship for the work. The user studies were partly supported by an unrestricted gift from Google intended to support "Software Tools for Rendering Haptics and Assessing Quality at Scale".

Finally, I am grateful to my friends for their companionship, encouragement, and support throughout this journey.

Table of Contents

Author's Declaration ii

Statement of contributions iii

Abstract iv

Acknowledgements v

List of Figures x

List of Tables xi

1 Introduction 1

1.1 Research Question . . . 2

1.1.1 Can Haptic Experience (HX) be reliably measured by a standardized questionnaire? . . . 2

1.1.2 What is the structure of Haptic Experience (HX)? . . . 2

1.2 Research Contributions . . . 3

1.3 Thesis Outline . . . 3

2 Related Work 6

2.1 Measuring User Experience . . . 6

2.2 Haptic Design and Experience . . . 8

2.3 Measuring Haptic Experience . . . 9
2.4 Chapter Summary - Addressing Gaps in Measuring Haptic Experience . . . 11

3 Theory and Item Construction **13**
3.1 Dimension Revisions . . . 13
3.1.1 Adopting Involvement in place of Immersion . . . 13
3.1.2 Retaining Autotelics . . . 14
3.2 Item Drafting Approach . . . 15
3.3 Autotelics . . . 17
3.4 Realism . . . 17
3.5 Harmony . . . 18
3.6 Expressivity . . . 18
3.7 Involvement . . . 19

4 Pre-testing and Refining Items **20**
4.1 Expert Review . . . 20
4.1.1 Procedures . . . 20
4.1.2 Results . . . 21
4.2 Cognitive Interviews . . . 22
4.2.1 Procedures . . . 22
4.2.2 Results . . . 23

5 Exploratory Factor Analysis **25**
5.1 Design and Procedure . . . 25
5.2 Participants . . . 26
5.3 Analyses . . . 27
5.4 Results and Discussion . . . 27

6 Confirmatory Factor Analysis **32**
6.1 Design and Procedure . . . 32
6.2 Participants . . . 33
6.3 Results and Discussion . . . 33

7 Psychometric Evaluation **34**
7.1 Internal consistency reliability . . . 34
7.2 Convergent, divergent and predictive validity . . . 34
7.3 Measurement invariance . . . 35

8 Discussion **37**
8.1 HXI: A Robust Questionnaire Ready for Haptic Experience Assessment . . . 37
8.1.1 HXI is better tailored for HX than general UX questionnaires . . . 38
8.1.2 HXI Applies to Various Haptic Contexts . . . 38
8.2 Understanding the Factor Structure of Haptic Experience . . . 39
8.2.1 Harmony and Discord: Two Sides of Harmony . . . 39
8.2.2 Expressivity: System Characteristic Rather than Experiential Factor 40

9 Limitations **41**
9.1 Exploring More Modalities and Contexts . . . 41
9.2 Further Testing of HXI's Validity . . . 42
9.3 Recommendations on Reporting for Continual Validation of the HXI . . . 43
9.4 Segment-Specific Insights in Assessing HX . . . 43
9.5 Leveraging HXI to Enhance Design Processes . . . 44

10 Administering HXI **45**
10.1 Scope of the HXI . . . 45
10.2 Structure of the HXI . . . 46
10.3 Form of the HXI . . . 46

10.4 Scoring the HXI . . . 47
10.5 Interpreting Scores . . . 47
10.6 Reporting HXI Results . . . 47

11 Ongoing Work & Conclusion **49**
11.1 Translation of the HXI into French . . . 49
11.2 Pilot Testing of the Translated Version . . . 51
11.3 Validation and Psychometric Evaluation of the Translated Version . . . 51
11.4 Conclusion . . . 52

References **54**

APPENDICES **64**

A Items drafted based on theoretical construction and revisions made during expert review **65**
A.1 Items for *Autotelics* . . . 65
A.2 Items for *Realism* . . . 66
A.3 Items for *Harmony* . . . 67
A.4 Items for *Expressivity* . . . 69
A.5 Items for *Involvement* . . . 70

B Items entering factor analysis, after Study 1 and Study 2 **72**
B.1 Items for *Autotelics* . . . 72
B.2 Items for *Expressivity* . . . 73
B.3 Items for *Harmony* . . . 73
B.4 Items for *Involvement* . . . 74
B.5 Items for *Realism* . . . 74
B.6 Attention Check Items . . . 74

C Descriptive for data collected in Study 3 **75**

List of Figures

1.1 Flowchart of the development process of HXI. 5

5.1 Parallel analysis scree plot: result suggests 5 or 6 factors, with simulated and observed eigenvalue overlapping for the 6th factor 28

5.2 Item loadings for EFA: the five factors each have 4 items with high loadings and the pattern is consistent with the theoretical model 31

11.1 Translation workflow 50

List of Tables

3.1 Summary of HXI Factors and Key Revisions: This table provides a concise summary of the HXI factors and the key revisions made when drafting items in the theory construction stage. 16

4.1 Haptic Devices and Applications Used: This table describes the haptic devices and applications used in in-person studies, as well as user tasks and haptic feedback in these applications. 24

7.1 Pearson Correlations Between HXI and UEQ-S Subscales 35

7.2 Measurement Invariance Testing Results. 36

C.1 Statistical Summary of Online and In-Person Scores in Study 3 76

Chapter 1

Introduction

Haptic technology is increasingly popular for its potential to enhance user experience (UX) across various applications. For example, mid-air and vibrotactile stimuli have been shown to make UX more pleasant, unpredictable, and creative [52]. When interacting with objects in a virtual environment, passive haptic props can also provide a satisfying sense of presence [4]. Vibrotactile haptic feedback can also enhance enjoyability, aesthetic appeal, immersion, and meaning in gaming experiences [74]. To help further support understanding, communication, and evaluation of experiences involving haptics, Kim and Schneider [45] introduced Haptic Experience (HX): a set of quality criteria that are the most important considerations for experiences involving haptic technology. These criteria provide a vocabulary for how haptics can support experiences. For instance, in a combat game, players might feel varying intensities of vibration with each hit or block, creating a heightened sense of *Involvement* by physically drawing them into the action. Additionally, well-timed vibrations can enhance *Harmony* by aligning feedback with visual and auditory cues, making the experience feel seamless and cohesive.

While research on HX continues to grow, there remains a lack of quantitative tools to systematically study and evaluate HX. This absence presents significant challenges for both haptic researchers and practitioners. Most existing UX questionnaires are ill-suited for capturing the unique experiential qualities of haptic interactions, leading to the reliance on qualitative approaches such as small-scale, in-person ad-hoc acceptance testing. While informative, these methods are time-consuming and impractical for large-scale deployment [67]. Meanwhile, most quantitative measures use self-developed questionnaire items, which lack proper validation and have questionable reliability and accuracy in capturing HX [45, 67]. Together, these limitations highlight an urgent need for a standardized quantitative instrument specifically tailored to measure HX.

1.1 Research Question

The lack of a validated quantitative measurement tool for HX has been noted in prior research. Several attempts have been made to develop questionnaires specifically tailored to HX (see Section 2.3); however, these instruments have not demonstrated sufficient validity and reliability [3, 65]. Building on these limitations, this project addresses the following two research questions.

1.1.1 Can Haptic Experience (HX) be reliably measured by a standardized questionnaire?

Despite the need for a standardized questionnaire specifically measuring HX, prior attempts have yet to produce a fully validated instrument. Whether this reflects limitations in previous methodological approaches or an inherent complexity in HX itself remains an open question. Section 2.3 examines these prior efforts and their potential shortcomings. By initiating a more rigorous, end-to-end questionnaire development process grounded in more rigorous and systematic methods, this work aims to determine whether a reliable and valid measure of HX is achievable, and to produce one if so.

1.1.2 What is the structure of Haptic Experience (HX)?

Since the original theoretical model of HX was proposed [45], both theoretical developments and empirical evidence from prior questionnaire development efforts have suggested that the structure of HX may be more nuanced than initially described. In particular, findings by Anwar et al. [3] indicate that HX measured across multiple haptic modalities may exhibit a structure different from that observed in vibrotactile-only contexts. In this thesis, we reviewed and revised the theoretical model of HX (elaborated in Chapter 3) by integrating insights from theoretical literature and previous questionnaire development efforts. The subsequent validation of the proposed questionnaire also serves as an empirical test of the revised model (elaborated in Chapter 6 and Chapter 7), contributing to a more up-to-date and empirically grounded understanding of HX.

1.2 Research Contributions

This thesis establishes a five-factor, 20-item *Haptic Experience Inventory* (HXI), encompassing the dimensions of Autotelics, Realism, Involvement, Harmony, and Discord.

Addressing the first research question, we developed the first validated questionnaire specifically measuring HX. The HXI showed strong psychometric properties, including high internal consistency reliability (overall Cronbach's $\alpha = 0.94$, with the lowest factor $\alpha = 0.84$), good convergent validity with UEQ-S ($r = 0.75$), the short version of the User Experience Questionnaire, a fast and reliable questionnaire to measure the UX of interactive products. It also demonstrates strong predictive validity, supported by a high correlation with overall satisfaction with the haptic interaction ($r = 0.82$). Configural, metric and scalar invariance were also established, confirming its effectiveness across different haptic modalities and study formats. As a standardized, easily deployable instrument, the HXI is tailored for HX, enables large-scale evaluation, and facilitates meaningful comparisons across different studies and applications.

Addressing the second research question, our revised HX model was supported by strong convergence between the theoretical model and the empirically derived HXI factor structure. The final HXI model contains five factors: Autotelics, Realism, Harmony, Discord, and Involvement. In addition to iterative refinement of factor definitions and item formulations, the factor Expressivity was dropped, and the original factor Harmony splits into Harmony and Discord. The dropping of Expressivity suggests it may be better categorized as a system characteristic rather than an experiential dimension (elaborated in subsection 8.2.2). The distinction between Harmony and Discord aligns with the concepts of hygiene factors and motivators in UX theory, where the absence of harmony may be more noticeable to users, while good haptics tend to remain subtle and less prominent.

1.3 Thesis Outline

- Chapter 2 reviews related work, including the research on UX measurement, haptic design and experience, as well as existing gaps and prior efforts in measuring HX.
- Chapter 3–Chapter 7 covers the development and initial validation of the HXI, following the questionnaire development process suggested by DeVellis and Thorpe [21] and Boateng et al [8]. A diagram illustrating this process is presented below in Figure 1.1.

  - Chapter 3 describes the review and revision of HX factors and the generation of an initial item pool of 84 items.
  - Chapter 4 reports expert reviews ($N = 10$) and cognitive interviews ($N = 9$) which evaluates the suitability of developed factors and items, resulting in a refined pool of 50 items.
  - Chapter 5 presents exploratory factor analysis (EFA, $N = 360$) used to examine the underlying structure of the items.
  - Chapter 6 reports confirmatory factor analysis (CFA, $N = 231$) to validate the derived factor structure.
  - Chapter 7 evaluates internal consistency reliability, convergent and predictive validity, and measurement invariance using data from EFA and CFA.
- Chapter 8 discusses theoretical and practical implications of the HXI.
- Chapter 9 outlines limitations and directions for future research.
- Chapter 10 provides guidance on administering and applying the HXI in research and practice.
- Chapter 11 documents ongoing work on extending the HXI by developing a French version of the HXI.

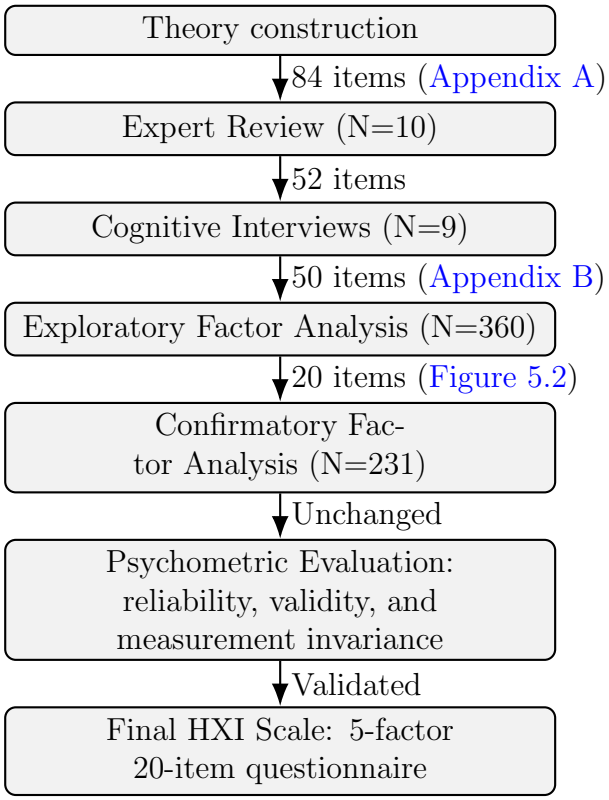

Figure 1.1: Flowchart of the development process of HXI.

Chapter 2

Related Work

This chapter discusses work related to the measurement of user experience, introduces important concepts and research in haptic design and experience, and describes the challenges in and the importance of measuring haptic experience, highlighting the need for a standardized quantitative questionnaire measuring haptic experience. In the end, we summarized the various reasons for the need to develop a questionnaire tailored for haptic experience, highlighting the importance of the research project described in this thesis.

2.1 Measuring User Experience

User experience, often shortened to UX, is a multifaceted concept that captures users' overall experience when interacting with a product, system or service. A widely adopted distinction in UX research separates experience into two primary dimensions: pragmatic and hedonic. Pragmatic qualities focus on getting tasks done, and refer to the practical aspects of interaction, including usability, utility, and efficiency. Hedonic qualities, on the other hand, focus on non-task aspects that contribute to the user's emotional and affective responses, such as enjoyment, emotions and pleasure during interaction [33]. This distinction has informed both theoretical models of UX and the development of widely used UX measurement instruments [34, 80].

Evaluating UX is an integral part of the design process, and typically employs various methodologies including observational studies, psychophysiological measures, and questionnaires or surveys. Among these, questionnaires remain the most widely used approach due to their ease of deployment and ability to capture user perceptions systematically [53].

Several questionnaires have been established to measure general UX, intended to apply to all products or contexts. The User Experience Questionnaire (UEQ) [48] is one such tool that categorizes UX into six dimensions: Perspicuity, Efficiency, and Dependability in the pragmatic category, and Novelty, Stimulation, and Attractiveness in the hedonic category. AttrakDiff2 [34] has a similar structure, having items from factors of Pragmatic, Hedonic and Attractiveness. Attractiveness in this context is a combination of ergonomic and hedonic qualities, reflecting the overall appeal. The meCUE questionnaire takes a modular approach to assess UX, which consists of modules related to the perception of different product characteristics (usefulness, usability, visual aesthetics, status, commitment), to users' emotions (both positive and negative emotions), and to consequences (product loyalty and intention to use) [57]. HX, as a specific variant of UX, could also have components from both pragmatic and hedonic aspects.

Beyond general UX measures, there are many questionnaires tailored to specific aspects of contexts of user experience. These questionnaires often have specific factor structures depending on the concept being measured and the settings it was designed to apply to. For example, the Presence Questionnaire (PQ) has been widely used to measure sense of presence in virtual environments. PQ is made up of Involvement, Adaptation/Immersion, Sensory Fidelity, and Interface Quality [84, 85]. The Player Experience Inventory (PXI) measures player experiences at the level of functional and psychosocial consequences, consisting of 10 factors, such as Meaning, Immersion, Autonomy and Challenge [1]. The Game Experience Questionnaire (GEQ) measures game experience and has seven core components: Immersion, Flow, Competence, Positive and Negative Affect, Tension, and Challenge [41]. These specialized questionnaires highlight the importance of tailoring assessments to the specific experiences users encounter in different contexts. Our efforts towards developing a questionnaire specifically measuring Haptic Experience align with this trend toward domain-specific assessment.

Development in UX theory has a more nuanced, rich description of the experience of technology. For example, eudaimonia (striving towards personal goals or flourishing) has been proposed and measured as a complementary experiential category from hedonia (momentary pleasures) [55]. Models of user experience involve a split view of the designer from the user, suggesting that individual context and the environment of use determines user experience in addition to the interface [33].

In the HX context, these theoretical advancements are relevant as they suggest that Haptic Experience, like other forms of UX, is multifaceted and sensitive to the user's context and the nuances of interaction. Recent work on unifying UX and HX has positioned the HX criteria as part of the apparent system character, and highlighted the need for more theoretical grounding as the field develops [20]. Continuing to integrate insights from UX

theory into the development of the HXI helps ensure that it captures not only functional and sensory dimensions but also the broader experiential outcomes that contribute to meaningful haptic interactions.

2.2 Haptic Design and Experience

Haptics refers to technologies and senses related to touch. Its two main components are taction and kinesthesia. Tactile sensors in the skin allow people to feel surface textures and qualities, while kinesthetic sensors provide the sense of body forces and motions [24]. Haptic technology can evoke these sensations through a variety of means, including passively through physical contact with surfaces and objects, and actively through programmatic stimuli such as vibrations, forces, and motions. The most common types of active haptic feedback are vibrotactile feedback which produces vibrations that could be felt by tactile sensors and are commonly used in mobile devices and game controllers [16], and force feedback devices such as robotic arms providing kinesthetic feedback [71]. Beyond these two most common types of feedback in commercial products, there is a large and increasing set of alternative modalities. For example, mid-air haptic feedback, which uses focused ultrasound to project haptic feedback to users' hands through ultrasonic transducers [10], has emerged as more recent technology that is both commercially successful and heavily researched for user experience [19, 29, 59].

Prior research has demonstrated that well-designed haptic feedback can meaningfully enhance user experience across applications and modalities [51]. For example, mid-air and vibrotactile stimuli contributes to a more pleasant, unpredictable, and creative UX [52]. Passive haptic props provides a satisfying sense of presence when interacting with objects in a virtual environment [4], as long as multisensory feedback is carefully coordinated, e.g., between visual and tactile feedback [7]. Vibrotactile haptic feedback leads to enhanced enjoyability, aesthetic appeal, immersion, and meaning in gaming experiences [74]. Psychophysiological signals also show affective feedback, e.g., through EEG, when haptic feedback is present in a vibro-kinetic setting [61].

Importantly, the design of haptic systems often involves a focus on technical expressiveness and parameters, such as the physical range of achievable sensations. While these characteristics are from a system design perspective, they do not directly translate into positive user experience. Richer haptic sensations do not guarantee better experiences if designed or integrated poorly. Differentiating between system characteristics and experiential qualities is vital to evaluating haptic interactions. Thus, evaluating haptic experience requires recognition of it emerging from the interaction between system characteristics,

user actions, and multisensory context, rather than assessing the technical properties of haptic systems.

2.3 Measuring Haptic Experience

Despite the benefits shown of haptic technology contributing to UX, hapticians need specific measurement tools to measure how much and in what aspect user experience is actually enhanced. The lack of measurement tools for HX was a major obstacle identified by Schneider et al. [67] in the Haptic Experience Design (HaXD) process. They found that general UX questionnaires often do not translate well to haptic interactions, leaving researchers to rely primarily on qualitative approaches such as in-person, ad hoc testing. While informative, these methods are poorly suited for large-scale deployment, as they are time-consuming and typically require direct researcher involvement during the evaluation process.

Meanwhile, when haptic researchers want to use quantitative measures for HX, the most deployed form is self-developed Likert-style items (e.g., [9, 14, 30, 42, 49, 79]). For example, researchers aiming to obtain a quantitative rating of realism might ask, "Do you feel the haptic sensations are realistic?" and provide response options ranging from "very unrealistic" to "very realistic." These items are typically drafted by the researchers themselves without undergoing any validation process. This raises concerns about their reliability and accuracy.

Furthermore, researchers may not report all the items they used, often presenting only one or two as examples, which can limit transparency and replicability. This also presents challenges to haptic researchers and developers in interpreting and comparing results across studies as items differ for each study [45, 67]. Additionally, haptic developers and users also face challenges due to a lack of proper language describing experiences in haptic interactions, which hampers effective communication about specific aspects of these experiences [67].

These gaps highlight the need for a framework of Haptic Experience and a standardized, quantitative questionnaire specifically designed to assess HX. In response to this need, several studies have focused on understanding and evaluating HX.

Kim and Schneider [45] proposed the HX model based on a comprehensive literature review and user studies. Their model includes 5 factors: Harmony, Autotelics, Expressivity, Immersion and Realism, each of which captures different dimensions of HX. This model shares some similar theme and terminology with general user experience, while capturing

unique constructs and emphasis within haptic technology (e.g. Autotelics and Realism in this context are very specific to haptics). Researchers have started to build on this model by exploring new design parameters [64]. This model offers valuable insights into HX, providing a foundation for researchers to generate items for developing a measurement tool specifically tailored to HX.

Based on this model [45], Sathiyamurthy et al. [65] developed an item pool for an HX questionnaire and conducted exploratory factor analysis. The items converged in a way consistent with the theoretical model. However, their study only used an online sample with vibrotactile modality (participants' smartphones or gaming consoles), which limits the scope of the questionnaire to only vibrotactile technologies. Further, their study did not include a confirmatory factor analysis, leading to limited robustness, applicability and validity.

This effort was furthered by Anwar et al. [3], who conducted both EFA and CFA in an effort to finalize a reliable and validated questionnaire, using the same item pool drafted by Sathiyamurthy et al. [65]. Their study was conducted in-person, and involved different haptic modalities (vibrotactile, force feedback, surface haptics, and mid-air haptics). A slightly different model was derived, with Autotelics failing to load and Immersion reshaped as Involvement. They believe the difference could be due to different modalities being used or the study conditions. However, reliability of items under two factors in the questionnaire fell below the acceptable threshold, possibly partly due to inadequate number of items as there are only 2 or 3 items per factor remaining in the end. Consequently the developed questionnaire is not fully reliable, indicating the complexities involved in capturing the multifaceted nature of HX.

Here are a few possible reasons why these efforts fell short. First, the shared item pool was originally developed in a study with vibrotactile stimuli only, which may not fully capture HX across other modalities. Second, the HX model itself may require revision, given that some factors failed to load. Third, the absence of data quality mechanisms — such as attention checks — may have introduced noise into the collected responses. Finally, the limited number of items in the pool leaves little room for selection during scale development, which is especially consequential given the challenges already present in the process.

Despite the questionnaire being not fully reliable and validated, and the authors' explicit cautions against its use, some studies already began employing it to measure HX [13, 64], as no other tools are available, demonstrating the pressing need for such a tool. Meanwhile, others have used the HX model more generally to structure their evaluations, drawing inspirations from the factor structure of HXI as it provides insights into the most important

aspects of haptic interactions that shape haptic experiences [60, 86]. This further highlights the pressing need for a robust HX instrument.

2.4 Chapter Summary - Addressing Gaps in Measuring Haptic Experience

The review of related concepts and prior work highlights a clear methodological gap in the measurement of Haptic Experience on several levels.

While general UX questionnaires offer well-validated tools for assessing overall user experience, they lack sensitivity to the modality-specific and experiential qualities of haptic interaction. Many of their items are not applicable to haptic contexts, and as a result, these instruments fail to capture the nuanced ways in which haptic feedback contributes to experience.

Without a standardized and rigorously validated measure of HX, researchers are often forced to develop custom and usually unvalidated items, or rely solely on qualitative, ad-hoc testing. These practices introduce inconsistencies in wording and interpretation, raise concerns over measurement validity, hinder communication and comparison across different studies, and poses challenges for large-scale assessment. These measures are typically unstructured, only covering selected aspects of HX, limiting their ability to systematically evaluate design choices or benchmark haptic systems.

Existing efforts to measure HX directly have established a promising conceptual starting point, but several limitations in prior work constrain their effectiveness as a foundation for a standardized instrument. Most notably, the reliance on a shared item pool originally developed for vibrotactile stimuli raises concerns about construct coverage across broader haptic modalities. In addition, issues with factor structure suggest that the underlying HX model itself may require refinement. Methodological constraints further compound these issues: the absence of data quality safeguards (e.g., attention checks) introduces potential noise, while the limited size of the item pool restricts item selection, weakening the robustness of scale construction.

Taken together, these limitations indicate that incremental adjustments to existing instruments may not be sufficient. Instead, they point to the need for a full restart of the scale development process, beginning with a careful re-examination of the theoretical construct and followed by theory-driven item generation.

The work presented in this thesis responds to this need by undertaking a complete questionnaire development and validation process aimed at producing a standardized tool

for measuring Haptic Experience. Such an instrument would enable more precise evaluation of haptic systems and support researchers and designers in creating more effective and satisfying user experiences.

We followed best practices in questionnaire development [8], undertaking a new full-cycle scale development process that spans scale conception (theory-driven construct definition, iterative item refinement), data-driven scale construction and comprehensive scale validation and psychometric evaluation. In particular, a robust HX instrument must demonstrate reliability, validity, and measurement invariance across different haptic modalities and study contexts.

Chapter 3

Theory and Item Construction

This chapter outlined the process of reviewing and revising the HX theory and developing items accordingly. The foundation of our scale was built upon the proposed five-factor HX model [45], which identified five key dimensions of HX: Autotelics, Realism, Harmony, Expressivity, and Immersion. We also drew upon discoveries made in prior studies to develop a haptic experience questionnaire [3, 65], and related concepts in HX and HCI research. This process resulted in an initial item pool for further evaluation and refinement.

3.1 Dimension Revisions

This section lays out groundings for two changes we made to the theoretical model – we adopted Involvement in place of Immersion, and retained Autotelics despite it being dropped previously [3]. As such, we adopted the constructs of Autotelics, Realism, Harmony, Expressivity, and Involvement in our study.

3.1.1 Adopting Involvement in place of Immersion

Rather than "Immersion", we opted for "Involvement", a concept related to Immersion, echoing the decision made in the work of Anwar et al. [3]. This choice reflects our view that Involvement more accurately captures the role of haptics in a variety of interaction contexts.

Immersion is commonly associated with a holistic psychological state characterized by deep absorption, presence, and a diminished awareness of the physical world [2]. This makes

immersion inherently context-dependent as it is more often elicited in prolonged experiences involving multiple sensory modalities. It's typically understood as a global experiential outcome that requires the interaction of the system and multiple sensory modalities. For example, immersion and immersive are typically used to refer to virtual reality (VR) and augmented reality (AR) systems rather than any single modality [56]. In contrast, many haptic interactions, such as mobile notifications, are brief, task-oriented, and non-narrative. In such contexts, they can still lead to meaningful engagement, but can hardly result in the experience of immersion in the traditional sense. As such, use of the word "immersion" may implicitly require incorporating other modalities, thereby confounding the focus on contributions of haptics. Items developed with this concept in mind may also more easily lead the responder to think more about other modalities.

To address these limitations, this work adopts Involvement as the alternative construct. It is defined as "a psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events" [85]. We believe it can apply to a wider variety of media and contexts, and better reflect the intended scope of this factor. We believe involvement is a better word choice to describe experiences that can arise in short, utilitarian interactions, making it more applicable to the wide range of contexts in which haptic feedback is deployed, like mobile notifications or productivity tasks.

Additionally, the term Immersion in HCI literature often denotes objective system properties, such as the extent to which a VR system can support natural sensorimotor contingencies for perception [50, 75]. The term presence [85] is used to refer to the sense of being immersed instead. However, presence carries similar limitations, as it is predominantly studied in the context of VR and multi-sensory environments, and implies a holistic experience that may not generalize to brief or single-modality haptic interactions. Adopting Involvement instead of Immersion thus also avoids potential confusion arising from these established but divergent uses of "immersion" in the field.

3.1.2 Retaining Autotelics

Autotelics is characterized as "the touch experience as an end in and of itself" as opposed to "instrumental" [62]. As a distinct hedonic dimension, Autotelics captures intrinsic enjoyment derived directly from tactile stimulation, rather than instrumental benefits such as task support or feedback clarity. This distinction is also critical for HX, as Autotelics stand out when other factors are more connected to other modalities or system status.

In previous HX questionnaire development efforts, Autotelics exhibited an inconsistent

empirical pattern. Though it was supported in the work by Sathiyamurthy et al.[65], Anwar et al. [3] eliminated Autotelics due to low loading. However, examination of the previously proposed Autotelics items suggests that it could be attributed to issues in item formulation rather than to a lack of theoretical relevance. Several earlier items (e.g., "The haptic feedback felt satisfying.", "I would prefer the system without the haptic feedback.") were not adequately constrained to reflect the nature of Autotelics and may have encouraged participants to evaluate overall usefulness or satisfaction of haptic sensations within the multi-modality context in connection with other modalities and system events, rather than the intrinsic qualities of the haptic sensation itself.

Given the central role of intrinsic pleasure in hedonic experience, Autotelics was retained. The primary challenge was to draft items in a way that reliably isolates the intrinsic enjoyment of haptic sensations.

3.2 Item Drafting Approach

Our item revision process involved almost completely rebuilding the items from scratch. We undertook a comprehensive re-evaluation of each construct and examined which items worked and which did not in the three previous studies [3, 45, 65]. We also received valuable external input and suggestions from other haptics researchers with experience in HX research and/or questionnaire development; we are especially thankful to two researchers, whom we name in Acknowledgements.

We began by identifying limitations in the previous items, including wording, phrasing, relevance to the theoretical factor, or possible overlap with other factors. These observations informed our decision to draft new items that better capture each factor's intended scope. For each factor, we examined existing literature on Haptic Experience identified by domain expert, as well as prior items, then drafted new items intended to be more detailed, closely aligned with the constructs, and applicable across different devices and use cases. We also developed examples that might elicit high and low presence of these constructs, to facilitate a better and more intuitive understanding of the factors.

We also went through multiple iterations and discussions before finally producing the drafted item set. A total of 84 items (approximately 17 per factor) were generated. The following sections outline the definitions of the five factors and key revisions compared to the items drafted in previous studies [3, 65] during the item generation process. A summary is presented in Table 3.1. The complete list of drafted items can be found in Appendix A.

Factor	High-level Description and Key Revisions
Autotelics	Represents the pleasure derived from haptic sensations alone, without considering their functional utility. Focuses on the intrinsic enjoyment of the touch experience. **Key Revisions:** Revised items to emphasize intrinsic pleasure and removed items that may include functional utility.
Realism	Refers to how well haptic sensations simulate or represent real-world experiences, including faithful re-creations and believable exaggerations. **Key Revisions:** Enriched item content to better convey realism, ensuring both realistic and exaggerated sensations were included in the measurement.
Harmony	Refers to the seamless integration of haptic sensations with other sensory modalities, resulting in a cohesive multisensory experience. **Key Revisions:** Revised items to fall more precisely onto Harmony and reduce cross-loading with other factors.
Expressivity	Captures the richness and diversity of haptic sensations and how well they convey information about user actions and system events. **Key Revisions:** Redesigned items to highlight meaningful variation and communication of information through diverse haptic feedback.
Involvement	Refers to the level of engagement and focus experienced by users due to haptic sensations, contributing to their psychological immersion. **Key Revisions:** Adjusted items to better capture the engagement and focus resulting from haptic feedback, enhancing measures of psychological immersion.

Table 3.1: Summary of HXI Factors and Key Revisions: This table provides a concise summary of the HXI factors and the key revisions made when drafting items in the theory construction stage.

3.3 Autotelics

Autotelics is a distinct hedonic dimension, characterized as "the touch experience as an end in and of itself" as opposed to "instrumental" [62]. It refers to the positive experience gained from haptic stimuli alone, without considering its functional utility as in the system. In short, it means the haptic sensation itself feels good, pleasant, or pleasurable.

Revision summary: Previously drafted items (e.g., "The haptic feedback felt satisfying.") may fail to properly constrain the scope of the factor, leading participants to erroneously factor in the functional roles. The newly devised items (e.g., "The haptic sensations were enjoyable on its own, regardless of its function."), on the other hand, all explicitly emphasize the constrained scope.

High-Autotelic Example: A relaxation app with a soft pneumatic vest providing gentle pressure sensations for a calming experience.

Low-Autotelic Example: A 4D movie with unpleasant jabs in the back, leading users to avoid the sensation.

3.4 Realism

Realism in HX can be understood as the extent to which sensory stimuli could simulate or represent sensations that one would expect to encounter in reality. In contrast, sensations that feel strange, unnatural or surreal are not considered to be realistic.

Realism encompasses not only faithful re-creation but also intentional design choices that may involve incorporating exaggerated or abstract elements to evoke believable sensations that align with users' expectations and intentions nonetheless. For example, "exaggerated simulated physics (which) has been a staple of digital games" [40], and are also considered realistic and believable in their own way. Haptic feedback is thus high realism when it represents the sensations expected in its context.

Revision summary: Previously drafted items may have been too brief to understand without a specific context (e.g., "The haptic feedback was believable."). We enriched the content of the statements to more clearly convey the meaning and to better reflect this factor of Realism (e.g., "The haptic sensations closely resembled how I feel in real-life.").

High-realism Example: An action-adventure game where haptic feedback closely mimics the tension of a bowstring and the release of an arrow.

Low-realism Example: A messaging app with long, low buzzes that feel strange and unnatural.

3.5 Harmony

In the realm of HX, Harmony emerges as a pivotal factor shaping the overall quality of sensory encounters. Within the context of haptic feedback, Harmony refers to the seamless integration of haptic sensations with other sensory modalities, resulting in a unified and cohesive multi-sensory experience. Previous research has also explored a similar concept known as congruence, described as the intuitive match between different modalities [38].

Revision summary: Many of the previously drafted items did not load onto any factor in previous studies (e.g., "The haptic feedback felt appropriate when and where I felt it."), possibly due to cross-loading with other factors like Involvement. We have revised the items to more closely follow the definition of Harmony (e.g., "The haptic sensations blended seamlessly with the overall experience.").

High-harmony Example: A game where vibrotactile footstep feedback is synchronized with visual and sound effects, matching different terrains.

Low-harmony Example: The same game with vibrations lingering too long and mismatching the character's movements, creating disjointed feedback.

3.6 Expressivity

Expressivity in the context of HX pertains to the system's ability to portray rich variations and convey information through haptic sensations. These variations can represent either user inputs and actions, or system events and status. Expressivity captures the richness and diversity of haptic sensations as well as how well information is communicated through touch.

Revision summary: Reflecting on which of these items did or didn't load in previous studies, we believe that Expressivity, in this context, demands more than mere variations, which the previous items tried to capture (e.g., "I felt adequate variations in the haptic feedback."). We redesigned the factor Expressivity to capture not only the diversity but the meaningful relationship between haptic sensations, user interactions and system events (e.g., "The diversity in haptic sensations allowed me to interpret changes in the system.").

High-expressivity Example: A cooking game where users feel dynamic resistance and texture feedback as they stir a pot, reflecting the ingredients' different levels of viscosity.

Low-expressivity Example: A fitness app with generic vibrations regardless of the exercise performed, lacking specificity and meaningful variation.

3.7 Involvement

In the context of HX, Involvement refers to a psychological state characterized by heightened engagement, focus and attention, as a result of the haptic sensations. Close to what was originally termed "Immersion" by Kim and Schneider [45], involvement primarily focuses on the level of engagement experienced by users. Involvement is closely related to psychological immersion and often regarded as a precursor to it.

Revision summary: No revisions summarized; new items were created, as involvement was introduced with only two items from other intended constructs [3].

High-involvement Example: A VR storytelling experience with subtle vibrations enhancing emotional connection with the plot.

Low-involvement Example: A virtual museum tour with generic haptic feedback that fails to enhance engagement with the virtual environment.

Chapter 4

Pre-testing and Refining Items

This chapter describes the pre-testing and refinement of the initial item pool developed based on the HX model, using expert review and cognitive interviews with novice users. The items were iteratively refined, resulting in a finalized item pool for exploratory factor analysis (EFA). The primary purpose of this step is to identify and eliminate potential misunderstandings, and to ensure that the items were clear, easy to understand, and relevant to the underlying theoretical constructs.

4.1 Expert Review

The items generated from the theoretical constructs were presented to experts in the field of haptic experience or design. These experts reviewed the overall model structure and each of the five constructs individually, as well as evaluated each item's relevance to the theoretical factors, and its clarity and phrasing.

The primary objective of this step is to assess the content validity of the questionnaire being developed. Content validity refers to whether a questionnaire is really measuring what it intends to measure. It covers content's relevance to the theoretical constructs and the representativeness of the items and constructs [8, 36].

4.1.1 Procedures

Experts were recruited through groups or mailing lists related to haptic experience or design, like the Eurohaptics Society mailing list and relevant Slack communities. Interested

individuals were directed to a sign-up survey for basic information including name, email and brief summary of their expertise. The criteria for expertise required participants to be either faculty members, PhD holders or senior PhD students with at least two publications involving evaluating haptic designs, or to hold senior or managerial positions in haptic development, design, or research within the haptic industry. Experience in questionnaire development was not a prerequisite.

Upon verification of their expertise, they were provided access to participate in the study in the form of an online survey. After giving consent, expert participants would be presented with instructions explaining the study procedures and were asked to review each theoretical factor and associated items. The theoretical factors and items within each factor were presented in random order. For each factor, we provided a definition and examples of high and low levels of this factor, followed by the associated items. Experts were asked to rate the relevance and fit of each item (by selecting from 5 ordered categorical options: "Fully relevant", "Mostly relevant", "Moderately relevant", "Somewhat relevant", and "Irrelevant" for relevance, "Ideal", "Suitable", "Acceptable", "Marginal", and "Inappropriate" for fit), and to comment on any issues they identified. They were also asked to comment on each factor and the model as a whole. Each participant was compensated $50 for their time and expertise.

This study and studies in the following sections had been reviewed and received ethics clearance through a University of Waterloo Research Ethics Board (REB #45455).

4.1.2 Results

Responses from 10 experts were collected, exceeding the typical 5-7 experts recruited for similar purposes [8]. Overall, the experts expressed satisfaction with the constructs and model structure, providing positive feedback on the factors and the overall framework. The items received an average of 4.24 (SD=0.35) for relevance, and 3.84 (SD=0.40) for fit, out of the scale of 1-5. Among these, retained items on which the revised items were built upon have an average of 4.44 (SD=0.23) for relevance, and 4.06 (SD=0.31) for fit.

Items with high ratings and no substantive concerns were retained. For items that received critical comments, revisions were made to maximize consensus on a qualitative basis. Some items were removed based on expert feedback due to comprehensibility issues, insufficient breadth of applicability, or questionable relevance to the intended theoretical factor. Revisions focused on improving word choice and phrasing to enhance clarity and precision; these changes are documented in Appendix A and summarized below.

For Autotelics, minor revisions were made to further emphasize that pleasure was derived from the sensations themselves. For Realism, one item was split to separate the concepts of realism and believability, based on expert feedback that these aspects, while related, may be interpreted differently by respondents. For Harmony, many items were revised to simplify language, and to achieve a more consistent reference to multisensory coordination by using "other senses" in place of other more complex or too specific expressions like "other modalities" and "what I was seeing or hearing". For Expressivity, many items were removed due to concerns that they overlapped conceptually with Harmony – not in terms of coordination with other sensory modalities, but rather in relation to actions and events within the system. For Involvement, revisions primarily focused on simplifying language and reducing the use of intensity modifiers. In particular, comparative and superlative forms (e.g., more, -er) were reduced, as experts noted that such wording can impose unnecessary cognitive burden by requiring respondents to make hypothetical comparisons, affecting their responses.

After incorporating expert feedback, we retained approximately ten items per factor based on the severity of issues identified and the degree of redundancy within each construct, resulting in a refined pool of 52 items.

4.2 Cognitive Interviews

Following the expert review, cognitive interviews were conducted to assess face validity, the extent to which end users perceive the items of an assessment instrument as appropriate for the targeted construct and assessment objectives [35]. The interviews also aimed to identify potential issues with item comprehensibility [8, 23, 54].

We targeted individuals with little or no experience with haptic technology for this study. While the target population of the HXI includes users with all levels of experience, the choice was made for two reasons: firstly, the majority of potential users are expected to have limited familiarity with haptic technology; secondly, the appropriateness of the items for experts has been addressed in the expert review stage.

4.2.1 Procedures

Participants were recruited on the university campus through posters specifically advertising for people with little or no familiarity with haptic technology. Participants were scheduled for in-person study session. Upon arrival, participants were informed about the

study procedures and provided consent. They first interacted with a haptic device and were then presented with the items shown in a random order. Participants were instructed to articulate their thoughts and interpretations of each item, especially any points of confusion or lack of clarity. Researchers also actively prompted participants for their interpretation of certain words or phrases when it was not clear. After each round, revisions were made based on the feedback, and participants in the next round would be shown the revised items. Each participant received $20 as remuneration. The devices and applications used in this study can be found in Table 4.1.

4.2.2 Results

Nine participants completed the study in three rounds of three participants each. The participants reported little to no familiarity with haptic technology, and included both native and non-native English speakers. This met the recommendation of "5–15 interviews in two to three rounds, or until saturation" [54]. Saturation was reached in the third round as no new insights emerged.

Feedback collected from the interviews highlighted comprehensibility issues with several items. The participants noted that some terminology was unfamiliar and potentially confusing, and a few words or phrases were considered overly complex. Based on this feedback, revisions were made to simplify the language and improve clarity. For example, "The haptic sensations themselves were pleasant." was revised to "The haptic sensations were pleasant on their own," as some participants found the term "themselves" unclear.

Following cognitive interviews and subsequent revisions, the refined item pool of 50 items proceeds into factor analysis.

Device	Application	Task	Haptic Feedback
3D Systems **"Touch"**	*Jenga demo, provided along with driver.*	Participants use the device arm as the control mechanism to remove blocks from a tall tower made of wooden blocks and place them on top without collapsing the tower.	Force feedback is used to simulate friction and collision forces during the block manipulation.
3D Systems **"Touch"**	*Surface demo, provided along with driver.*	Participants use the device arm to explore the different surfaces and effects associated with different objects or areas in the environment (e.g. moving through liquid, high/low fraction surface, feel the edge of solid object, elasticity effect of a spring).	Force feedback is used to simulate the different surfaces or effects.
Haply **"2diy Gen3"**	*Maze demo, adapted from example in official repo.*	Participants control two on-screen objects using the Haply device and must move them to a designated destination while navigating environmental constraints, including walls and water regions.	Force feedback is used to simulate collision and gravity forces when interacting with the objects and environment.
Ultraleap **"Haptics Development Kit"**	*Sensation Designer.*	Participants use the sensations editor to switch between different types of stimulation or change parameters of stimulation.	Mid-air haptic stimulation is produced accordingly.

Table 4.1: Haptic Devices and Applications Used: This table describes the haptic devices and applications used in in-person studies, as well as user tasks and haptic feedback in these applications.

Chapter 5

Exploratory Factor Analysis

Following the pre-testing and refining of items, Exploratory Factor Analysis (EFA) was conducted to uncover the underlying structure of the item set, enabling the creation of a measurement model with appropriate items. EFA helps researchers determine how different items cluster together, indicating which items measure similar aspects of the concept under study. By examining these patterns, the EFA allows for the refinement of a questionnaire by retaining items that load strongly onto distinct factors and removing those that do not contribute meaningfully.

5.1 Design and Procedure

A survey containing the 50 items was administered to participants both in-person and online. Before completing the survey, in-person participants interacted with a haptic device, while online participants reflected on and reported a recent gaming experience involving haptic components.

In-person study: Participants were recruited in several locations on a large university campus, where they could participate immediately. After giving consent, participants interacted with one randomly assigned haptic device and application for around 5 minutes.

They then completed demographic questions and responded to the 50 items on a 7-point Likert scale, options are "Strongly disagree", "Disagree", "Somewhat disagree", "Neither agree nor disagree", "Somewhat agree", "Agree" and "Strongly agree". Additionally, participants rated their overall haptic experience (*"On a scale of 1 to 100, how would you rate the haptic component of your experience interacting with the device?"*). As remuneration,

they received a coffee voucher valued at approximately $2. The devices and applications used in this study are all four combinations in Table 4.1.

Online study: Participants meeting all the following eligibility criteria were recruited through Prolific.com: (1) played a console or VR game within the past 12 hours; (2) the gaming experience included haptic components; and (3) they have a good recall of this gaming experience. Participants were screened again before entering the survey. After giving consent, they reported the name and device/platform of the game and provided an open-ended description of the best or worst aspect of the haptic experience. The remainder of the survey was identical to that of the in-person study. Participants were compensated GBP£1.5 for their time. Reported devices used include PlayStation (45%), Nintendo Switch (15%), Xbox (12%), and Oculus Quest (10%). The remaining 18% of participants used either Valve Index, PC with controllers, or provided non-specific responses such as "console" or "VR" without specifying the exact devices. A wide range of games were involved, among which most popular ones include GTA (7%), Beat Saber (6%), Fortnite (4%), and Animal Crossing (4%).

Attention checks: Two attention checks — one instructional and one nonsensical — were included to filter out inattentive participants. The description of experience provided is also used to filter out invalid responses. These practices align with recommendations to gather high quality data in online studies [28, 46, 47]. The HXI items and attention checks were mixed and presented in random order. A complete list of HXI items and attention check items can be found at Appendix B.

The attention checks apply to both online and in-person studies. In additional, a threshold of 5 minutes for completion time is also used to filter out participants who likely didn't carefully read all the items before responding. The description of the gaming experience was also checked by the researchers to filter out participants who did not describe any specific haptic experiences. Failing any of these checks would invalidate the data from this participant. Quality of the online sample and the combination of in-person and online samples is further discussed in subsection 8.1.2.

5.2 Participants

In-person study: A total of 152 participants completed the in-person study, with a gender distribution of 60.53% male, 36.84% female, and 2.63% non-binary)[1]. The majority were

[1]A sum of less than 100% means some participants prefer not to say, same for below.

ages 18-24 (89.47%) and received some university education without degree (77.63%, e.g., current undergraduate students).

Online study: A total of 208 participants completed the online study, with a gender distribution of 59.62% male, 36.54% female, and 2.88% non-binary. The majority were ages 18-24 (39.42%) and 25-34 (42.79%). Education backgrounds varied, with 21.63% completed high school or less, 25.96% with some university, and 51.44% holding a bachelor's degree or higher.

5.3 Analyses

To assess the suitability of data for factor analysis, Bartlett's test of sphericity [5] was used to test that the correlation matrix was not random, and the Kaiser-Meyer-Olkin (KMO) [44] was calculated to test sampling adequacy. Skewness and kurtosis were also tested to detect extreme skewness or kurtosis [82].

After confirming the suitability of the data, EFA was conducted. Eigenvalues, parallel analysis [39], and the visual scree test [11] were used to determine the appropriate number of factors to retain. Theoretical convergence were also considered.

Common factor analysis was selected over Principal Component Analysis (PCA) as our intent was to identify the latent factor structure [25, 82]. Iterated principal axis (PA) was selected for extraction, as it allows a more precise estimate of communalities that are iterated until a final solution is reached, and for PA being a least-squares estimation method that made no distributional assumptions, while ML is sensitive to normality which likely won't hold up according to previous HX questionnaire studies [3, 77, 82]. Promax, an oblique rotation method [37], was chosen as inter-factor correlations are expected by both theoretical assumptions and prior work [3, 45, 65].

Factors with a minimum of three salient pattern coefficients, internal consistency reliability $\geq$.70, and that were theoretically meaningful were considered adequate [82].

5.4 Results and Discussion

Participant Filtering and Data Cleaning: In the in-person study, 15 were filtered out for failing either of the attention checks. In the online study, 44 were filtered out on following criteria: (1) not providing a detailed description of a game experience with haptic

components (N=14); (2) failed either attention check (N=16); (3) completing the survey in less than 5 minutes (N=14). Combined, the final sample consisted of 302 participants. All negatively phrased items were reverse-coded to ensure consistency in the analysis.

Data Suitability and Descriptives: The data was suitable for factor analysis. Bartlett's test of sphericity indicated that the correlation matrix was not random, $\chi^2(1225) = 11{,}482, p < .001$. The KMO statistic was 0.96, well above the minimum standard of 0.7, indicating sampling adequacy. While there were no extreme univariate skewness nor kurtosis, Mardia's multivariate skew and kurtosis were both significant ($p < .001$), indicating non-normality. Means for each item range from 4.10 to 5.88. Standard deviations for each item range from 0.78 to 1.78. The means and standard deviations for all items can be found in Appendix C.

Determining Factor Number: The number of factors were determined using a combination of eigenvalues, scree plot analysis, parallel analysis and theoretical constructs. Eigenvalues suggested five factors with eigenvalue greater than 1. The visual scree test suggested five or six factors. Parallel analysis, shown in Figure 5.1, indicated five factors, or potentially six, with the sixth factor having an eigenvalue of 0.59 compared to the eigenvalue of 0.60 for the sixth simulated factor. Combined with considering the five-factor theoretical model, we explored both five- and six-factor models.

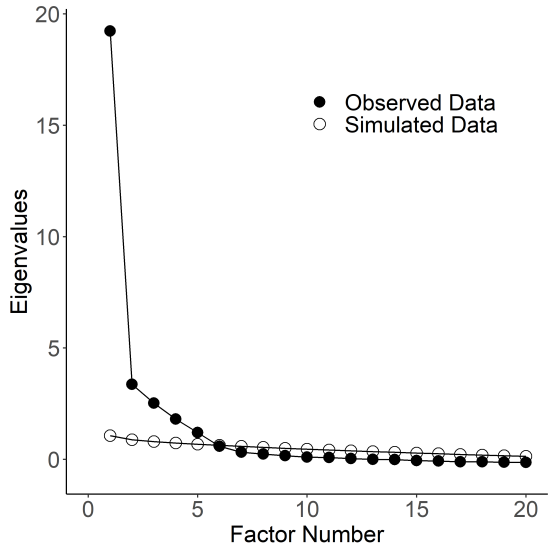

Figure 5.1: Parallel analysis scree plot: result suggests 5 or 6 factors, with simulated and observed eigenvalue overlapping for the 6th factor

EFA Results: EFA was conducted starting with six and five factors respectively, fol-

lowed by an iterative item selection process. Items with weak primary loadings or high cross-loadings were removed using a threshold of at least 0.4 for primary loading, and a primary loading at least 0.2 greater than the next highest loading [8, 82]. Items within a factor are also further selected based on their loading and phrasing to reduce redundancy when needed [22]. Redundancy is evaluated by the researchers based on the phrasing of the items. For example, I7 "The haptic sensations supported my overall engagement with the system." is dropped as it is too similar to I6 "I found the haptic sensations strengthened my engagement with the system."

With the initial extraction of five- and six-factor models, all items except R5 loaded in accordance with their theoretical constructs. Items from the theoretical factor of Expressivity initially loaded onto a distinct factor in both models but were subsequently removed due to cross-loadings, yielding final five- and four-factor solutions.

The five-factor solution (Figure 5.2) retained the same four factors as the four-factor solution, with the addition of a factor composed of a subset of negatively phrased items (e.g. "The haptic sensations felt out of sync with the other senses.") originally intended to measure Harmony. In contrast, the positively phrased Harmony items consistently loaded onto a single factor across both solutions.

This additional factor demonstrates clear theoretical distinction from the positively phrased Harmony construct. It also shows strong psychometric properties, with substantial factor loadings (0.647–0.752) and high internal consistency.

Model fit indices were compared between the two models. The four-factor solution had a Comparative Fit Index (CFI) of 0.981, Tucker–Lewis Index (TLI) of 0.976, and Root Mean Square Error of Approximation (RMSEA) of 0.045. The five-factor solution had slightly lower CFI (0.978) and TLI (0.974), and also a lower RMSEA (0.043). These differences were not significant, indicating that both models had similarly good fit [2].

Based on these considerations, we decided to include this additional factor. We retained the original names of the theoretical factors for most factors derived from the EFA: Autotelics, Realism, and Involvement. The additional factor with negatively phrased items from Harmony was named Discord, while the factor with positively phrased Harmony items retained the name Harmony. Figure 5.2 presents the EFA model and item loadings.

EFA was also conducted separately on the online and in-person samples, yielding consistent results. The same factor structure emerged across both conditions, and the items selected in the general model exhibited satisfactory loadings when analyzed separately

[2]The meanings of the indices are explained in Section 6.3, as they are more vital to CFA and optional here

for the online and in-person samples. The chosen five-factor solution demonstrated good model fit for both the online (CFI = 0.973, TLI = 0.968, RMSEA = 0.049) and in-person (CFI = 0.966, TLI = 0.960, RMSEA = 0.054) sample. The full results and factor loadings are not presented as they show minimal differences from those reported above and would substantially increase the length of the paper. Furthermore, concerns regarding the applicability of the proposed model to the online and in-person conditions are addressed through measurement invariance analysis, as detailed in Section 7.3.

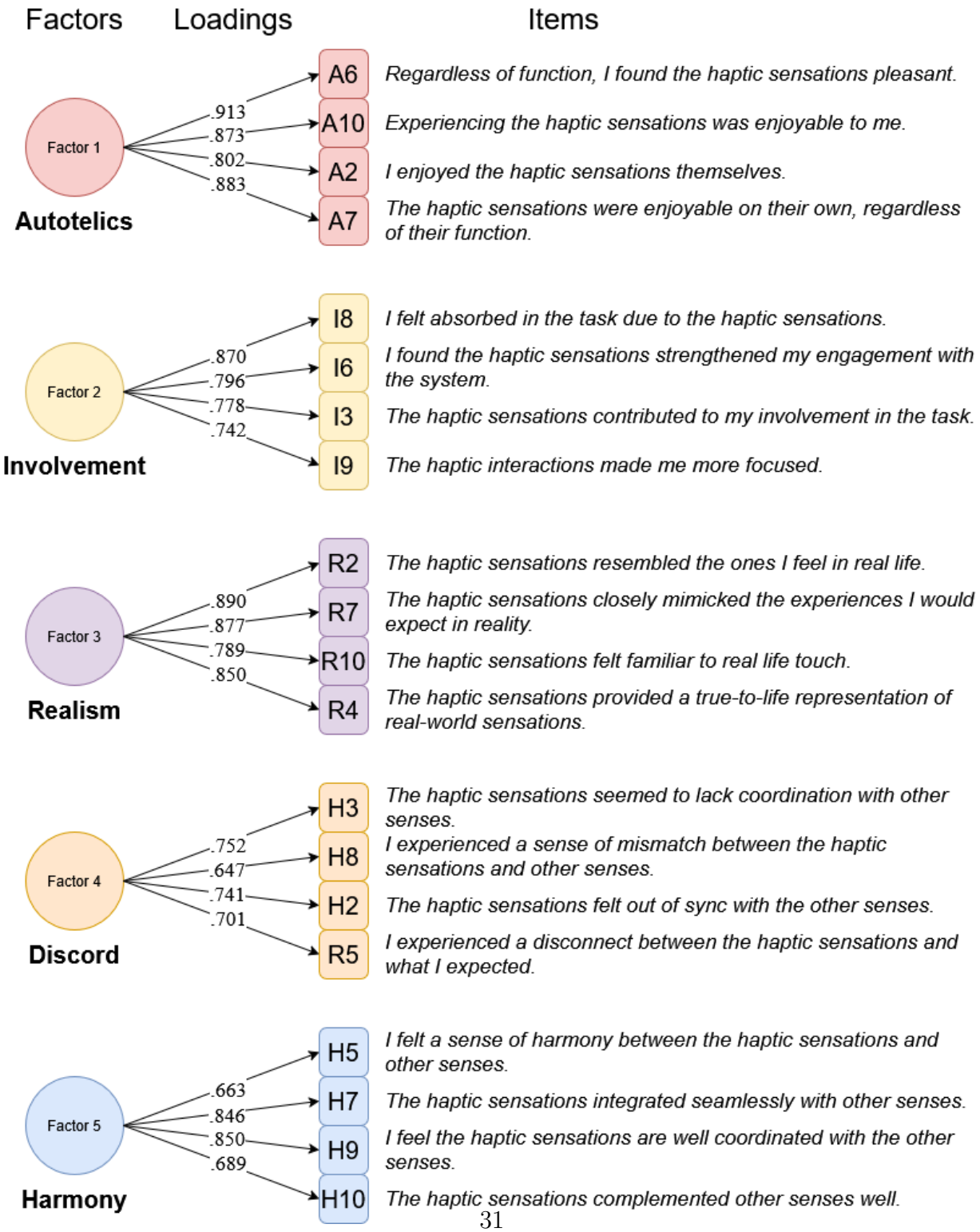

Figure 5.2: Item loadings for EFA: the five factors each have 4 items with high loadings and the pattern is consistent with the theoretical model

Chapter 6

Confirmatory Factor Analysis

Following exploratory factor analysis (EFA), confirmatory factor analysis (CFA) is used to validate the factor structure identified in EFA by confirming if the items group into the expected dimensions when tested using a new set of data. It provides evidence that the measurement model generalizes beyond the exploratory sample rather than reflecting patterns only observed in the EFA sample.

6.1 Design and Procedure

The CFA step was conducted using a design and procedure similar to the EFA, with both in-person and online participants recruited in identical ways and following similar procedures. The only differences are in the content of the questionnaire, with reduced number of items for HXI resulted from removal of items in EFA.

An eight-item UEQ-S [70] was also included at the end of the questionnaire, presented in random order, and participants are asked to "rate the haptic sensations" for UEQ-S. The UEQ-S was included to assess the HXI's convergent and discriminant validity while keeping the study length manageable due to its concise format. As an established and widely accepted measure of general UX across diverse applications, the UEQ-S is well-suited for testing convergent validity. The UEQ-S includes two factors, Pragmatic and Hedonic, which can offer insights into how HXI factors correlate with these aspects, while also providing an opportunity to examine discriminant validity. A small proportion of participants completed the HXI, but did not complete the UEQ-S scale. A quota for gender with male participants capped at 50% was also implemented in the online study to achieve a more balanced gender distribution.

6.2 Participants

In-person Study: 93 participants completed the in-person study. Gender distribution was 60.22% male, 38.71% female, 1.08% non-binary. Most participants were ages 18-24 (92.47%) and received some university education without degree (83.87%).

Online study: 138 participants completed the online study. Gender distribution was 50.00% male, 44.93% female, 5.07% non-binary. Most participants were ages 18-24 (39.86%) and 25-34 (31.16%). Education backgrounds varied, with 15.94% who completed high school or less, 34.06% with some university, 49.26% with bachelor's or higher degree.

6.3 Results and Discussion

Study 4 used the same process for participant filtering and data cleaning. Six were filtered out for failing either of the attention checks in the in-person study. In the online study, 27 were filtered out on the following criteria: (1) not providing a detailed description of a game experience with haptic components (N=7); (2) failed either attention check (N=14); (3) completing the survey in less than 5 minutes (N=6).

The measurement model derived from the EFA study demonstrated excellent goodness-of-fit: $\chi^2(160) = 214.6$, below the conservative limit of twice the DoF. Alternative fit indices also supported the model's fit. The Comparative Fit Index (CFI), which measures the model's fit with respect to a more restrictive baseline model, was 0.981. The Tucker-Lewis Index (TLI), a more conservative measure similar to the CFI that penalizes complex models, was 0.977. Both indices are above the commonly used threshold of 0.95 [8], indicating good fit. The Standardized Root Mean Square Residual (SRMR), which measures the difference between the observed correlation and the model implied correlation matrix, was 0.042. This value is below the threshold of 0.08, suggesting good fit. The Root Mean Square Error of Approximation (RMSEA), which measures the discrepancy between the population and model-implied population covariance matrices, was also 0.042. This value is below the threshold of 0.05, indicating close fit. The items loadings are all satisfactory, with standardized loadings ≥ 0.86.

CFA was not conducted separately for the online and in-person samples, primarily due to the larger sample size requirements of CFA procedures. Instead, measurement invariance analysis was employed to assess the applicability of the model across both conditions, as detailed in Section 7.3.

Chapter 7

Psychometric Evaluation

Following the CFA study, additional psychometric assessments were conducted, including internal consistency reliability, convergent, divergent and predictive validity, as well as measurement invariance, to further examine the robustness of the model.

7.1 Internal consistency reliability

Internal consistency is a measure of reliability. It measures whether items measuring the same construct produce similar scores, and is usually measured with Cronbach's alpha [17]. Using the data collected in the CFA study, all factors demonstrated good internal consistency reliability. Autotelics, Involvement, Realism, Harmony and Discord have a Cronbach's α of 0.92, 0.87, 0.92, 0.90, 0.84 respectively. The overall Cronbach's α is 0.94. These are all well above the acceptable threshold of 0.7 [58].

7.2 Convergent, divergent and predictive validity

Convergent and predictive validity were also assessed. Convergent validity refers to how closely a test is related to other measures of the same or similar constructs, and is assessed by through the correlation between scores of HXI, UEQ-S and their subscales. Predictive validity refers to how well a score on a scale or test predicts scores on some criterion measure [18], and is assessed by correlation between the score of HXI and participants'

	AUT	INV	REA	DIS	HAR	UEQ-P	UEQ-H
Autotelics	1.000					0.595	0.521
Involvement	0.610	1.000				0.451	0.510
Realism	0.377	0.447	1.000			0.330	0.573
Discord	0.455	0.455	0.346	1.000		0.573	0.363
Harmony	0.589	0.648	0.551	0.644	1.000	0.578	0.546

Table 7.1: Pearson Correlations Between HXI and UEQ-S Subscales

overall rating of the haptic experience. Data used are participants in the CFA study that completed both the HXI and the UEQ-S scales.

The correlations are presented in Table 7.1. The five factors within the HXI model shows significant inter-correlations (0.377 – 0.648). The overall HXI scores have a correlation of 0.748 with UEQ-S. Furthermore, the five HXI factors correlated with both the Pragmatic (0.330 – 0.595) and Hedonic (0.363 – 0.573) subscales of the UEQ-S, supporting convergent validity.

Divergent validity was also primarily supported by the pattern of correlations between the HXI factors and the UEQ-S subscales. For example, Realism has the highest correlation (0.573) with the Hedonic UEQ-S subscale while having the lowest correlation (0.330) with the Pragmatic UEQ-S subscale, aligning with realism being often non-functional and optional. Also, Discord had a lower correlation with the Hedonic subscale (0.363), compared to both its correlation with the Pragmatic subscale (0.573) and Harmony's correlation with the Hedonic subscale (0.546), consistent with the idea that Discord often has functional implications while Harmony are more closely related to Hedonic.

HXI also demonstrates good predictive validity through a correlation of 0.817 with the overall rating of the haptic experience, suggesting that the HXI is a good predictor of the overall satisfaction with haptic experience. As the UEQ-S scores and overall ratings are collected at the same time as HXI, the correlations also support concurrent validity.

7.3 Measurement invariance

Measurement invariance assesses whether this instrument is interpreted in the same way across different groups [32, 63]. Measurement invariance across different haptic modalities was tested using combined samples from the EFA and CFA study to ensure adequate sample size. Since online studies use vibrotactile devices, while in-person studies employed

	df	χ^2	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
Configural	480	655.51	0.974	0.047	0.042			
Metric	510	701.82	0.972	0.048	0.049	-0.002	0.001	0.007
Scalar	540	771.47	0.966	0.051	0.050	-0.006	0.003	0.001

Table 7.2: Measurement Invariance Testing Results.

mid-air and force-feedback devices, the comparison between haptic modalities also reflects the difference between online and in-person study forms.

There are four steps for measurement invariance [63, 83]: (1) configural: equivalence of the model form (structure); (2) metric: equivalence of factor loadings; (3) scalar: equivalence of item intercepts; and (4) residual: equivalence of item residuals. In our analysis, achieving the first two levels of measurement invariance would indicate the HXI's model structure and loadings hold well across different groups. The first three levels are tested, and the residual invariance test is omitted since it's not part of the latent factor and is inconsequential to the interpretation of latent mean differences [63, 81].

Each invariance level was evaluated by comparing the fit of a more constrained model against its less constrained predecessor (i.e. adding the constraint of loadings, intercepts, and residuals in the steps described above). Model fit was evaluated using changes in fit indices rather than relying on the χ^2 test, which is often overly sensitive to small deviations from a "perfect" model [63]. Model fit was evaluated using change thresholds of ΔCFI $\leq$ .01, ΔRMSEA $\leq$ .015, and ΔSRMR $\leq$ .030 (metric invariance) or $\leq$ .015 (scalar or residual invariance) [12, 63]. When changes in fit indices fall within these thresholds, the additional constraints are considered tenable, supporting invariance at that level.

The results are presented in Table 7.2. The configural model tests the general model fit across different groups. The model achieved good model fit, supporting configural invariance. In the metric model, factor loadings are constrained. The changes in all fit indices are below the criteria, supporting metric invariance. In the scalar model, intercepts are constrained in addition to constrains in the metric model. Again, the changes in all fit indices are below the criteria, supporting scalar invariance.

Chapter 8

Discussion

We discuss the overall outcome of the developed HXI, insights from the resulting five-factor structure, as well as limitations and future work.

8.1 HXI: A Robust Questionnaire Ready for Haptic Experience Assessment

The development of HXI followed a rigorous and systematic process, ensuring the robustness and reliability of the instrument. The HXI has demonstrated a strong underlying theoretical model, internal consistency reliability, convergent and predictive validity, and measurement invariance across different contexts. This comprehensive development process provides a solid foundation for the HXI, establishing it as the first fully validated questionnaire specifically designed for assessing HX.

This new instrument can benefit researchers and developers by providing an easily deployable tool that offers insights into the most predominant aspects of experiences with haptic technology. As a standardized quantitative tool, it would also enable direct comparisons across different iterations of the same product, or across different products, possibly even those from different modalities. For example, when developing a new haptic application, the HXI could be deployed to evaluate whether there are substantial improvements after major changes. It could also be compared to the HXI scores of previous similar applications, enabling the developers to quickly gain a sense of how well are they doing, and identify if any of these factors requires more attention. While future work is needed to see

if benchmarking might be possible, the ability to compare and isolate aspects of HX can aid in both research and practice.

8.1.1 HXI is better tailored for HX than general UX questionnaires

Hapticians have previously expressed a need for "something similar [to existing questionnaires] that focuses on our needs" [45]. General UX questionnaires like UEQ and AttrakDiff, while well-established and useful for general UX, are not designed for the nuanced assessment of haptic interactions and often do not fit well for measuring HX. For example, some items in UEQ [48] may not apply in the context of HX: "secure – not secure", "impractical – practical", "friendly – unfriendly", "easy to learn – difficult to learn".

In contrast, the HXI was developed specifically for haptic interactions, with its items intended to fit various haptic contexts. The factor structure of HXI also aligns with existing HX research [45], captures the unique aspects of HX that are not fully addressed by general UX questionnaires. The HXI also showed slightly higher predictive power compared to the UEQ-S.

8.1.2 HXI Applies to Various Haptic Contexts

The stability of the HXI's factor structure across different forms of administration (online and in-person) and haptic modalities (e.g., vibrotactile, mid-air, force-feedback) demonstrates its versatility and reliability. The ability of the HXI to maintain consistent factor loadings and item intercepts across these varied conditions supports its applicability in diverse haptic contexts. This consistency is crucial for researchers and practitioners who need a reliable instrument that can be used across different studies and applications, ensuring that the HXI is a robust measure for capturing HX in various environments.

It is worth noting that in-person and online groups in our study differed in several ways laid out below. The measurement invariance across these contexts supports the ecological validity of the HXI, indicating that it can reliably capture HX across diverse settings and populations.

First, there are differences in the type of devices and the duration of the interactions. The interaction duration with the haptic device in our in-person studies was only a few minutes, which may not provide an in-depth haptic experience. In contrast, participants typically spend significantly more time, sometimes hours when playing games, allowing for

a more comprehensive haptic experience. The devices used in our in-person studies (force feedback and mid-air haptic devices) are often novel to most participants, which may have lead to a novelty effect that could affect the way participants experience these interactions. Conversely, participants in the online study were likely already familiar with the haptic sensations in the games.

Second, participants recruited on campus are predominantly university students, with the majority falling into the 18-24 age group, whereas the sample recruited online are more demographically diverse in terms of age, education level, employment status, and residing country. This contributed to a more diverse and representative sample.

Lastly, there is evidence supporting quality and reliability of the data collected online for haptic interactions. Research in other fields has shown the recalling performance of experienced events after 24 hours is still pretty accurate [15]. Furthermore, prior work suggests remote vibrotactile studies can yield similar ratings to in-person studies [69]. Questionnaire development based on online reported evaluations of experience has also been done on (game) player experience [66]; in our case, in-person experiences were necessary to test the scale on rarer modalities. Our results further support the validity of online studies with haptic feedback from the description of their experience with specific details. For example, one participant reported that *"(she) really enjoy haptics in gaming as it contributes significantly to my experience. For example, when the cat in the game gets startled by a human or by water, the strong vibration is a nice touch as it serves as an indicator that the action is negative, which is definitely satisfying."*

8.2 Understanding the Factor Structure of Haptic Experience

The factor structure of the HXI largely aligns with theoretical expectations. The identified factors reflect key dimensions of haptic interaction that are important for understanding how users perceive and evaluate haptic feedback. Two major changes observed and adopted during the questionnaire development process were the splitting of Harmony into Harmony and Discord, and the exclusion of Expressivity.

8.2.1 Harmony and Discord: Two Sides of Harmony

We found that positively and negatively phrased items under the theoretical factor Harmony loaded onto two separate factors: one describing the achievement of Harmony and

the other describing the lack of Harmony. The factor describing the lack of Harmony was named Discord. This discovery aligns with the previous HX questionnaire development attempt [3], where positively and negatively phrased items under Harmony loaded onto factors composed purely of positive or negative items, respectively. However, due to the number of items in their study, the patterns are not clear as there are only two or three items per factor.

In our study, the distinction is clear, with most items from Harmony loading in this manner onto Harmony or Discord. This reflects a more nuanced understanding of HX than previously proposed. While Harmony captures the pleasurable and seamless aspects of haptic interactions, Discord represents the less desirable or jarring elements that can occur. This also relates to the notions of hygiene and motivators, as adapted to UX [78]. Hygiene contributes to dissatisfaction but not satisfaction, while motivators contribute to satisfaction but not dissatisfaction; in other words, hygiene must be addressed to avoid a poor experience while motivators can improve experiences after hygiene is met. This is consistent with prior work in haptics that suggest the absence of Harmony may be more noticeable, while "good haptics are subtle and less prominent" [45, 67]. This is supported by the fact that Harmony and Discord both correlate in a similar manner with the pragmatic subscale of UEQ-S, but Harmony has a stronger correlation with the hedonic subscale, indicating Harmony could be more closely related to hedonic aspects, possibly acting as a motivator when hygiene is met.

8.2.2 Expressivity: System Characteristic Rather than Experiential Factor

Expressivity was excluded from the final HXI model. The items from Expressivity did show a pattern of loading onto a factor, but were all subsequently dropped due to severe cross loading. This suggests that this dimension might be on a different level from other factors.

We believe this could be due to Expressivity being more relevant to the design process rather than the user experience. This leads to Expressivity being more of a design parameter and system characteristic than one of the dimensions of HX. While future assessment tools or modeling might include Expressivity as a way of assessing a haptic system's capabilities, excluding Expressivity from the HXI ensures that the instrument remains tightly aligned with its intended purpose of assessing HX from the user's perspective.

Chapter 9

Limitations

Despite the strengths of the HXI, there are several limitations that future research should address. This chapter addresses important limitations and future directions for the project.

9.1 Exploring More Modalities and Contexts

While different haptic modalities and contexts were used in this study, it's still a limited subset of all available modalities and applications. For example, neither surface haptic devices [6] or passive VR props [4] are represented in this study. In our study, we used three of the most commonly applied or studied modalities. Force-feedback is one of the longest running modalities of haptic feedback, frequently chosen as a focal point for mature haptic feedback. For example, the Haptipedia project targeted it in part because it has "considerable maturity in both research and commercial settings" [71]; grounded force feedback devices are similarly the target of recent complementary work in evaluation like the Haptify benchmarking effort [26]. Vibrotactile feedback is commonly found in commercial devices ranging from pager motors to game controllers, and the target of many design tool efforts in research [68, 72]. Mid-air haptic feedback has a long history of academic research [10], including design tools [43, 76], design guidelines [29], and, along with vibrotactile feedback, early measurement of user experience with haptics [52]. As such, the HXI should be suitable for the most commonly studied modalities in Haptic Experience.

Furthermore, because the HXI is interpreted consistently across these modalities, which differ significantly in their sensory mechanisms and physical sensations, we believe there are reasonable grounds to be confident in HXI's applicability to more modalities and applications not covered in this study. Nevertheless, future work is important to empirically

confirm such applicability. This will help ensure the reliability and validity of the HXI across an even wider spectrum of haptic technologies.

9.2 Further Testing of HXI's Validity

Our work established several important forms of validity for the HXI, including convergent, predictive, and concurrent validity, as part of the HXI's initial validation. Still, further testing of the HXI's validity is warranted to provide further and more comprehensive validation.

Discriminant validity: To assess discriminant validity, we examined correlations between the HXI and the Pragmatic and Hedonic factors of the UEQ-S. While this provided a preliminary exploration, it may not fully demonstrate the HXI's discriminant validity. A key challenge for further investigation is selecting an appropriate questionnaire capturing a construct different from HX but is still of interest to researchers. While we considered using scales measuring usability, the overlap between the HXI's experiential focus and certain usability elements suggests that they may not provide the "different construct" needed for rigorous discriminant validity testing. Further research could identify suitable comparison measures to more thoroughly establish discriminant validity.

Criterion validity: For criterion validity, we used participants' overall ratings of the haptic components in the interactions, constrained by our study design. This subjective rating may lack ecological validity. Future work could incorporate objective measures, such as purchase intention for haptic products or retention rates in haptic applications, to provide an examination of criterion validity that is more robust and has greater ecological validity.

Test-retest validity: Test-retest validity was not assessed in this study. Despite test-retest validity being commonly used for stable traits such as personality, we believe it is not applicable to the HXI as haptic experiences are not expected to be stable over time, but rather likely to change over time. Haptic experiences are inherently dynamic, subjective, and context-sensitive. Factors such as mood, recent interactions, and environmental context can all affect one's experience of haptic feedback and complicate efforts to capture stability over time. Moreover, in interactive systems, replicating identical interaction conditions across sessions is difficult, as user inputs can alter system events and responses. Consequently, we believe test-retest methods commonly used for stable traits, such as personality, may not be suitable for UX instruments like the HXI. For instance, Haider et al. [31] assessed the test-retest reliability of the miniPXI, which measures player experience.

The test-retest reliability of some factors fell below conventional "acceptable" thresholds, which the authors attributed to the aforementioned challenges. An alternative approach considered was administering the HXI twice on the same experience at different time points (e.g., immediately post-interaction and after a delay). While this could reveal temporal effects on HX and provide insights into HXI's stability over time, it diverges from HXI's intended temporal scope, which is in-the-moment assessment immediately or shortly after interaction. Although further research on test-retest validity could provide valuable insights, it would likely require methodological adaptations and an expanded study design to account for the inherent variability in user experience.

9.3 Recommendations on Reporting for Continual Validation of the HXI

We strongly encourage future studies involving the HXI to include the reporting of its Cronbach's alpha, as well as how the HXI relates to other measurements (e.g., other UX questionnaires, overall haptic experience ratings, application logs, and psychophysiological measures like electrodermal activity and eye-tracking data). These complementary measures could help further test the validity of the HXI and contribute to understanding which aspects of the haptic experiences it best captures. Additionally, identifying measures that complement the HXI could provide a fuller picture of the Haptic Experience, guiding improvements in both research and product development.

9.4 Segment-Specific Insights in Assessing HX

When applying the HXI, researchers and developers should carefully consider the scope of what is being evaluated. Some participants in cognitive reviews reported varying levels of agreement with certain items depending on different segments of the experience. Future studies could explore how to best utilize the HXI regarding the evaluation scope (e.g., different scenes in a game, distinct phases of an interaction, subsets of feedback effects).

For experiences that involve diverse components, it may be beneficial to include brief interviews or open-ended questions that allow users to express such concerns, at least in the pilot study stage. This approach would enable those using the questionnaire to determine the appropriate scope and consider segmenting different parts of the experience based on user feedback, which could improve the clarity and relevance of the evaluation.

9.5 Leveraging HXI to Enhance Design Processes

There is a need to explore how the HXI can better support the design of haptic experiences. By understanding the specific aspects of HX that contribute to user satisfaction, designers can use the HXI to create more effective and enjoyable haptic interactions. Future research could help establish best practices to how and when the HXI can be integrated into the design process to provide actionable insights that enhance the quality of haptic feedback and its alignment with user expectations.

Chapter 10

Administering HXI

The Haptic Experience Inventory (HXI) is the first validated tool to measure Haptic Experience across five core factors: Autotelics, Realism, Harmony, Discord, and Involvement. Each factor represents a unique dimension of the user's experience with haptic interactions, offering a comprehensive framework for assessing the quality and impact of haptic feedback. This section outlines its scope, structure, administration, scoring, interpretation, and reporting guidelines for the HXI. While some content may overlap with prior sections, this segment consolidates all relevant information to ensure clarity.

We also recommend checking online instructions and aids for deploying the HXI at https://get-hxi.org/ for the most up-to-date guidance and best practices. Future deployments and shared results will continue to refine our understanding of the questionnaire's performance and enhance its utility in diverse contexts.

10.1 Scope of the HXI

HXI is designed to be broadly applicable across different haptic modalities, though caution advised when applying it to new modalities. While only three (vibrotactile, mid-air, force-feedback) are covered in our study, the consistency of the HXI's factor structure, loadings and item intercepts across these modalities, which vary significantly in their mechanisms, supports its diverse applicability. Nevertheless, for modalities not covered in the study, we recommend exercising caution. Researchers should adhere to the complete form of the HXI and report reliability metrics, as well as other relevant measures, to further examine its suitability.

HXI is intended for administration immediately or shortly after the haptic interaction. In our study, the HXI was administered within 12 hours for online participants and immediately after the interaction for in-person participants. Prolonged recall periods may compromise the reliability of results from the HXI.

10.2 Structure of the HXI

The HXI measures five factors through 20 items, with 4 items for each of the five factors.

Autotelics – Evaluates the pleasure derived from haptic sensations independent of functional value.

Realism – Measures how closely haptic sensations simulate or represent real-world experiences.

Harmony – Assesses the seamless integration of haptic feedback with other sensory modalities.

Discord – Identifies aspects of haptic feedback that are disjointed with other senses.

Involvement – Captures the level of engagement and focus haptic sensations create.

10.3 Form of the HXI

Presentation: Items should be randomized to avoid order effects.

Instructions: Before administering, briefly explain terms such as "haptic sensations" and "other senses" to avoid confusion. We used the following definitions:

"Haptic sensations": *Anything that you feel with the sense of touch. It could be vibrations, force, temperature, pressure, or any other physical sensation.*

"Other senses": *Sensory modalities beyond touch, such as vision and hearing.*

Response Format: Participants rate each item on a 7-point Likert scale from 1 (Strongly Disagree) to 7 (Strongly Agree) based on their agreement with each statement.

On using selected subscales: Sometimes people use only a subset of an established questionnaire, but there have been cautions against doing so and one study showed psychometric properties may change when only using a subset of a questionnaire [27]. We recommend using the complete standard form of the HXI if possible.

10.4 Scoring the HXI

For each item, the seven options should be scored from 1 (Strongly Disagree) to 7 (Strongly Agree). We recommend against scoring items on a scale of -3 to 3, as a score of 0 does not correspond to a meaningful condition, which may lead to interpretive inconsistencies.

Factor Scores: Average the scores of the four items for each factor to obtain individual factor scores.

General Score: Reverse the score for Discord first, then add it with the other four factors scores. Thus, General Score = Autotelics + Involvement + Realism + Harmony + (8 - Discord). This provides a comprehensive view of the overall haptic experience.

10.5 Interpreting Scores

Context-specific Priorities: The general score offers an overall assessment, and each factor score provides insight into specific attributes of the haptic experience. It is important to note, though, that haptic applications have different goals and contexts, and the relative importance of each factor may vary across each product. This was also echoed in expert interviews. Depending on the context, not all factors are targeted equally. For example, a relaxation app may prioritize Autotelics, while haptic design for a game may prioritize Involvement.

Comparative Analysis: The HXI enables quantitative comparisons across different variations or iterations of a product to highlight the impact of haptic design improvements when applicable. Measurement invariance is also established for the HXI, supporting comparisons between different products, even if they are using a different haptic modality! Choosing proper products to compare against is important to produce the most useful insights.

10.6 Reporting HXI Results

Reliability Metrics: It's recommended to report Cronbach's alpha for each factor to demonstrate scale consistency. Some rules of thumb recommend values > 0.6 or > 0.7 as sufficient. If the alpha coefficient of a factor is low, you should interpret the results with caution, it may indicate that some items within the factor are misinterpreted by many of the participants, or these items are not relevant to the experience being evaluated. We

also recommend that you share such findings with us to facilitate better understanding of HXI when the Cronbach's alpha is low.

How the HXI relates to other measurements: Reporting how the HXI relates to other measurements (e.g., other UX questionnaires, overall ratings, and psychophysiological measures like electrodermal activity and eye-tracking data) could help the community gain a better understanding of the ecological validity of the HXI and how to best use HXI in practice.

Chapter 11

Ongoing Work & Conclusion

Ongoing work is underway to produce a high-quality French version of the Haptic Experience Inventory (HXI). This project follows established best practices for questionnaire translation and cross-cultural adaptation and consists of three main stages: (1) translation of the HXI into French, (2) pilot testing of the translated version, and (3) validation and psychometric evaluation of the French HXI.

11.1 Translation of the HXI into French

A forward–backward translation approach was employed to produce a high-quality French version of the HXI. This process consists of forward translation, backward translation, and reconciliation by an expert committee.

In the forward translation stage, two independent French versions of the HXI were produced. The first version was developed by collaborators on this project through an iterative process that involved internal discussion and external feedback. These collaborators are fluent in both English and French, reside in Quebec, and have experience conducting research in haptics, making them familiar with both the domain and the intended constructs of the HXI. The second version was produced by an independent translator based in and certified (Expert Traducteur près la Cour d'Appel, expert translator at the Court of Appeal) in France, who was not familiar with haptic research or the content of the HXI, providing a complementary perspective focused on linguistic naturalness and general comprehensibility.

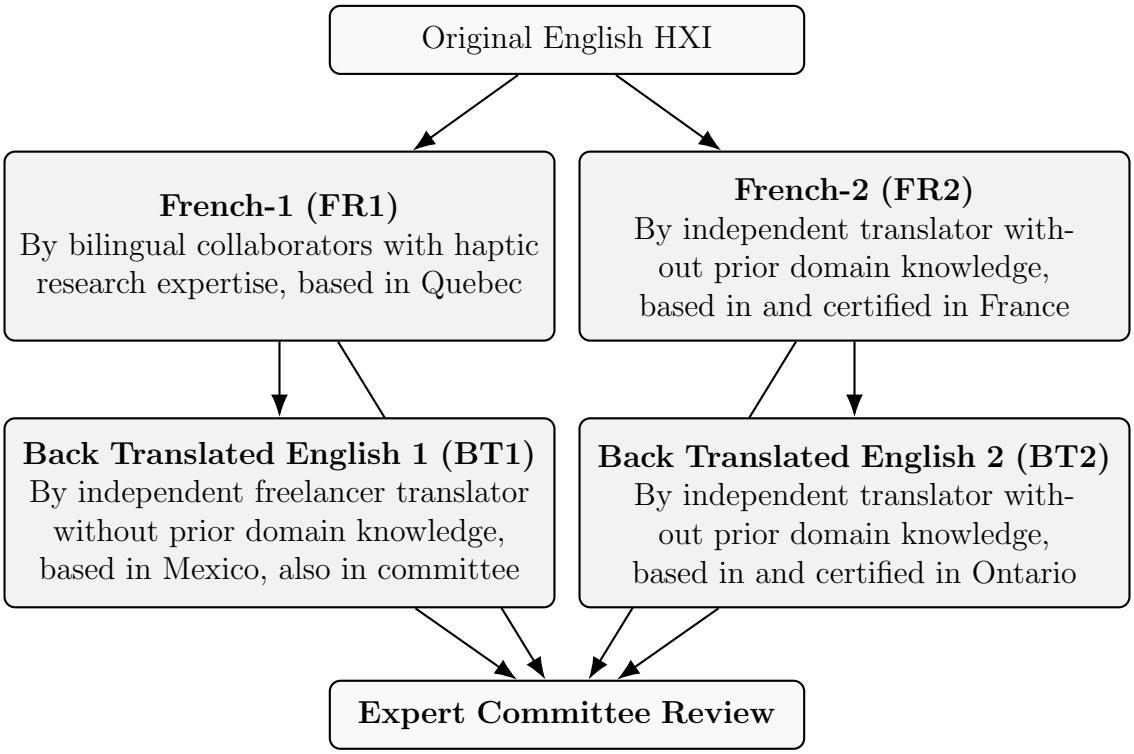

Figure 11.1: Translation workflow

In the backward translation stage, each of the two forward-translated French versions was independently translated back into English by a certified translator. Both back-translators were from a Canadian French (Quebec) language background and had no prior knowledge of the original instrument or the intent of the translation, in order to minimize bias.

All versions produced during the forward and backward translation stages were then reviewed by an expert committee for reconciliation. The committee consisted of the two collaborators involved in producing the first forward translation; the two authors of the original HXI; an external haptics researcher fluent in both English and French; a researcher with expertise in questionnaire translation methodology; and a professional translator involved in one of the backward translations. The committee reviewed discrepancies across versions and resolved them through discussion, with the goal of achieving semantic, idiomatic, experiential, and conceptual equivalence. This process resulted in a single pre-final French version of the HXI, ready for pilot testing.

11.2 Pilot Testing of the Translated Version

Following the translation stage, cognitive interviews will be conducted to pilot test the pre-final French version of the HXI. This step aims to assess face validity, defined as the extent to which end users judge the items of an assessment instrument to be appropriate for the targeted construct and assessment objectives, and to identify potential issues with item comprehensibility.

Participants fluent in French will be recruited from a Canadian university campus in Quebec. Both native and non-native French speakers are included to ensure that the questionnaire could be interpreted consistently by individuals with varying levels of French proficiency. Participants first interact with a haptic device and are then presented with the questionnaire items in random order. During the interviews, participants are instructed to articulate their thoughts and interpretations of each item, with particular attention to any points of confusion, ambiguity, or difficulty in understanding. Researchers also actively prompt participants to explain their interpretations of specific words or phrases when clarification is needed. Feedback from the cognitive interviews is used to further refine item wording and improve clarity prior to large-scale validation.

11.3 Validation and Psychometric Evaluation of the Translated Version

Following pilot testing, the French version of the HXI will undergo a comprehensive psychometric evaluation. Unlike the development of the original HXI, exploratory factor analysis (EFA) will not be conducted for the French version. This decision is motivated by strong theoretical and empirical support for the five-factor structure established in the original validation of the HXI, as well as by practical considerations related to data collection. As such, confirmatory factor analysis (CFA) will be used as the primary method to evaluate the fit of the hypothesized five-factor measurement model.

Internal consistency reliability will be assessed for each subscale using Cronbach's alpha. Test-retest reliability may also be examined using data collected online. While exact replication of haptic experiences over time is challenging, this analysis will provide preliminary insight into the temporal stability of the instrument when applied to the same combination of applications and device.

Content validity is inherited from the original HXI through the translation and expert reconciliation process. Construct validity will be evaluated through CFA, while criterion-

related validity will be examined by assessing correlations between the French HXI and related measures, such as overall ratings of the haptic experience and established user experience questionnaires (e.g., UEQ, PXI), to test convergent and predictive validity.

Finally, measurement invariance analyses will be conducted to assess equivalence between the English and French versions of the HXI. Configural, metric, and scalar invariance will be tested to evaluate whether the instrument measures the same constructs in the same way across languages, supporting cross-linguistic comparability of HXI scores.

11.4 Conclusion

Haptic technology continues to grow in popularity and sophistication. Yet until now, haptic researchers and practitioners lacked a standardized, validated tool to measure the haptic experiences they create. This thesis addresses this gap. The Haptic Experience Inventory (HXI) is the first fully validated questionnaire with the purpose of evaluating haptic experience, and this thesis documents a rigorous end-to-end questionnaire development process spanning theoretical construction, expert review, cognitive interviews, exploratory and confirmatory factor analysis, and comprehensive psychometric evaluation, involving 610 participants and multiple haptic modalities.

The work in this thesis demonstrates that the questionnaire development pipeline — spanning theory construction, expert review, cognitive interviews, EFA, CFA, and psychometric evaluation — is transferable to other experiential domains in HCI that lack standardized measurement tools. The systematic approach taken here, including the iterative revision of theoretical constructs, multi-stage item refinement, and validation across diverse study contexts, provides a replicable template for developing domain-specific questionnaires in other areas. Moreover, the development process itself yielded insights beyond the instrument: the emergence of Discord as a psychologically distinct factor from Harmony, and the reclassification of Expressivity as a system characteristic rather than a user experience dimension, illustrate how rigorous scale development can advance theoretical understanding of the construct under study. These findings suggest that applying similar methods to other underexplored domains could not only produce reliable measurement tools, but also surface latent structural properties of experience that remain invisible to qualitative or ad hoc approaches.

The result of the work is a 20-item, five-factor instrument that is reliable, valid, and measurement-invariant across contexts, consisting of five factors: Autotelics, Realism, Involvement, Harmony, and Discord. The validation process also yielded empirical insights

into the structure of HX: that Expressivity is better positioned as a system characteristic than a user experience dimension, and that Harmony and Discord are psychologically distinct in ways that matter for experience and design. These findings advance the theoretical understanding of HX.

For researchers, the HXI enables systematic, large-scale evaluation and meaningful comparison across studies, devices, and modalities. For practitioners, it offers a practical diagnostic tool to identify which dimensions of a haptic design are working and which are not. The ongoing development of a French version, and potential future adaptations into more languages, will hopefully further extend its reach.

The HXI is not the final word on measuring HX, but it is a major step and a solid foundation, on which the community can further build, validate, and refine as haptic technology and the instrument see broader adoption.

References

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies*, 135:102370, March 2020.

[2] Sarvesh Agreval, Adèle Simon, Søren Bech, Klaus B Bærentsen, and Søren Forchammer. Defining immersion: Literature review and implications for research on immersive audiovisual experiences. In *AES 147th Convention*, page 10275. Audio Engineering Society, 2019.

[3] Ahmed Anwar, Tianzheng Shi, and Oliver Schneider. Factors of Haptic Experience across Multiple Haptic Modalities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pages 1–12, New York, NY, USA, April 2023. Association for Computing Machinery.

[4] Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 1968–1979, New York, NY, USA, 2016. Association for Computing Machinery.

[5] M. S. Bartlett. Tests of significance in factor analysis. *British Journal of Psychology*, 3:77–85, 1950. Place: United Kingdom Publisher: British Psychological Society.

[6] Olivier Bau, Ivan Poupyrev, Ali Israr, and Chris Harrison. Teslatouch: electrovibration for touch surfaces. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, pages 283–292, 2010.

[7] Christopher C Berger, Mar Gonzalez-Franco, Eyal Ofek, and Ken Hinckley. The uncanny valley of haptics. *Science Robotics*, 3(17):eaar7010, 2018.

[8] Godfred O Boateng, Torsten B Neilands, Edward A Frongillo, Hugo R Melgar-Quiñonez, and Sera L Young. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health*, 6:149, 2018.

[9] Stefan Josef Breitschaft, Alexander Pastukhov, and Claus-Christian Carbon. Where's my button? evaluating the user experience of surface haptics in featureless automotive user interfaces. *IEEE Transactions on Haptics*, 15(2):292–303, 2022.

[10] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 505–514, 2013.

[11] Raymond B. Cattell. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2):245–276, April 1966. Publisher: Routledge.

[12] Fang Fang Chen. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3):464–504, 2007.

[13] Yang Chen, Hamed Alimohammadzadeh, Shahram Ghandeharizadeh, and Heather Culbertson. Force-feedback through touch-based interactions with a nanocopter. In *2024 IEEE Haptics Symposium (HAPTICS)*, pages 271–277, 2024.

[14] Lung-Pan Cheng, Thijs Roumen, Hannes Rantzsch, Sven Köhler, Patrick Schmidt, Robert Kovacs, Johannes Jasper, Jonas Kemper, and Patrick Baudisch. Turkdeck: Physical virtual reality based on people. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pages 417–426, 2015.

[15] Chantal Chevroulet, Helen M. Paterson, Angus Yu, Emily Chew, and Richard I. Kemp. The impact of recall timing on the preservation of eyewitness memory. *Psychiatry, Psychology, and Law*, 29(3):471–486, 2021.

[16] Seungmoon Choi and Katherine J Kuchenbecker. Vibrotactile display: Perception, technology, and applications. *Proceedings of the IEEE*, 101(9):2093–2104, 2012.

[17] Lee J Cronbach. Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3):297–334, 1951.

[18] Lee J Cronbach and Paul E Meehl. Construct validity in psychological tests. *Psychological bulletin*, 52(4):281, 1955.

[19] Tor-Salve Dalsgaard, Joanna Bergström, Marianna Obrist, and Kasper Hornbæk. A user-derived mapping for mid-air haptic experiences. *International Journal of Human-Computer Studies*, 168:102920, 2022.

[20] Tor-Salve Dalsgaard and Oliver Schneider. A unified model for haptic experience. *ACM Transactions on Computer-Human Interaction*, 2025.

[21] Robert F DeVellis and Carolyn T Thorpe. *Scale development: Theory and applications*. Sage publications, 2021.

[22] Robert F. DeVellis and Carolyn T. Thorpe. *Scale Development: Theory and Applications*. SAGE Publications, August 2021. Google-Books-ID: FX0_EAAAQBAJ.

[23] Andrea Egger-Rainer. Enhancing validity through cognitive interviewing. A methodological example using the Epilepsy Monitoring Unit Comfort Questionnaire. *Journal of Advanced Nursing*, 75(1):224–233, January 2019.

[24] Mohamad Eid, Mauricio Orozco, and Abdulmotaleb El Saddik. A guided tour in haptic audio visual environments and applications. *International Journal of Advanced Media and Communication*, 1(3):265–297, 2007.

[25] Leandre Fabrigar, Duane Wegener, Robert MacCallum, and Erin Strahan. Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods*, 4:272, September 1999.

[26] Farimah Fazlollahi and Katherine J. Kuchenbecker. Haptify: A measurement-based benchmarking system for grounded force-feedback devices. *IEEE Transactions on Robotics*, 39(2):1622–1636, 2023.

[27] Gabriele Helga Franke. "The Whole is More than the Sum of its Parts": The Effects of Grouping and Randomizing Items on the Reliability and Validity of Questionnaires. *European Journal of Psychological Assessment*, 13(2):67–74, May 1997. Publisher: Hogrefe Publishing.

[28] Ujwal Gadiraju, Gianluca Demartini, Ricardo Kawase, and Stefan Dietze. Human Beyond the Machine: Challenges and Opportunities of Microtask Crowdsourcing. *IEEE Intelligent Systems*, 30(4):81–85, July 2015. Conference Name: IEEE Intelligent Systems.

[29] Orestis Georgiou, William Frier, and Oliver Schneider. User experience and mid-air haptics: Applications, methods, and challenges. *Ultrasound Mid-Air Haptics for Touchless Interfaces*, pages 21–69, 2022.

[30] Janet K Gibbs, Marco Gillies, and Xueni Pan. A comparison of the effects of haptic and visual feedback on presence in virtual reality. *International Journal of Human-Computer Studies*, 157:102717, 2022.

[31] Aqeel Haider, Gunter Wallner, Kathrin Gerling, and Vero Vanden Abeele. An investigation of the test-retest reliability of the minipxi. *Proceedings of the ACM on Human-Computer Interaction*, 8(CHI PLAY):1–24, 2024.

[32] Kyunghee Han, Stephen M. Colarelli, and Nathan C. Weed. Methodological and statistical advances in the consideration of cultural diversity in assessment: A critical review of group classification and measurement invariance testing. *Psychological Assessment*, 31(12):1481–1496, December 2019.

[33] Marc Hassenzahl. The thing and i: understanding the relationship between user and product. *Funology 2: from usability to enjoyment*, pages 301–313, 2018.

[34] Marc Hassenzahl, Michael Burmester, and Franz Koller. Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. *Mensch & Computer 2003: Interaktion in Bewegung*, pages 187–196, 2003.

[35] Stephen N Haynes, David Richard, and Edward S Kubany. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment*, 7(3):238, 1995.

[36] Stephen N. Haynes, David C. S. Richard, and Edward S. Kubany. Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3):238–247, 1995. Place: US Publisher: American Psychological Association.

[37] Alan E. Hendrickson and Paul Owen White. Promax: A Quick Method for Rotation to Oblique Simple Structure. *British Journal of Statistical Psychology*, 17(1):65–70, 1964. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2044-8317.1964.tb00244.x.

[38] Eve Hoggan, Topi Kaaresoja, Pauli Laitinen, and Stephen Brewster. Crossmodal congruence: the look, feel and sound of touchscreen widgets. In *Proceedings of the 10th international conference on Multimodal interfaces*, ICMI '08, pages 157–164, New York, NY, USA, October 2008. Association for Computing Machinery.

[39] John L. Horn. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185, 1965. Place: Germany Publisher: Springer.

[40] Perttu Hämäläinen, Joe Marshall, Raine Kajastila, Richard Byrne, and Florian "Floyd" Mueller. Utilizing Gravity in Movement-Based Games and Play. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '15, pages 67–77, New York, NY, USA, October 2015. Association for Computing Machinery.

[41] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. The game experience questionnaire. 2013.

[42] Seungwoo Je, Myung Jin Kim, Woojin Lee, Byungjoo Lee, Xing-Dong Yang, Pedro Lopes, and Andrea Bianchi. Aero-plane: A handheld force-feedback device that renders weight motion illusion on a virtual 2d plane. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pages 763–775, 2019.

[43] Kevin John, Yinan Li, and Hasti Seifi. Adaptics: A toolkit for creative design and integration of real-time adaptive mid-air ultrasound tactons. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery.

[44] Henry F. Kaiser. An index of factorial simplicity. *Psychometrika*, 39(1):31–36, March 1974.

[45] Erin Kim and Oliver Schneider. Defining Haptic Experience: Foundations for Understanding, Communicating, and Evaluating HX. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–13, New York, NY, USA, April 2020. Association for Computing Machinery.

[46] Aniket Kittur, Ed H. Chi, and Bongwon Suh. Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 453–456, Florence Italy, April 2008. ACM.

[47] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1301–1318, San Antonio Texas USA, February 2013. ACM.

[48] Bettina Laugwitz, Theo Held, and Martin Schrepp. Construction and evaluation of a user experience questionnaire. In Andreas Holzinger, editor, *HCI and Usability for*

Education and Work, pages 63–76, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[49] Shi-Hong Liu, Pai-Chien Yen, Yi-Hsuan Mao, Yu-Hsin Lin, Erick Chandra, and Mike Y Chen. Headblaster: a wearable approach to simulating motion perception using head-mounted air propulsion jets. *ACM Transactions On Graphics (TOG)*, 39(4):84–1, 2020.

[50] Tuva Fjærtoft Lønne, Håvard R Karlsen, Eva Langvik, and Ingvild Saksvik-Lehouillier. The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: A randomized controlled study. *Heliyon*, 9(6), 2023.

[51] Karon E MacLean, Oliver S Schneider, and Hasti Seifi. Multisensory haptic interactions: understanding the sense and designing for it. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations-Volume 1*, pages 97–142. 2017.

[52] Emanuela Maggioni, Erika Agostinelli, and Marianna Obrist. Measuring the added value of haptic feedback. In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6, 2017.

[53] Camila Loiola Brito Maia and Elizabeth Sucupira Furtado. A systematic review about user experience evaluation. In Aaron Marcus, editor, *Design, User Experience, and Usability: Design Thinking and Methods*, pages 445–455, Cham, 2016. Springer International Publishing.

[54] Keith Meadows. Cognitive Interviewing Methodologies. *Clinical Nursing Research*, 30(4):375–379, May 2021. Publisher: SAGE Publications Inc.

[55] Elisa D Mekler and Kasper Hornbæk. Momentary pleasure or lasting meaning? distinguishing eudaimonic and hedonic user experiences. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 4509–4520, 2016.

[56] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[57] Michael Minge, Manfred Thüring, Ingmar Wagner, and Carina V Kuhr. The mecue questionnaire: a modular tool for measuring user experience. In *Advances in Ergonomics Modeling, Usability & Special Populations: Proceedings of the AHFE 2016*

International Conference on Ergonomics Modeling, Usability & Special Populations, July 27-31, 2016, Walt Disney World®, Florida, USA, pages 115–128. Springer, 2017.

[58] Jum C. Nunnally. *Psychometric Theory*. McGraw-Hill, 1978. Google-Books-ID: WE59AAAAMAAJ.

[59] Marianna Obrist, Sue Ann Seah, and Sriram Subramanian. Talking about tactile experiences. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1659–1668, 2013.

[60] Chaeyong Park, Jeongwoo Kim, Dong-Geun Kim, Seungjae Oh, and Seungmoon Choi. Vibration-augmented buttons: Information transmission capacity and application to interaction design. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2022.

[61] Horea Pauna, Pierre-majorique Léger, Sylvain Sénécal, Marc Fredette, François Courtemanche, Shang-Lin Chen, Élise Labonté-Lemoyne, and Jean-François Ménard. The psychophysiological effect of a vibro-kinetic movie experience: The case of the d-box movie seat. In *Information Systems and Neuroscience: Gmunden Retreat on NeuroIS 2017*, pages 1–7. Springer, 2018.

[62] Joann Peck and Terry L. Childers. Individual Differences in Haptic Information Processing: The "Need for Touch" Scale. *Journal of Consumer Research*, 30(3):430–442, December 2003.

[63] Diane L. Putnick and Marc H. Bornstein. Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental review : DR*, 41:71–90, September 2016.

[64] Tom Roy, Yann Glémarec, Gurvan Lécuyer, Quentin Galvane, Philippe Guillotel, and Ferran Argelaguet. Towards end-user customization of haptic experiences. In *EuroHaptics*, pages 1–14, 2024.

[65] Suji Sathiyamurthy, Melody Lui, Erin Kim, and Oliver Schneider. Measuring Haptic Experience: Elaborating the HX model with scale development. In *2021 IEEE World Haptics Conference (WHC)*, pages 979–984. Institute of Electrical and Electronics Engineers, July 2021.

[66] Owen Schaffer. Development and Preliminary Validation of the Enjoyment Questionnaire and the Sources of Enjoyment Questionnaire. In *Extended Abstracts of the 2022*

CHI Conference on Human Factors in Computing Systems, CHI EA '22, pages 1–7, New York, NY, USA, April 2022. Association for Computing Machinery.

[67] Oliver Schneider, Karon MacLean, Colin Swindells, and Kellogg Booth. Haptic experience design: What hapticians do and where they need help. *International Journal of Human-Computer Studies*, 107:5–21, 2017. Multisensory Human-Computer Interaction.

[68] Oliver S. Schneider and Karon E. MacLean. Studying design process and example use with macaron, a web-based vibrotactile effect editor. In *2016 IEEE Haptics Symposium (HAPTICS)*, pages 52–58, 2016.

[69] Oliver S. Schneider, Hasti Seifi, Salma Kashani, Matthew Chun, and Karon E. MacLean. HapTurk: Crowdsourcing Affective Ratings of Vibrotactile Icons. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3248–3260, New York, NY, USA, May 2016. Association for Computing Machinery.

[70] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4:103, January 2017.

[71] Hasti Seifi, Farimah Fazlollahi, Michael Oppermann, John Andrew Sastrillo, Jessica Ip, Ashutosh Agrawal, Gunhyuk Park, Katherine J Kuchenbecker, and Karon E MacLean. Haptipedia: Accelerating haptic device discovery to support interaction & engineering design. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pages 1–12, 2019.

[72] Hasti Seifi, Kailun Zhang, and Karon E. MacLean. Vibviz: Organizing, visualizing and navigating vibration libraries. In *2015 IEEE World Haptics Conference (WHC)*, pages 254–259, 2015.

[73] Tianzheng Shi and Oliver Schneider. Development and initial validation of the haptic experience inventory (hxi). In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, 2025. Association for Computing Machinery.

[74] Tanay Singhal and Oliver Schneider. Juicy haptic design: Vibrotactile embellishments can improve player experience in games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.

[75] Mel Slater. Immersion and the illusion of presence in virtual reality. *British journal of psychology*, 109(3):431–433, 2018.

[76] Karthikan Theivendran, Andy Wu, William Frier, and Oliver Schneider. Rechap: An interactive recommender system for navigating a large number of mid-air haptic designs. *IEEE Transactions on Haptics*, 17(2):165–176, 2024.

[77] Howard EA Tinsley and Steven D Brown. *Handbook of applied multivariate statistics and mathematical modeling*. Academic press, 2000.

[78] Alexandre N. Tuch and Kasper Hornbæk. Does herzberg's notion of hygienes and motivators apply to user experience? *ACM Trans. Comput.-Hum. Interact.*, 22(4), jun 2015.

[79] Luca Turchet, Travis West, and Marcelo M Wanderley. Touching the audience: musical haptic wearables for augmented and participatory live music performances. *Personal and Ubiquitous Computing*, 25:749–769, 2021.

[80] Maurice HPH van Beurden, Wijnand A Ijsselsteijn, and Yvonne AW de Kort. User experience of gesture based interfaces: a comparison with traditional interaction methods on pragmatic and hedonic qualities. In *International gesture workshop*, pages 36–47. Springer, 2011.

[81] Robert J. Vandenberg and Charles E. Lance. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1):4–70, 2000.

[82] Marley W. Watkins. Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3):219–246, April 2018. Publisher: SAGE Publications Inc.

[83] Keith F Widaman and Steven P Reise. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. 1997.

[84] Bob G. Witmer, Christian J. Jerome, and Michael J. Singer. The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3):298–312, June 2005.

[85] Bob G. Witmer and Michael J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence*, 7(3):225–240, June 1998. Conference Name: Presence.

[86] Gyeore Yun, Hyoseung Lee, Sangyoon Han, and Seungmoon Choi. Improving viewing experiences of first-person shooter gameplays with automatically-generated motion effects. In *Proceedings of the 2021 CHI conference on human factors in computing systems*, pages 1–14, 2021.

APPENDICES

Appendix A

Items drafted based on theoretical construction and revisions made during expert review

This appendix lists all 84 items arising from theory constructing, and revisions made during expert review.

A.1 Items for *Autotelics*

Retained without revision

A1 I enjoyed the haptic sensations themselves

A2 The haptic sensations themselves were pleasant.

A9 The haptic sensations themselves were delightful to experience.

A12 I like how the haptic sensation itself feels, regardless of its role in the system.

A13 Regardless of function, I found the haptic sensations pleasant.

A14 The haptic sensations provided a pleasurable sensory experience regardless of its utility.

Retained with revision

A4 The feeling of the haptic feedback alone was satisfying to me.

Revised to "The feeling of the haptic feedback itself was satisfying to me."

A5 I disliked the haptic sensations themselves.

Revised to "I disliked how haptic sensations felt."

A10 I was not satisfied with the haptic sensations themselves.

Revised to "I was not satisfied with the haptic sensations."

A11 The haptic sensations were enjoyable on their own, regardless of its function.

Revised to "The haptic sensations were enjoyable on their own, regardless of their function."

A17 Simply experiencing the haptic feedback was enjoyable to me.

Revised to "Experiencing the haptic feedback was enjoyable to me."

Dropped

A3 I appreciated the haptic sensations for the sheer pleasure they offered.

A6 I can experience a sense of joy purely from the tactile sensations themselves.

A7 The haptic sensations were satisfying in and of themselves.

A8 The haptic sensations themselves were rewarding.

A15 The haptic sensations provided a pleasurable experience independent of their purpose.

A16 I found the haptic sensations delightful independently of their purpose.

A18 I appreciated the haptic sensations solely for the enjoyment they provided.

A.2 Items for *Realism*

Retained without revision

R8 The haptic sensations closely mimicked the experiences I would expect in reality.

R9 The haptic sensations made me feel like I was experiencing something real.

R13 The haptic sensations felt realistic.

R14 I experienced a disconnect between the haptic sensations and what I expected.

Retained with revision

R1 The haptic sensations closely resembled the ones I feel in real life.

Revised to "The haptic sensations resembled the ones I feel in real life."

R3 The haptic sensations provided a lifelike representation of real-world sensations.

Revised to "The haptic sensations provided a true-to-life representation of real-world sensations."

R5 The haptic sensations elicited a strong sense of realism and believability.

Split into "The haptic sensations were believable." and "The haptic sensations elicited a strong sense of realism."

R6 The haptic sensations feel consistent with real-life touch experiences.

Revised to "The haptic sensations felt consistent with real-life touch experiences."

R10 The haptic sensations triggered a sense of familiarity, resembling real life touch.

Revised to "The haptic sensations felt familiar to real life touch."

Dropped

R2 The haptic sensations were true-to-life

R4 I was impressed by how accurately the haptic sensations represented real-world interactions.

R7 The haptic sensations matched my expectations of real-world touch.

R11 The haptic sensations feel natural in the context of the virtual environment or system.

R12 The haptic sensations felt unnatural.

R15 The haptic feedback felt strange.

R16 The haptic sensations seem out of the ordinary.

A.3 Items for *Harmony*

Retained without revision

H4 The haptic sensations blended well with the overall experience.

H7 I feel the haptic sensations are well coordinated with other senses.

H8 The haptic sensations felt disconnected from the rest of the experience.

H9 The haptic sensations clashed with the overall sensory experience.

Retained with revision

H1 The haptic sensations integrated seamlessly with other sensory cues.

Revised to "The haptic sensations integrated seamlessly with other senses."

H2 The haptic sensations complemented other modalities effectively.

Revised to "The haptic sensations complemented other senses well."

H3 I felt a sense of unity and cohesion between the haptic sensations and other senses.

Revised to "I felt a sense of cohesion between the haptic sensations and other senses."

H10 The haptic sensations seemed to lack coordination with what I was seeing or hearing.

Revised to "The haptic sensations seemed to lack coordination with other senses."

H11 I experienced a sense of incongruity between the haptic sensations and other senses.

Revised to "I experienced a sense of mismatch between the haptic sensations and other senses."

H14 The sensory elements didn't work together, creating a sense of discord..

Revised to "The sensory elements did not work together, creating a sense of discord."

H15 The haptic sensations felt out of sync with the other senses.

Revised to "The haptic sensations felt out of sync with the other senses."

Dropped

H5 The haptic sensations felt like they belonged in the overall experience.

H6 The haptic sensations felt like an integral part of the experience.

H12 The haptic sensations didn't fit in the context.

H13 The haptic sensations felt awkward and disruptive together with the other senses.

H16 The haptic sensations seemed out of place in the system or virtual environment.

A.4 Items for *Expressivity*

Retained without revision

E2 The haptic sensations felt the same all the time.

E3 The haptic sensations felt indistinguishable from each other.

E4 The haptic sensations change depending on how things change in the system.

E7 I can perceive a range of haptic variations that distinctly correspond to different inputs or events.

E9 I could discern specific events or actions through haptic sensations.

Retained with revision

E1 The haptic sensations were rich and dynamic.

Revised to "The haptic sensations were rich."

E10 The diversity in haptic sensations allows me to interpret changes in the system.

Revised to "The diversity in haptic sensations allowed me to interpret changes in the system."

E13 Different haptic sensations help me distinguish between different elements or actions within the virtual environment.

Revised to "Different haptic sensations helped me distinguish between different elements or actions within the system."

E14 The haptic sensations dynamically change according to my actions or changes in the system.

Revised to "The haptic sensations change dynamically according to my actions or changes in the system."

E17 The haptic feedback changes appropriately based on my actions and interactions.

Revised to "The haptic sensations changes appropriately based on my actions and interactions."

Dropped

E5 The haptic sensations reflect varying inputs, actions, or events.

E6 The haptic sensations effectively communicate changes in the virtual environment.

E8 The haptic stimuli effectively conveyed different types of information.

E11 The haptic sensation conveys specific messages related to the system.

E12 Different haptic sensations make it clear when there is a change in the system.

E15 The haptic sensations effectively convey information about different interactions with the system.

E16 I notice a range of haptic sensations that correspond to different inputs and events.

A.5 Items for *Involvement*

Retained without revision

I17 I felt engaged with the system due to the haptic sensations.

Retained with revision

I1 The haptic sensations increased my involvement in the task or system.

Revised to "The haptic sensations contributed to my involvement in the task."

I3 I felt more deeply involved in the task because of the haptic sensations.

Revised to "I felt deeply involved in the task because of the haptic sensations."

I4 I felt more absorbed in the task due to the haptic sensations.

Revised to "I felt absorbed in the task due to the haptic sensations"

I7 I found the haptic elements fostered my engagement with the system or task.

Revised to "I found the haptic sensations fostered my engagement with the system."

I9 The haptic interactions made me more attentive and focused.

Revised to "The haptic interactions made me more focused."

I13 I felt a heightened sense of connection and engagement due to the haptic sensations.

Revised to "I felt a strong sense of engagement due to the haptic sensations."

I14 The haptic sensations heightened my overall engagement in the task or system.

Revised to "The haptic sensations supported my overall engagement with the system."

I15 I felt more deeply connected to the task or system because of the haptic sensations.

Revised to "I felt deeply connected to the system because of the haptic sensations."

I16 I found myself deeply engaged and connected with the task or system through the haptic interactions.

Revised to "I found myself deeply connected to the system through the haptic interactions."

Dropped

I2 The haptic sensations contributed significantly to my overall involvement in the system.

I5 I lost track of time and place due to the immersive haptic sensations.

I6 I experienced a profound level of engagement and focus through the haptic sensations.

I8 The haptic sensations captured my attention and increased my focus on the system or task.

I10 I devoted more attention and energy to the task or system because of the haptic feedback.

I11 The haptic interactions captivated my attention, fostering a stronger connection with the task or system.

I12 The haptic sensations made the task or system more captivating and engaging.I'm motivated to explore more by the haptic sensations.

Appendix B

Items entering factor analysis, after Study 1 and Study 2

This appendix lists all 50 revised items that were used for factor analysis, along with the two attention checks implemented.

B.1 Items for *Autotelics*

A1 The haptic sensations were pleasant on their own.

A2 I enjoyed the haptic sensations themselves.

A3 The feeling of the haptic sensations itself was satisfying to me.

A4 I like how the haptic sensation itself feels, regardless of its role in the system.

A5 The haptic sensations themselves were delightful to experience.

A6 Regardless of function, I found the haptic sensations pleasant.

A7 The haptic sensations were enjoyable on their own, regardless of their function.

A8 I disliked how the haptic sensations felt.

A9 I was not satisfied with the haptic sensations.

A10 Experiencing the haptic sensations was enjoyable to me.

A11 The haptic sensations provided a pleasurable sensory experience regardless of its utility.

B.2 Items for *Expressivity*

E1 The haptic sensations felt the same all the time.

E2 The diversity in haptic sensations allowed me to interpret changes in the system.

E3 Different haptic sensations helped me distinguish between different elements or actions within the system.

E4 The haptic sensations changes appropriately based on my actions and interactions.

E5 I could distinguish specific system events or actions through haptic sensations.

E6 The haptic sensations change dynamically according to my actions or changes in the system.

E7 The haptic sensations felt indistinguishable from one another.

E8 The haptic sensations change depending on changes in the system.

E9 The haptic sensations have a rich range of variations.

B.3 Items for *Harmony*

H1 The haptic sensations clashed with the overall sensory experience.

H2 The haptic sensations felt out of sync with the other senses.

H3 The haptic sensations seemed to lack coordination with other senses.

H4 The haptic sensations blended well with the overall experience.

H5 I felt a sense of harmony between the haptic sensations and other senses.

H6 The haptic sensations felt disconnected from the rest of the experience.

H7 The haptic sensations integrated seamlessly with other senses.

H8 I experienced a sense of mismatch between the haptic sensations and other senses.

H9 I feel the haptic sensations are well coordinated with the other senses.

H10 The haptic sensations complemented other senses well.

B.4 Items for *Involvement*

I1 I felt deeply connected to the system because of the haptic sensations.

I2 I felt deeply involved in the task because of the haptic sensations.

I3 The haptic sensations contributed to my involvement in the task.

I4 I felt engaged with the system due to the haptic sensations.

I5 I found myself deeply connected to the system through the haptic interactions.

I6 I found the haptic sensations strengthened my engagement with the system.

I7 The haptic sensations supported my overall engagement with the system.

I8 I felt absorbed in the task due to the haptic sensations.

I9 The haptic interactions made me more focused.

I10 I felt a sense of engagement due to the haptic sensations.

B.5 Items for *Realism*

R1 The haptic sensations felt realistic.

R2 The haptic sensations resembled the ones I feel in real life.

R3 The haptic sensations felt consistent with real-life touch experiences.

R4 The haptic sensations provided a true-to-life representation of real-world sensations.

R5 I experienced a disconnect between the haptic sensations and what I expected.

R6 The haptic sensations made me feel like I was experiencing something real.

R7 The haptic sensations closely mimicked the experiences I would expect in reality.

R8 The haptic sensations were believable.

R9 The haptic sensations evoked a strong sense of realism.

R10 The haptic sensations felt familiar to real life touch.

B.6 Attention Check Items

D1 Please select somewhat disagree for this statement.

D2 I am a computer program, not a real human.

Appendix C

Descriptive for data collected in Study 3

	ONLINE				INPERSON			
Item	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**
A1	1	7	5.46	1.190	2	7	5.25	1.040
A2	1	7	5.64	1.161	2	7	5.28	1.106
A3	1	7	5.62	1.210	2	7	5.26	1.210
A4	2	7	5.26	1.386	1	7	5.17	1.171
A5	1	7	5.37	1.297	2	7	5.22	1.092
A6	1	7	5.42	1.263	2	7	5.19	1.162
A7	1	7	5.37	1.234	2	7	5.24	1.206
A8	2	7	5.88	1.284	2	7	5.36	1.255
A9	1	7	5.49	1.568	1	7	5.26	1.314
A10	1	7	5.66	1.159	2	7	5.38	1.095
A11	1	7	5.37	1.229	2	7	5.23	1.129
R1	1	7	4.71	1.570	1	7	4.88	1.417
R2	1	7	4.18	1.571	1	7	4.62	1.501
R3	1	7	4.47	1.525	1	7	4.71	1.451
R4	1	7	4.34	1.575	1	7	4.80	1.429
R5	1	7	5.33	1.371	1	7	4.78	1.414
R6	1	7	4.49	1.588	2	7	5.04	1.243
R7	1	7	4.17	1.553	1	7	4.70	1.453
R8	1	7	4.80	1.406	2	7	5.09	1.181
R9	1	7	4.54	1.584	1	7	4.80	1.393
R10	1	7	4.10	1.602	1	7	4.73	1.492
H1	1	7	4.72	1.778	1	7	5.09	1.377
H2	1	7	5.18	1.483	1	7	4.96	1.326
H3	1	7	5.21	1.394	1	7	4.77	1.369
H4	2	7	5.78	1.016	2	7	5.55	1.121
H5	1	7	5.06	1.290	2	7	5.12	1.155
H6	1	7	5.46	1.385	1	7	5.15	1.261
H7	1	7	5.07	1.165	2	7	5.00	1.250
H8	1	7	5.15	1.521	1	7	4.91	1.334
H9	1	7	5.20	1.194	2	7	5.21	1.180
H10	2	7	5.29	1.139	2	7	5.24	1.057

Table C.1: Statistical Summary of Online and In-Person Scores in Study 3

	ONLINE				INPERSON			
Item	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**
E1	1	7	4.45	1.625	1	7	5.12	1.582
E2	2	7	5.30	1.239	2	7	5.38	1.049
E3	2	7	5.57	1.114	3	7	5.37	1.011
E4	2	7	5.68	1.038	2	7	5.42	1.031
E5	1	7	5.34	1.322	2	7	5.43	0.801
E6	2	7	5.33	1.229	3	7	5.60	0.932
E7	1	7	4.47	1.630	1	7	5.05	1.320
E8	2	7	5.21	1.207	2	7	5.46	0.990
E9	1	7	4.97	1.354	2	7	5.14	1.181
I1	1	7	4.95	1.428	2	7	5.35	1.001
I2	1	7	5.24	1.419	2	7	5.62	1.028
I3	1	7	5.48	1.265	2	7	5.74	0.907
I4	1	7	5.45	1.210	3	7	5.64	0.896
I5	1	7	5.02	1.363	2	7	5.33	1.063
I6	1	7	5.29	1.351	2	7	5.72	0.861
I7	2	7	5.45	1.087	2	7	5.72	0.781
I8	1	7	5.18	1.358	2	7	5.62	0.873
I9	1	7	5.10	1.501	2	7	5.68	0.989
I10	1	7	5.18	1.502	2	7	5.63	0.907

Table C.1: Statistical Summary of Online and In-Person Scores in Study 3 (continued)