EVALUATING EMBODIMENT IN MUSICAL INSTRUMENT MODIFICATION AND AUGMENTATION

ANDREA GUIDI

Supervisors: Professor Andrew McPherson & Dr Tony Stockman

Examiners: Dr Tom Mitchell & Professor Joshua Reiss

Submitted in partial fulfilment of the
University of London Degree of Doctor of Philosophy
School of Electronic Engineering and Computer Science
Media and Arts Technology
Queen Mary University of London

Abstract

This PhD seeks to identify key aspects that optimise the learning process of new musical interfaces by professional musicians. Modifying or extending an existing musical instrument can impact players' skills. Fluency of execution or pitch accuracy can deteriorate due to demands on the performer's attention from the unfamiliarity of the instrument. As a result, players may require additional training on a modified instrument before they regain their fluency. The problem is that performers, especially professional players who have already invested many years in the unmodified musical instrument, might prefer to start from a high level. Thus, designing a new instrument that builds upon existing skills can be appealing.

However, which design strategies might support such a goal? Which aspects of the original design should be preserved? How can we assess whether the resulting modified instrument allows the performer to retain their skills? This research presents four studies that tackle these questions. Results from the first two studies suggest that the design strategy should focus on participants' sensorimotor imagery rather than the instrument's auditory feedback. During these studies, participants were still able to retain their fluency and pitch accuracy even in the presence of disrupting or irrelevant auditory feedback. Two additional studies propose quantitative methods to evaluate skill retention in instrument modification. This research can advise designers on whether they are on the right track in crafting an interface that builds upon existing skills. This challenge may apply to augmented instruments, the modification of existing musical instruments, or new digital instruments.

Acknowledgements

I would like to thank my primary supervisor, Prof. Andrew McPherson, for the support, time and enthusiasm he dedicated to supervising this research. Additional people provided technical, intellectual, and psychological support throughout this journey, like the Bela Team (Giulio, Robbie, Adan, particularly), the AIL lab (Charlotte and Giacomo, especially), the MAT crew (Jack and Jorge, particularly), Jonathan, Nick and my family and friends including Agnese, Anna, Flavio, Lorena, Jelena and Orazio (alphabetic order). Throughout this research, I was financially supported by the Engineering and Physical Sciences Research Council, Arts and Humanities Research Council via the Media and Arts Technology Doctoral Training Centre grant, and by my friends and family. Without this support, I could not have undertaken and completed this research.

Declaration

I, Andrea Guidi, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original and does not, to the best of my knowledge, break any UK law, infringe any third party's copyright or other Intellectual Property Rights, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Milano, Italia, November 2022

0.1 Associated publications

Portions of the work detailed in this thesis have appeared previously in the following conference papers.

- Chapter 5: Andrea Guidi, Andrew McPherson: Quantitative Evaluation of Aspects of Embodiment in New Digital Musical Instruments. NIME 2022.
- Chapter 3: Andrea Guidi, Fabio Morreale, Andrew McPherson: design for auditory imagery: Altering Instruments to Explore Performer Fluency. NIME 2020.
- Chapter 4: Andrea Guidi, Fabio Morreale, Andrew McPherson: design for auditory imagery: Altering Instruments to Explore Performer Fluency. NIME 2020.
- Chapter 5: Fabio Morreale, Andrea Guidi, Andrew P. McPherson: Magpick: an Augmented Guitar Pick for Nuanced Control. NIME 2019.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This document presents an exploration and evaluation of aspects of instrument design that retain virtuosity[1] on a modified or extended musical instrument during music performances.

Players can achieve subtle, detailed, and quick adjustments on a musical instrument during a virtuosic performance. This ability relies on years of training. During the training process, players develop skills that allow for the real-time performance of complex music. The modification or extension of an existing musical instrument can lower those skills. Fluency of executions or pitch accuracy can deteriorate as performers' attention is drawn by learning the new version of their instrument. As a result, players may require additional training.

The problem is that performers, and perhaps especially professional players who have already invested many years in the unmodified musical instrument, might prefer to start from a high level. Which design considerations should be regarded to optimise the learning process of using an unfamiliar musical instrument modification? Or how to design the modification to afford the use of the existing performing skills related to the original interface? The goal of transferring expert players' abilities is to make learning the modification or extension of a musical instrument less time-consuming.

It may be reasonable to design instruments that use existing skills to achieve this goal, as we can see from the many examples in recent history, like the Moog synthetic keyboard or the electric guitar. Expert performances with new instruments are rare, especially outside of the design process [1, 2]. The expertise to perform with such interfaces is rarely achieved [3].

Hence we may need to do that more with digital instruments. During this

[1] In this document, virtuosity claims to simply name the presence of performing skills such as the ability to achieve subtle, detailed and quick adjustments on a musical instrument during a performance

study, skilled musicians encountered musical interfaces that could allow them to leverage their existing abilities. The main research goal is to understand "does the interface facilitate it?" And more importantly, "how do we tell?". Even before we get to the instrument design stage, it is not obvious how to assess whether people are using their skills or not. How do we even know how far people's skills can transfer? How do we know to what extent they are drawing on their existing musical skills if confronted with a modified instrument?

This thesis does not address situations where individuals are willing to invest significant time in learning a new instrument. Rather, the challenge arises when individuals, particularly professional musicians who have already invested substantial years in mastering an instrument, desire to begin at an advanced level. Rather than observing players developing new skills over extensive periods of time, I'm looking at whether players manage to use or adapt their existing skills. Consequently, the primary focus of this research is not on examining the process of extensive instrument learning. The research outlined in this thesis seeks to explore the experiences of professional performers when confronted with unfamiliar musical technology and to evaluate their early responses.

The number of participants for each study was decided according to the examined literature and available funding. I deliberately created artificial situations by giving people "new or unfamiliar instruments" to play traditional repertoire. Corrective effects and mistakes are considered as data to test a series of research hypotheses. Each study aims to create a laboratory to observe performers' behaviour using a modified or augmented musical instrument in the outlined musical context. The subjective feeling of participants was not measured; For example, I measured the quantity of movement in their picking gestures. In these studies, I tried to bring an external view to whether somebody can execute skilled actions on an unfamiliar interface. The goal is not to privilege an objective method against subjective methodologies but rather to complement existing methods with something that is outwardly observable and repeatable.

This research advises designers on whether they are on the right track in designing an interface that builds upon existing skills. In addition to augmented instruments, this challenge may also apply to modifications of existing musical instruments and new digital instruments.

1.1 The Field of Research

There are many musical contexts and different motivations for making and playing instruments.

This research primarily focuses on the experiences of musicians who are well-versed in traditional Western acoustic instruments and are transitioning to or exploring another related or modified instrument. Such individuals are already skilled in one instrument but are in the early stages of encountering another related or modified instrument. The study's scope is outlined by several boundaries:

1. **Score-Based Performance Context**: This PhD is focused on score-based music performances rather than composition and improvisation. Specifically, it centres on performances situated within the Western classical tradition. In this context, the instrument generally responds to the performer's gestures in real-time reliably and predictably [4].

2. **Musicians' Background**: Tailored to musicians familiar with traditional Western acoustic instruments, those outside of this domain may not find the findings wholly applicable.

3. **Skill Level**: Targets participants who have already achieved a high level of expertise on one instrument. The research might not resonate with absolute beginners or musicians without foundational proficiency.

4. **Instrument Modification**: Considers musicians' first encounter with a related or modified version of a familiar instrument such as the augmented plectrum described in chapter 5.

5. **Exclusions**: Findings are not intended to generalize to genres involving shared creative agency with the instrument or musicians from diverse skill levels and different cultural backgrounds.

In essence, while the insights derived from this research can be enlightening for many, they are most applicable within the defined constraints.

I am aware of different musical values often found in communities like NIME[2]. Such values might include uncertainty, exploration, and the instrument as a co-creative agent [5, 6, 2, 7]. While intriguing, it is not the focus of this research. In fact, this research does not address performance cases characterised by attributes like surprise and uncertainty. This PhD also does not question whether virtuosity should be a goal for all music practices and instrument learning.

The research described in this document encompasses many fields, leading to a series of focused experiments. Its intellectual foundation is rooted in theories of music psychology embodiment, addressing embodiment, transparency, auditory imagery, and post-phenomenology. Part of this research will build on recent findings of music psychology, which in turn are based on neuroscience.

[2]New Interfaces for Musical Expression conference

1.2 Research Questions

This thesis investigates the design of modified or augmented instruments in terms of their impact on performers' skills. It also provides evaluation methods to assess the design's ability to achieve the use performer's existing skills. Here I define the fundamental research questions.

- RQ1: How can designers leverage performers' existing skills to play something new or unfamiliar? How does the ability to imagine the target sound affect accuracy and response time compared to playing familiar sounds?
- RQ2: To what extent does unfamiliar auditory feedback affect performers' pitch accuracy and fluency?
- RQ3: Does performance improve when participants focus on their internal representation of sound instead of the instrument's auditory feedback?
- RQ4: How can we assess the transferability of existing motor skills to a modified or extended musical instrument? To what extent do players utilize their existing training?
- RQ5: Does transferring existing musical skills to a new musical interface require conscious attention, and does it impact other performance aspects?
- RQ6: How easily can a gesture producing an unfamiliar sound in real-time integrate into a music performance compared to a gesture modifying the sound in the near future?
- RQ7: How can we evaluate the integration of instrument modifications into players' execution?

Performers often describe the experience of needing to be able to hear a sound in their head before they can accurately play it on their instrument. Green, in his book The Inner Game of Music, reports, "When you can hold the sound and pitch of the music clearly in your head... performing it accurately becomes easier. Your body has a sense of its goal (...). Effectively, you are playing a duet between the music in your head and the music you are performing [8]." This phenomenon is known in the literature as auditory imagery (see section 2.2). In new musical instruments design and performance, there is often a tension between familiarity and creative novelty. Doing something completely novel might entail abandoning any pre-acquired performance expertise, presenting a challenge for players. **RQ1** investigates **how, and how far, can designers pull performing skills that somebody already has to play something new or unfamiliar? Moreover, how accurately can professional players**

reproduce sounds for which they do not have developed auditory imagery? If the target sound cannot be imagined, how does that differ in terms of accuracy and response time to being able to play familiar sounds that can be imagined? The first study in this thesis addresses these questions and discusses the role of auditory imagery in musical instrument design for skills transfer.

Auditory feedback may be defined as the sound produced in response to musicians' physical actions on a musical instrument. Such feedback allows players to tune their performance and achieve subtle adjustments concerning intonation and articulation [9]. Performers imagine a sound, perform a gesture to play it on their instrument, listen to the resulting auditory feedback and compare it against their initial expectations to achieve an evaluation of the result [9]. Modifying an instrument can imply that familiar gestures produce unfamiliar sounds. **RQ2** asks **to what extent performers' pitch accuracy and fluency deteriorate in the presence of unfamiliar auditory feedback and thus mismatched auditory imagery. RQ3** asks **does players' performance improve if the task allows participants to ignore the sound coming from the instrument and to focus on their internal representation of sound (i.e. their expectations based on the sound the instrument used to reproduce)?** The second study in this research addresses these questions and queries the role of auditory feedback in supporting the production of novel or unfamiliar sounds on a modified instrument.

While learning a musical instrument, performers develop abstract representations of the movements they need to play it [10]. These representations are known as motor programs [11] and are essential to regulate players' gestures on the instrument. A modified or extended musical instrument may require a performer to adapt the gestures needed to perform sounds. **RQ4** asks **How can we assess the extent of transferability of existing motor skills to a modified or extended musical instrument, and how can we determine the level to which players utilize their existing training? RQ5** asks **Does the transfer of existing musical skills to a new musical interface require conscious attention, potentially leading to neglect of other performance aspects such as articulation and fluency, or is it possible for players to achieve the desired outcome without incurring a high cost of conscious attention?** The third study this thesis presents discusses an evaluation method that addresses questions **RQ4** and **RQ5**.

A designer may decide to extend an existing musical instrument to **instantly** produce **a new sound** in response to a gesture. Otherwise, they may decide that a gesture affects the **subsequent sounds** produced by the **underlying original instrument**. **RQ6** asks **Does a gesture producing an unfamil-**

iar sound in real-time easily integrate into a music performance compared to a gesture modifying the sound of the instrument in the near future? Beyond the specificity of RQ6, RQ7 aims to set the stage for a research methodology asking how to evaluate whether the modification of an instrument can be integrated into players' execution? The fourth study in this document addresses this question and proposes principles to identify a research method to evaluate skills transfer on a modified version of a familiar instrument.

These research questions explore the role of auditory imagery, the impact of unfamiliar auditory feedback, the transferability of motor skills, the integration of gesture-based modifications, and the evaluation of skills transfer on modified instruments.

1.3 Statement of Contribution

Recipients

The contributions of this thesis will be valuable for new musical instruments designed for the kind of performances that value learnability, perfectibility and repeatability of musical interaction. Performers who might adopt this kind of instrument are generally interested in subtle control over musical aspects such as timing, volume, timbre, accents, and articulation [12]. For this reason, performers that participated in the research have a professional background in instruments that allow nuanced control over these parameters and value predictability of interaction. Examples of performers who joined the experiments include professional violinists, electric guitar players, and pianists. They were asked to play pre-defined music on notation using modified versions of the kind of instruments they studied. Instruments that were altered for the experiments were changed while maintaining predictable, repeatable and accurate pitch, volume, timbre, and articulation control.

A different set of musical values exists in communities like NIME. Morreale et al. published a study highlighting the values of NIME digital musical interfaces [2]. The study documented the general prevalence of music performance with exploratory and experimental aims. Performers in this area are interested in developing an "agency of the instrument within the performance process" [6, 2, 7]. Players have "a sense of shared authorship with the instrument" when they make music with it [7]. This and other sets of musical values are not the focus of this thesis.

The goal of this thesis is not to disfavour the improvisation of music using instruments that allow shared creative agency. The goal is rather to provide guidelines and evaluation methods to inform the design of new musical designs for the kind of players who participated in this research.

Contributions

This research presents a new, quantitative technique aimed at evaluating the process of skill transfer in musical performances involving augmented or modified instruments. The method is presented in chapter 5 through a case study involving professional musicians. The groundwork for a supplementary quantitative method is hinted in chapter 6. These methods, collectively, are aimed to inform the design of instrument modifications and augmentations for skills transfer.

A further contribution of this research is design principles for instruments' modification or augmentation to achieve skills transfer. The first principle emphasizes the need to shift the design lens from auditory feedback to sensorimotor imagery. This involves recognizing the significance of performers' ability to imagine the desired sound and activates corresponding motor programs. Designing instruments that facilitate this connection is crucial for retaining players' skills and enabling them to perform on modified or augmented instruments. The second principle highlights the diminished importance of auditory feedback in supporting performances on modified or augmented instruments. It suggests that performers can achieve enhanced fluency and skills retention by placing greater reliance on their internal representation of the gestures required for execution and the intended sound they aim to produce, as opposed to relying heavily on auditory feedback. These findings agree with the examined literature concerning the role of sensorimotor imagery in musical performance and have substantial evidence. The described design principles are the outcome of two studies presented in chapters 3 and 4. The specific context and conditions of these experiments suggest that further research may be necessary to fully affirm the universal applicability of these findings to any musical instrument within the realm of design modification. More information is provided in the discussion of the studies chapters 3 and 4.

1.4 Structure of the Document

Chapter 2 provides the background for this PhD. Chapter 3, chapter 4, and chapter 6 describe the four research studies designed to tackle the research questions presented in section 1.2. Chapter 7 summarises the studies' findings with respect to the initial research questions and provides insights for future work.

Chapter 2 provides the background, reviewing theories of embodied cognition and theories of sensorimotor control and an account of musical instruments' augmentation and modification. It reviews the literature on musical performance, musical instruments' embodiment, auditory imagery, skills transfer, and augmented and modified musical instruments.

Chapters 3 and 4 present two complementary case studies that look into design principles for skills transfer. Particular attention is devoted to the ability to perform pitch material accurately and to preserve fluency of execution. The study considers that a musical instrument designed to play tones from the 12-tone equal temperament may be modified or extended to allow the performance of micro-intervals or new sounds. However, players trained in musical repertoire based on the 12-tone equal temperament (such as the Western classical musical repertoire) may not have developed auditory imagery on micro-intervals. How well can their existing auditory imagery support skill transfer between the original musical instrument and its modified version? This question is answered in the first case study, where professional violinists are asked to perform micro-intervals.

An instrument modification may also change the mapping between a musical gesture and the resulting note (i.e. playing a particular note on a violin corresponds to a different note). This disrupts the connection between professional players' expectations and resulting auditory feedback. In the second case study, expert violinists performed short musical excerpts with a re-tuned violin in different conditions. How well and in which conditions could expert performers retain their fluency with a mismatched mapping between auditory expectations (imagery) and auditory feedback? Is the instrument's auditory feedback helpful in tackling this challenge?

Chapter 5 presents a third research study. The study discusses a quantitative method to evaluate whether expert players can execute skilled actions on an unfamiliar interface while keeping the focus of their performance on the musical outcome rather than on the technology itself. During the study, twelve professional electric guitar players used an augmented plectrum to replicate prerecorded timbre variations in a set of musical excerpts. The task was undertaken in two experimental conditions: a reference condition and a subtle, gradual change in the augmented plectrum's sensitivity, designed to affect the guitarist's performance without making them consciously aware of its effect. It is proposed that players' subconscious response to the disruption of changing the sensitivity and their overall ability to replicate the stimuli may indicate the strength of the relationship they developed with the new interface. The case study presented in this study highlights the strengths and limitations of this method.

Chapter 6 describes the fourth and last research study of this document. An instrument extension could add a button, a sensor or even a pedal to a traditional instrument. The technological extension could produce a sound immediately corresponding to physical interaction. Otherwise, it could influence the subsequent sounds produced by the instrument itself. Which kind o mapping and sonic actuation better preserve the performing quality of the execution? In other words, does a sound produced immediately in response to a physical gesture better integrate into the set of actions needed to perform a music passage? Do the performance timing precision and fluency deteriorate when the same physical gesture modifies the instrument's sound? The study looks into design principles for skills transfer and proposes a method to evaluate it.

Chapter 7 summarises the findings from the studies and presents future steps to query skills transfers in new musical instruments.

Chapter 2

Background

Learning to play a musical instrument is a complex task, involving the acquisition of sensorimotor skills over many years. The resulting relationship between performer and instrument is notable for its richness of experience, with some performers reporting that the instrument feels like an extension of their body.

At the outset of this chapter, it is pertinent to highlight the core concepts around which the upcoming discourse and exploration will center on: hermeneutic, embodiment, and auditory imagery. These concepts form the foundation of the investigations and findings presented throughout this thesis.

- The term **hermeneutic** taken in this thesis is concerned with the interpretation and understanding of musical performance and experience. It will be used for understanding a type of context that shape musicians' experiences with modified or extended instruments.
- **Embodiment**, as another essential concept, considers the intimate relationship between a musician and their instrument. It reflects the cognitive relation, the intricate, physical connection and the manner in which a musician's movements and physical interactions with an instrument influence the produced sound.
- **Auditory imagery** pertains to a musician's ability to internally generate and manipulate sound representations.

In the studies presented, we confront the challenge of evaluating embodiment and the impact of auditory imagery on musical performances externally, through the observable attributes of musicians' executions - such as pitch accuracy, dynamics, and timing. By investigating these aspects in a series of musical tasks, this research aims to gain a nuanced understanding of the relation that forms during the encounter between a professional musician and a modified or augmented instrument. The goal is not to privilege an objective method against subjective methodologies but rather to complement existing methods with something that is outwardly observable and repeatable.

In this chapter, each of these concepts — hermeneutic interpretation, embodiment, and auditory imagery — will be looked at in-depth, both individually and in relation to one another. They form the theoretical backbone of our exploration, offering distinct but interrelated lenses through which this research will examine the experiences of professional performers confronted with unfamiliar musical technology. Through this varied approach, this thesis aims to present a rich and comprehensive view of musicians' skills' adaptability for musical instruments' innovation.

2.1 Learning a Modified Instrument

J. O'Connor suggests four stages of perceptual-motor learning of a musical instrument [11]: unconscious incompetence, conscious incompetence, conscious competence, and unconscious competence.

A player will spend years advancing through these stages. First, they become aware of things they cannot do and then learn to do them through considerable effort and attention. Finally, they internalize the skills to the extent that minimal conscious attention is required [13, 10], freeing that attention for higher-level musical interpretation.

Changes to the instrument's physical or sonic characteristics can heavily impact performers' ability to use their existing motor skills[1] [15] and lead to impaired fluency, where it is impossible to play something in tempo with proper rhythm and intonation. In these cases, the instrumental modifications mean that conscious attention is again required for each operation, with a corresponding reduction in speed and precision [16].

Since developing new expertise on an instrument can take years, digital musical instrument designers have turned to strategies to repurpose existing skills on modified, or new instruments [17, 18, 19], often through the augmentation of familiar instruments.

Reasons to have a familiar instrument as a starting point can involve cultural references and familiarities. Overholt describes augmented instruments as tools to create "music that explores new sound worlds, yet still follows in the traditional musical training to a certain degree" [20]. According to such a per-

[1] A motor skill is a function that allows people to perform specific movements of the body's muscles to perform a particular task efficiently [14]. An example of this is playing a musical instrument. If the musical note played by a skilled violinist has a different pitch from what they read on the notation, they can accurately move the finger on the violin fretboard to fix it.

spective, augmented instrument performers should be allowed to do whatever they could do and more naturally. In addition to building on existing sensorimotor skills[2]. The augmentation should extend the instrument while keeping similar cultural reference points. For instance, while considering an augmented guitar, it should be possible to recognize it like a guitar [23], and it should be possible to associate it with certain kinds of music styles and performing techniques [17].

2.2 Embodiment

It is a common experience amongst skilled musicians that the instrument behaves as an extension of the body. However, the term *embodiment* has different notions according to different (though connected) research fields.

In this section, I start by focusing on the origin of the term, which is rooted in twentieth-century phenomenology. I then present its use in the cognitive science and embodied music cognition research fields, and I describe how the term is used in this thesis.

2.2.1 The Term Embodiment in Phenomenology

The term embodiment is rooted in the philosophy of phenomenology by philosophers including Merleau-Ponty [24], and Husserl [25]. The concept of embodiment implies that the human body acts as a mediator of the world (*Lebenswelt*) [26] and thus contributes to shaping our perception in combination with our senses [27]. Merleau Ponty argues that the body participates in creating a perceptual experience [24] and contributes to the formation of meaning.

Embodiment is also discussed as a concept that applies to the relationship between people and technological tools. Ihde argues that the act of playing a musical instrument enables an embodiment related to the world. Through repeated practice [28], the instrument becomes an extension of the body [29, 30].

Nijs argues that the musical instrument becomes an extension of the expert performer [31] in that the actual operations of manipulating it recede from conscious attention due to acquired sensorimotor skills [13]. "The feeling of having merged with the instrument is based on the incorporation of the instrument, which is characterized by the so-called perceptual illusion of non-mediation" [31].

The idea of a transparent tool that becomes an extension of the body seems connected to Heidegger's notion of technology as *present-at-hand* [32]. Ham-

[2]sensorimotor skills [21, 22] are the skills we use to take in information about the world through our senses and to respond to that information through the body's movement. For example, if we see a red traffic light, we stop walking.

mering with a hammer, for example, the focus of the interaction is centred on the goal: hitting the nail. The more competent the workman, the less aware he is of the hammer at all: he simply hammers away. The movements in his hand are realized in the movements of the hammer in such a way that the hammer serves as an extension of the workman's hand. In this way, the hammer and the workman are together, entangled [32].

2.2.2 The Term Embodiment in Music Cognition

Empirical research conducted in the field of cognitive science produced evidence that the body contributes to the structure of human experience [33] and highlighted the influence between perception, action and cognition [34]. These elements concur in constructing and defining human experience [35].

Arguing for an intrinsic connection of perception and bodily action, Fuchs proposes that *what the organism senses are a function of how it moves, and how it moves is a function of how it senses* [36]. Drawing on these principles, embodied music cognition [37] established itself as a research field studying the role of the human body concerning musical activities, including performing. Like other musical activities, music performance relies on cognitive processes dependent on the link between *cognition, human motor system, gestures and body movements* [38, 39].

2.2.3 Use of the Term Embodiment in this Thesis

In both phenomenology and cognitive science, the term embodiment addresses the importance of the body in mediating and constructing perceptual experiences. As such, embodiment seems to represent an interface between disciplines. Scholars like Ihde and Nijs extend the meaning of the term embodiment to the relation between players and instruments where the instrument becomes an extension of the performer's body (hence it becomes embodied). In this sense, the instrument senses, mediates, and participates in constructing and defining the musical experience as it becomes a part of the performer's body.

Drawing on the research by Ihde, Nijs, and Leman, this thesis uses the term embodiment to describe what I define as *a relational property between performer and their musical instrument that depends on both performer and instrument and an appropriate balance between skill level and challenge of the activity.* The actions of manipulating the instrument recede from consciousness and become, to a certain extent, automatic (similarly to the workman hitting the nail described by Heidegger). Performers gain the ability to adjust their actions rapidly and precisely to correct errors or to shape their performance [13, 40]. As the instrument disappears, to some extent, from the performer's conscious attention and becomes an extension of their body [41, 28, 29, 13], the performer's attention moves to the outcome of their sensorimotor activity: music.

In this thesis, I'm particularly interested in outwardly observable performance patterns, in contrast to the internal subjective experience of the performer. The research studies I conducted query the impact of instruments' modification on sensorimotor skills that contribute to the emergence of embodiment relationships [3].

2.2.4 Hermeneutic Relation between Performer and Instrument

Embodiment is one of several possible performer-instrument relationships elicited by digital musical instruments [28, 42]. Don Ihde, in his account of hermeneutic phenomenology [43], discusses a further possible relationship. In a *hermeneutic* relationship, an instrument is a tool external to the body whose information we have to interpret [28]. An example of hermeneutic relation can be found in reading a thermometer, or in a musical context, with specific digital musical interfaces like a guitar tuner[4]. When technology is hermeneutically encountered, it offers readings of the world that the user has to interpret. Technology is no longer an extension of the body. Instead, the immediate perceptual focus is the technology itself [23, 17]. Performers think in *analytical mode*, and their attention is directed towards analyzing and decoding the information they perceive. The performer's attention shifts toward the mediating elements of the activity [44]. Musicians are requested to continually move the focus of their performance between the interface and the sonic results, leading to disruptions in the performing fluency and accuracy which can be observed and registered as auditory and visual cues to identify such relation.

The concept of embodiment and hermeneutic relationships, although related, are distinct phenomena within the context of performer-instrument interactions. While embodiment refers to the cognitive and experiential integration of the instrument into the performer's body and the immediate perceptual focus on the music produced, hermeneutic relationships involve the interpretation of external instrument information that requires conscious interpretation.

In the context of hermeneutic relationships, the performer's attention shifts towards a more conscious analysis and decoding of the information provided by

[3]Such relationships could include adjusting the pitch on the fretboard of a violin, regulating the strength in plucking the strings with a plectrum, tapping a foot pedal with accurate time precision within a musical performance

[4]Guitar tuners generally report the pitch of the instrument as a number to be interpreted rather than a direct experience of sharp or flat.

the instrument. This can be observed in scenarios such as reading a thermometer or using digital musical interfaces like guitar tuners, where the performer interprets numerical readings rather than directly experiencing the musical qualities.

By presenting the concept of hermeneutic relationships as an alternative to embodiment, this thesis aims to explore the performer's attentional focus during musical interactions. In this thesis, I propose that it is important to investigate where the performer's attention is directed to understand the extent to which a musical interface is being embodied.

The characteristics of the interface can influence the type of relationship that emerges. In fact, technological objects are never neutral. In them, we find programs of action, manuals of behaviour, and political and socio-cultural constructions, including aesthetic tendencies[28]. For example, the Buchla synthesizer tells us that a keyboard featuring a fixed-pitch structure is unnecessary to play it. In contrast, a piano tells us that microtones are of little importance [45]. Feenberg also writes about the impossibility of neutral tools [46]. Musical instruments have a specific affordance that reflects the design choices, values, and the cultural and historical context in which they were developed as they are refined over years of development [20, 28]. However, solely relying on the physical form of the technology may lead to oversimplified conclusions. For instance, a violin can be encountered hermeneutically depending on the performer's cognitive state and attentional focus, even though it is traditionally associated with embodiment. A simple modification, such as re-tuning the violin's strings, can result in performers establishing a hermeneutic relationship with the violin [16].

To determine whether a player embodies a musical instrument or engages in a hermeneutic relationship, it is crucial to consider where the performer's attention is directed and how their perception and action are linked [13]. These questions serve as important factors in understanding the nature of the performer-instrument relationship.

Overall, this thesis aims to explore the interplay between embodiment and hermeneutics in performer-instrument interactions and delve into the role of attention and cognitive state in shaping these relationships and it provides a comprehensive examination of these phenomena within the context of musical interfaces.

2.2.5 Embodiment and Instrument Modification

Embodiment is one condition that allows expert musicians to perform fluently with their musical instruments. Reaching an embodied relation with a musical instrument can require years of training.

Although embodiment is one of several possible performer-instrument relationships [42], there is significant interest in developing new instruments which support the level of functional transparency achieved by skilled performers of acoustic instruments. Since developing expertise on an instrument can take years, NIME designers have turned to strategies to repurpose existing skills on new instruments [17, 18, 19], often through the augmentation of familiar instruments.

However, modifying a traditional instrument does not guarantee the embodiment of the new design. Changes to an instrument's physical or sonic characteristics can easily disrupt embodiment [15]. Compounding the challenge, the same pairing of performer and instrument may give rise to both embodied and hermeneutic relations [47]. In particular, performers of a modified instrument might have an embodied relationship with the underlying traditional instrument, which was developed over many years of training, while the modified instrument behaviour remains unfamiliar. The motor skills and the abilities developed on the traditional side of the instrument (which allows its embodiment) does not necessarily translate to its modification. An instrument modification may require new skills, such as performing certain gestures that are not generally performed on the original design.

To evaluate the success of a new design concerning embodiment, it is essential to identify the type of relationship musicians establish with the instrument and where their attention is directed. Are they embodying the extension of their instrument? In other words, do the actions of manipulating the instrument recede from consciousness and become, to a certain extent, automatic? Otherwise, are they consciously thinking about the instrument, the actions they need to perform, or the sound they need to play?

2.3 Auditory Imagery

The concept of auditory imagery is crucial to the understanding of this thesis, primarily in the context of how musicians approach new or modified instruments and how designers can leverage existing musical skills.

Auditory imagery is a critical component of musical performances [48, 49, 21], acting as a mental tool that musicians utilize to hear or simulate sounds in their minds without the presence of actual, external sound [50]. This ability is crucial for instrumental performances as it forms part of a predictive mechanism in which performers first imagine the desired sound within their minds, the gestures they need to perform, and shape their physical movements on their instruments to achieve that sound[51].

Auditory imagery, motor imagery, and sensorimotor imagery are interrelated yet distinct concepts that play significant roles in understanding human cognitive processes, especially in the context of musical performance. Auditory imagery pertains to the mental representation or “hearing” of a sound in the absence of any external auditory stimulus. It allows individuals to recreate and manipulate complex auditory experiences internally. Motor imagery, on the other hand, is the mental rehearsal or simulation of a specific action without any overt physical movement, effectively “feeling” or “envisioning” movement in one’s mind. Sensorimotor imagery seamlessly integrates both auditory and motor components. This integrated form enables an individual to mentally rehearse a task, considering both the sounds associated with the task (e.g., playing a musical note) and the motor actions required to produce those sounds. This comprehensive imagery process plays a vital role in the preparation and execution of complex tasks, especially in disciplines where precise coordination of sound and movement, such as in musical performances, is essential.

A further element of the musical process is auditory feedback, which is the resultant sound from the musician’s physical interaction with the instrument. This feedback acts as an immediate form of self-assessment for musicians, enabling them to make delicate adjustments in their performance to refine elements such as intonation and articulation.

In traditional musical settings where musicians have a strong, embodied connection with their instrument, auditory imagery and feedback play a vital role [52]. During performances, musicians can instinctively generate auditory expectations or predictions about the sound outcome [51]. These predictions guide them in executing intricate sensorimotor tasks on their instrument to realize the predicted sound [9], and then modify their performance by comparing these initial expectations with the actual sound feedback received from the instrument [9].

Research has shown that auditory imagery can assist musical performances even when there’s an absence of auditory feedback or in instances where the feedback is disrupted. For instance, studies have found that musicians with advanced auditory imagery capabilities were less affected by a lack of auditory feedback when practising novel musical pieces [53]. It was also observed that musicians who possessed highly developed auditory imagery skills demonstrated greater pitch accuracy and consistency in timing, even in the presence of auditory interference [54]. Furthermore, the absence of auditory feedback seemed to intensify the use of imagery [55].

It’s important to note that apart from pitch, expert musicians can mentally imagine other musical aspects like duration, timbre, and loudness, even in the absence of auditory feedback [56].

Regarding articulation, researchers propose a correlation between auditory imagery and dynamic representation - a concept referring to changes over time [57]. They suggest that internal or imagined speech functions similarly to overt or audible speech, serving many of the same purposes [58]. This concept of functional continuity implies a connection between the characteristics of an imagined sound and the actual sound stimulus. This continuity is seen as evidence of the parallels between auditory imagery and auditory perception, supporting theories that link perception with motor actions, and demonstrating that auditory imagery can prefigure subsequent perception and action [57]. Notably, auditory stimuli often involve dynamic representation, suggesting that auditory imagery might likewise involve dynamic representation. In line with this, it has been argued that temporal information is a vital component of an auditory image [59].

For the research questions of this thesis (see section 1.2), auditory imagery becomes important in a few key ways:

- **Transference of Skills**: musicians who have developed a high level of proficiency on a specific instrument will likely have strong auditory imagery associated with their actions on that instrument. When confronted with a new or modified version of the instrument, they may leverage this auditory imagery to predict the outcome of their actions. This might aid in a more effective transfer of their existing skills to the new context (as proposed in research questions **RQ1** and **RQ4**).

- **Performance of New Sounds**: when an instrument is modified to produce unfamiliar sounds, the performer's ability to imagine these target sounds (auditory imagery) could affect their accuracy and response time (**RQ2**). Strong auditory imagery might help performers more quickly adapt to these new sounds.

- **Performance Improvement**: **RQ3** raises the possibility that focusing on an internal representation of sound (auditory imagery) instead of the actual auditory feedback from the instrument could improve performance. This might be particularly true in situations where the feedback from the instrument is unfamiliar or confusing.

In summary, the importance of auditory imagery in this thesis lies in its potential to aid musicians in quickly adapting to new or unfamiliar instruments, making the design process more effective by enabling better use of existing skills. It's an intrinsic cognitive ability that musicians rely on, and understanding its role can significantly enhance the field of instrument design.

2.3.1 Measuring Auditory Imagery

The existing literature mentions the Bucknell Auditory Imagery Scale (BAIS) as a self-report measure to query aspects of Auditory Imagery [60, 61]. Measurable features of auditory imagery include vividness, which refers to the clarity of an image; controllability, which relates to the ease and accuracy with which a person can manipulate an image; and precision of reference, or the extent to which the image accurately reflects the object it represents. Vividness is rated from 1 (no image) to 7 (as vivid as actual sound), and control of auditory imagery is rated from 1 (no image) to 7 (extremely easy to change) [60]. High scores on vividness predict fewer source memory errors from imagined tunes on a recognition test and better performance on pitch imitation tasks. Halpern suggests that *high scores are related to hemodynamic response and grey matter volume in several brain areas that are involved in auditory imagery* [60]. People vary in these skills, mainly developed by musicians with more than seven years of training [62]. Professional musicians scored higher than non-musicians when asked to recognize reversed familiar melodies and to identify exact reversals [62].

Examining auditory imagery can be challenging. Subjective methods like BAIS may lead to incomplete or biased data as they rely on participants' questionnaires. Compounding the challenge, auditory imagery is part of a multimodal process [9], and it is linked to motor imagery as well as phenomena like audiation [63]. Auditory imagery is an internal mental process whose effects might be observable but seems difficult to isolate and measure directly.

The studies presented in this thesis tackle the challenge of evaluating auditory imagery *externally* by looking at aspects of musicians' performances (pitch accuracy, dynamic, timing) in a series of musical tasks. As for embodiment, the goal is not to privilege an objective method against subjective methodologies but rather to complement existing methods with something that is outwardly observable and repeatable.

2.3.2 Auditory Imagery and Motor Imagery

Auditory imagery in trained musicians is strongly connected with motor imagery. Auditory images function as a bridge between perception and action [64], leading to activation in motor planning areas that guide planning movement [65].

According to Clark et al., imagery, as used by musicians, involves not only the melodic and temporal contours of music (which we might refer to as auditory imagery) but also a sense of the physical movements required to perform the music [51]. Likewise, Keller talks about Musical Imagery while defining a broader multimodal process. "Individuals generate the mental experience of auditory features of musical sounds, visual, proprioceptive, kinesthetic, and tactile properties of music-related movements" [9].

Auditory imagery, or the ability to imagine sound events, is a component of an interconnected process also featuring motor imagery and including the sensorimotor system used to achieve the actual sonic result. In a study concerning the integrated roles of interpretation, imagery and memorization [66], Holmes reported a strong connection between auditory imagery and motor imagery, resulting in "a mental impression of what the music will sound like can generate a sensation of what it would feel like to play the music."

On the same topic, Godøy reported how gestural imagery could be, on the other hand, instrumental in triggering and sustaining mental images of the sound [67]. Then auditory and motor imagery are connected and can influence each other during a music performance. Keller pointed out that these mechanisms "support the generation of anticipatory images that enable thorough action planning and movement execution that is characterized by efficiency, temporal precision, and biomechanical economy." [9]

This process particularly characterizes musical executions of performance majors [52] who use imagery to maintain focus, predict mistakes, recover from errors, and manage mental and physical fatigue representing an effective component of music execution [9]. Clark et al. reported how many famous musicians had stressed the importance of imagery in traditional music history [51]. For example, Anton Rubinstein was known to practice on a paper keyboard to further develop his auditory representation of his music. Mellet et al. and Kosslyn et al. demonstrated a functional equivalence between imagined and played executions by studying the auditory and motor systems involved in musical performance [68] [69]. Also, they found that musical perception and imagery engage similar regions within the auditory cortex.

According to Connolly et al., auditory imagery is beneficial both during music practice and music performance [70]. It presents a series of significant benefits, including improving learning and memory, overcoming technical difficulties and developing skills, heightening sensory awareness, and enhancing general confidence and resilience on stage.

It has also been demonstrated that auditory images function as a bridge between perception and action [64], leading to activation in motor planning areas and constituting a bridge that serves to guide planning movement [65] [71]. Specifically, it was assessed that having a deficit of mental imagery and a lack of vividness (measured using the BAIS) lead to poor singing (i.e. a poor imitation of pitch through singing). These results support the hypothesis that auditory imagery is coupled with motor imagery and that pitch imitation (at least through singing) relies on auditory imagery.

María Herrojo Ruiz et al. propose that pianists also use a mental representation of the performance to predict errors [72]. This prediction reflects on their actions as wrong keys are played with lower velocity showing that they know they will be wrong even before that happens.

2.3.3 Auditory Imagery and Notation

Simoens et al. suggest that a music performance achieved by reading a score requires an initial translation of the visual material into auditory information [73]. According to Berz and William, such a translation involves a tonal loop in working memory for pitch [74]. Simoens et al. also suggest such translation to auditory imagery might be connected to motor preparation to perform the music on the score [73].

Schön et al. highlight the importance of auditory imagery in creating an auditory representation [75]. A group of musicians were asked to compare the final note in a novel short melody with the last note in a previously viewed melody notated as a musical score. Participants should assess whether the final note of each excerpt was identical. The content of the notation influenced participants' responses, supporting the idea that visual notation can contribute to creating auditory representations.

Thanks to auditory imagery, musicians can imagine the continuation of a melody and anticipate a consequent auditory stimulus [76, 77, 78]. This ability is related to familiarity with the music material involved in the activity. Kraemer and colleagues asked musicians to listen to familiar and unfamiliar musical passages. Musical excerpts had a short duration and were characterized by silent gaps. During the gaps, participants reported hearing a continuation of the music in imagery for familiar music passages. These results are in line with [79], suggesting that auditory imagery can help predict musical content.

From a neural perspective, Hope and colleagues reported that when a subsequent auditory sequence is compared to a previously viewed musical score, the locus of activation shifts from the visual cortex to the auditory cortex [80]. This may suggest that reading a graphic score can evoke related auditory imagery.

2.4 Performance Based Studies

The decision to use a performance studies-based approach in this research is informed by the nuanced understanding of musical performance as an elaborate intersection of cognition, embodied actions, and instrumental interactions as it emerges from the literature review.

The existing literature covers different methods to query musical performance and skills transfer. These include self-report using questionnaires, experience sampling [81], measurement of physiological responses and naturalists' performance-based tasks [51].

Nonetheless, examining music performances can be challenging. Subjective methods like self-reports, questionnaires and experience sampling methods can report incomplete or biased data, while physiological measurements can be intrusive and difficult to interpret. Compounding the challenge, performer-instrument relationships rarely display only one mode. Performers of augmented traditional instruments might be able to retain expertise (unconscious competence) with the underlying traditional instrument while the augmented behaviour remains unfamiliar. Morreale et al.(2019) observed two patterns of behaviour when traditional instruments and augmentation are closely intertwined [82]. In the first case, the performer lets the augmentation partially or totally disrupt their playing of the traditional instrument since the new techniques are not yet familiar. In the second case, the performer ignores the sonic result of the augmentation and focuses on regulating their performance according to what they would normally do on the traditional instrument.

Performance-based studies stand out as particularly valuable within a quantitative research framework. Through the lens of the existing literature, it becomes clear that these studies have been instrumental in delivering a nuanced understanding of musical performance. They make use of quantifiable metrics to query complex phenomena, offering a quantitative comprehension of the aspects involved in musical expression.

The research reviewed in section 2.3.3 shows the utility and significance of performance-based studies. Schön et al., Kraemer and colleagues, and Hope et al. leverage this approach to uncover the role of auditory imagery in musical performance.

In Schön et al.'s study, a performance-based approach allowed for the empirical examination of how musicians employ auditory imagery in the reading and interpretation of a musical score [75]. In a similar vein, Kraemer and colleagues used a performance-based study to explore how auditory imagery facilitates the continuation and anticipation of musical passages [76, 77, 78]. By observing musicians as they listened to familiar and unfamiliar musical excerpts, the study revealed that the participants reportedly heard a continuation of the music during silent gaps in familiar passages, indicating the predictive power of auditory imagery. Lastly, Hope et al.'s research employed a performance-based approach to investigate the neural correlates of music reading and auditory imagery [80]. By examining the shift of neural activation from the visual to the auditory cortex during the comparison of a subsequent auditory sequence to a previously viewed musical score, the study shed light on the cognitive and neural underpinnings of music performance, enriching our understanding of the process.

More generally, performance-based studies are compelling from a quantitative research perspective in the context of musical instrument design and design and instrument design modifications.

- **Objective Metrics:** Performance-based studies often generate objective, quantifiable metrics like timing, accuracy, speed, and efficiency.

- **Interface Evaluation:** The performance of a user interacting with a system can often reveal the usability and effectiveness of an interface. Therefore, performance-based studies can quantify the impact of different design decisions and guide evidence-based recommendations.

- **User Behavior Analysis:** By capturing and measuring user behaviour in response to various system designs or tasks, performance-based studies can provide valuable insights into user preferences, abilities, and strategies. These insights can then inform the design of musical interfaces.

- **Task Completion:** Performance-based studies often involve tasks that simulate real-world scenarios. These tasks can be used to evaluate how effectively users can accomplish these tasks using the system, interface, or prototype being examined.

- **Iterative Design:** Performance metrics can provide tangible and practical suggestions for refining designs in an iterative process, which can be a crucial part of musical instrument design and research.

- **Benchmarking:** Performance-based studies provide a means to establish benchmarks for specific tasks or interfaces. This allows for a quantitative comparison of different systems or design iterations.

2.5 Conclusions

The concepts reviewed in this chapter were presented and examined to provide a theoretical framework for this research, presenting interconnected perspectives through which this research assesses the experiences of professional performers when faced with unfamiliar musical technology.

The comparative exploration of embodiment and hermeneutic interpretation is instrumental to this thesis. Both concepts, each powerful in their own right, offer complementary lenses to examine the relationship between musicians and their instruments. Embodiment reflects the physical and cognitive connection that performers have with their instruments, shaping the outcomes of their musical expressions. It emphasizes the inherent physicality and sensory experience of playing music, and how that corporeal interaction can influence and be influenced by the musical output. On the other hand, hermeneutic interpretation reflects the act of decoding the layers of meaning, context, and experiences that musicians bring to and derive from the performances of their instruments. It underscores the interpretive, subjective aspects of musical instruments and performance. The comparative examination of these concepts allows us to gain an understanding of musicians' interaction with their instruments offering a more comprehensive picture of the musician-instrument relationship.

One of the main challenges of this thesis is to assess embodiment and auditory imagery from an external perspective, considering observable performance attributes such as pitch accuracy, dynamics, and timing. Through engaging expert players in various musical tasks, this research inquires about the nuanced relationships that evolve when a professional musician interacts with a modified or augmented instrument.

The concepts presented in this chapter established the theoretical ground of this investigation, supplying interconnected perspectives for querying the experiences of professional performers when faced with unfamiliar musical technology. By adopting a performance-study research approach, this thesis strives to provide a comprehensive, objective, and detailed understanding of musicians' skill adaptability in the face of musical instrument modifications.

Chapter 3

Quarter tones study

The experiment described in this chapter has been produced from the paper *Design for Auditory Imagery: altering instruments to explore performer fluency* [16].

3.1 Introduction

Instrument modification can achieve new sonic possibilities, including extending the chromatic palette[1] of an existing instrument. The resulting design can maintain features of the traditional instrument while affording new or unfamiliar pitch elements like microtones. As discussed in the introduction and background of this thesis, transferring existing skills to perform the instrument modification may be desirable for shortening the training process and possibly facilitating the design adoption.

However, to what extent can a familiar interface support transferring pitch expertise and skills to perform unfamiliar pitch material? What players' skills are challenged when a familiar instrument is modified to extend its chromatic palette? Auditory feedback, auditory and motor imagery, and kinaesthetic ability are features of musical executions that are generally discussed in musical performance. Which of these features and skills are particularly important to perform unfamiliar pitch material? Therefore which of them should be mostly considered in the design of a modified instrument?

Twelve professional violinists trained in the 12-tone equal temperament are asked to play micro-tones on their violin. The experiment covers the following research questions: *how accurately can violinists use a familiar interface to play pitch material for which they have not developed auditory imagery? If the actual target sound is unfamiliar (quarter tones), how does that differ in*

[1] The variety of pitches that an instrument affords to perform

terms of accuracy and response time to being able to play ordinary chromatic music on a violin? To what extent is auditory feedback useful to adjust their pitch accuracy when playing quarter tones? The hypothesis in this case study is that even expert musicians will perform with lower accuracy even on a familiar interface (the violin) when confronted with pitch material sitting outside their auditory and motor imagery. In this sense, the experiment outcomes focus on a fundamental limitation of making a new instrument. No matter what physical interface or mapping is chosen. Regardless of the physical form of the instrument and its configuration, if a musical feature does not feed into a culturally defined structure of how people think they hear and play music, they struggle to play it.

3.2 Participants

Twelve professional violinists answered an open call sent through music schools and participated in the experiments. Each violinist filled out a questionnaire before the start of the study indicating: their demographics, years of training, and repertoire. The participant group consisted of individuals whose ages ranged from 22 to 46 years old. The gender distribution positions towards females, who constituted 67 per cent of the group, while males represented the remaining 33 per cent. In terms of their study duration, it spanned from 12 to 36 years.

The questionnaire included questions to verify whether participants had prior familiarity with microtones. The microtones study challenges players' sensorimotor imagery because it requires them to play quarter tones. A pre-existing ability to play them may impact the study data. Participants replied they were unfamiliar with microtones before the beginning of the study. During the study, they played their own instrument.

Participants were trained in western classical music. They had ABRSM's Grade 7 [2] and above. They were paid an hour at a professional rate (50 GBP). The two studies took place in a music studio within the Queen Mary University of London (Mile End campus) and lasted one hour.

The Queen Mary Ethics of Research Committee granted approval for this study, along with all other investigations discussed in this thesis.

3.3 Study Description

Violinists were asked to play semitones and quarter tones notes using their violins. The term *semitone* describes tones from the 12-tone equal temperament.

[2] ABRSM (Associated Board of the Royal Schools of Music) is an accredited board awarding exams and diploma qualifications in music within the UK

Semitones are usually found in music passages in the Western tonal music tradition and are familiar to participants for their coherence with the type of studies and training they followed. The term *quarter tones* describes tones from the 24-tone equal temperament. Quarter tones were chosen for being rare in the Western traditional music repertoire. They are a type of pitch material for which the selected violinists did not have developed imagery.

Players were introduced to the study by presenting the tasks and an explanation of the quarter tone notion and notation (what is a quarter tone and how it is visually notated on a score in this study). Quarter tones were notated using the symbols ♭ ♮ ♯ indicating respectively to play: one quarter tone below a flat note, one quarter tone below a natural note and one quarter tone below a sharp note. Figure 3.1 shows a semitone stimulus example while fig 3.2 quarter tone note example where the symbol indicates to play a pitch equal to one quarter tone below a C sharp. Musicians were asked to avoid using vibrato in their executions[3].

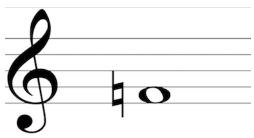


Figure 3.1: a semitone stimulus.

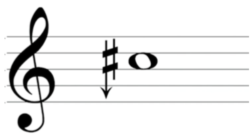


Figure 3.2: a quarter tone stimulus.

Performers were asked to play a chromatic scale to warm up and a quarter tone scale to further familiarise themselves with the quarter tone notation. Consequently, players tackled the challenge of playing quarter tones and semitones in three subsequent conditions: reading notation on display connected to a laptop ($notation_a$), listening to audio recordings through audio speakers (*audio playback*), and again reading notation on display ($notation_b$). The second notation section was included to account for the potential familiarisation with the sound of quarter tones that could have happened during the audio section. The Notation$_\mathrm{a}$ section comprised twelve stimuli, the audio playback section encompassed twenty-four stimuli, and the Notation$_\mathrm{b}$ section contained twelve stimuli. Notation$_\mathrm{a}$ section included twelve stimuli, the audio playback section included twenty-four stimuli, and the notation$_\mathrm{b}$ section included twelve stimuli.

The notation sections provide visual instructions for what pitch needs to be played. Because they were asked to play the stimuli in the first position[4].This indication and the notation could allow them to identify the spot on the fret-

[3]Vibrato is a musical effect produced by a regular and subtle pitch variation. The study focuses on measuring pitch accuracy in participants' executions. Executing quarter tones and semitones characterised by vibrato would have biased the data and influenced analysis results

[4]Violin, as well as other string instruments such as guitars, allow to play the same pitch on different strings and positions on the fretboard.

board where the stimuli need to be played.

The notation allows players to connect with their imagination of how quarter tones and semitones should sound. It also gives a visual reference about where to locate them on the violin's fretboard to start the executions. The notation does not provide an actual auditory reference for the stimuli. Participants' adjustments *after* the start of the note rely on their sensorimotor imagery and kinaesthetic feedback. There are two notation sections. The first starts before the audio section, and the latter once the audio section is completed. Data collected in the latter notation section is meant to assess the impact on players' performance from being exposed to auditory reference quarter tones.

The audio section provides a clear auditory reference to what needs to be played. Audio stimuli indicate how they should sound. However, they do not provide any instructions about their position on the fretboard.

The study aims to investigate participants' utilization of sensorimotor imagery (encompassing both auditory and motor components) in playing quarter tones, examining their accuracy at the note onset and adjustments throughout the note duration.

Comparison between semitones and quarter tones, as well as between the audio and notation sections, aims to elucidate participants' utilization of auditory and motor imagery in playing quarter tones. If participants can effectively engage their sensorimotor imagery in playing quarter tones, we would expect their accuracy to be similar between the notation and audio sections throughout the duration of the note. Their ability to imagine how the stimuli should sound enables them to make adjustments to their pitch accuracy during the note's duration, regardless of the presence or absence of auditory references for the stimuli.

Alternatively, if participants rely primarily on motor imagery to play quarter tones, their pitch accuracy at the beginning of the notes could be comparable between the two sections, irrespective of the presence of visual indications (notation) specifying the finger position on the fretboard for initiating the execution. We may also expect pitch adjustments during the notes' life to be higher in the audio section because the audio stimuli provide an auditory reference that can be memorised and compared to what they are playing. However, since the audio section does not provide participants with a clear indication of where to position their finger on the fretboard, their pitch deviation performance would be higher at the beginning of the notes compared to their notation section performances. In the second notation section, we may observe some improvement in pitch accuracy because of players being exposed to auditory stimuli in the audio section.

The first three stimuli in section $\text{Notation}_{\text{a}}$ and section Audio were regarded as trial stimuli and were not included in the data. They helped participants familiarise themselves with the study's task. During the whole study, every note was presented only once to each participant to prevent potential learning effects and related biases. Each time the performance of a stimulus was completed, I manually activated the following stimulus (more details in the apparatus section). The order of quarter tones and semitones stimuli was randomised for each participant to prevent biases coming from a certain sequence of stimuli.

In Notation_a and Notation_b sections, participants were instructed to perform the stimuli as soon as they felt ready. In the audio section, they were asked to wait until the end of the audio stimuli playback. They were asked to wait until the end of the playback so they could not adjust their pitch by playing on top of the stimuli playback.

3.3.1 Apparatus

The apparatus can be divided into devices and software to generate and display the stimuli and devices and software to collect participants' performance data.

A display and audio speakers were connected to a laptop to present the stimuli. The display showed the notation stimuli while the speakers presented the audio stimuli. The Laptop ran a custom Processing [83] code to generate, display, and log the notation stimuli and to log the audio stimuli. The audio stimuli were generated and outputted using a custom Super Collider [84] FM script synthesising a string sound. The script is based on the example *20 70 1* included in the Amsterdam Catalog of Csound Computer Instruments [85]. The portion of the study apparatus to record performance data included: a DPA microphone[5] to capture the audio coming from the violin, a light sensor to detect the moment in which the notation was displayed on the screen, and a *Bela*[86]. Figure 3.3 displays the elements included in the apparatus.

The light sensor was taped to the monitor to detect a blinking box on display generated by Processing and visualised on the right low corner of the screen each time a new notation stimulus was presented. Logging the stimulus timestamp with the sensor was preferred to logging the moment in which the stimulus was generated or outputted by Processing. The sensor is not influenced by potential delays caused by the computer itself and detects the moment the stimulus is displayed more accurately. The box size was 50 square pixels and barely noticeable to the participants because of its size and because it was covered by the sensor, which was, in turn, mounted using a piece of tape. Because of the tape, the ambient light did not influence the sensor, and the blinking box was not visible to the participants. Turning the box from white to black and then

[5]The microphone was applied to the violin body.

white over a 1-second period was enough to cause the sensor signal to change from 1 (the previous stimulus is being displayed) to 0 (a new stimulus is being displayed) to 1 (the new stimulus is the current stimulus). Figure 3.3 shows the change in colour of the box according to the experiment status.

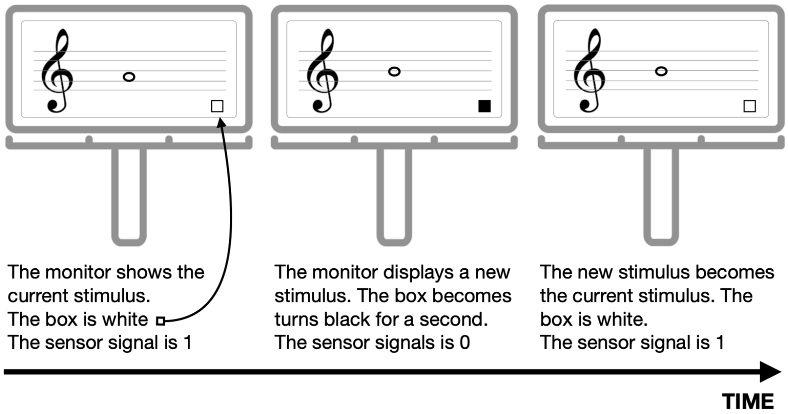


Figure 3.3: change of state of the blinking box.

The Bela platform hosted a C++ script to record and store the violin audio signal. The board receives and records the light sensor signal while keeping it in sync with the violin audio signal. Keeping the two signals in sync was crucial to measure the time interval between when a notation stimulus was presented and when the performer started its execution.

Figure 3.4 displays the elements included in the apparatus.

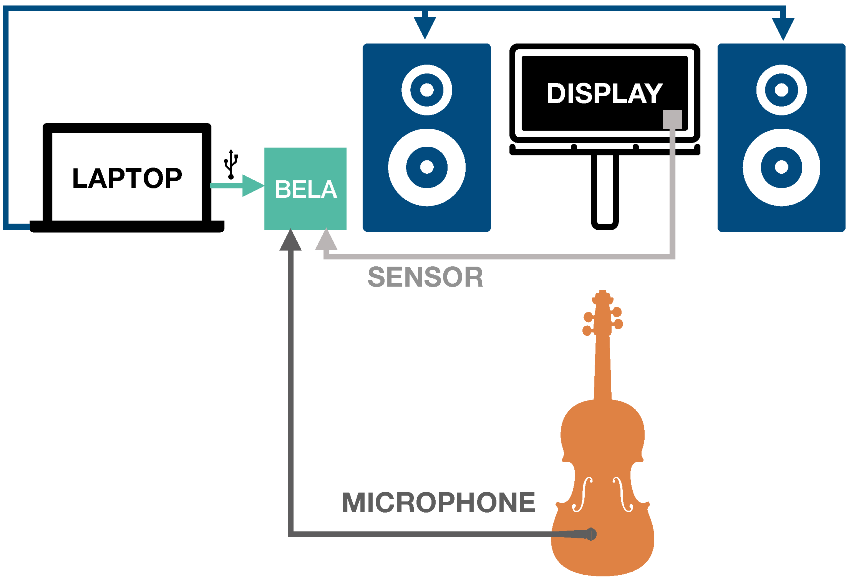


Figure 3.4: study 1 apparatus.

3.4 Data analysis

3.4.1 Building the database for analysis

The following paragraphs describe the extraction of the data from the recorded audio and sensor signals and the organisation of the data in a single data frame. The data published in this chapter is in line with the data and findings discussed in the publication by Guidi et al. [16]. However, some of the figures differ as they result from an advanced and thorough analysis achieved after two years of additional PhD training. Experience and new skills related to R Studio allowed for an enhanced, clean and organised dataset. R studio [87] helped identify and correct minor errors caused by the *pYin* algorithm (like pitch frequencies displayed as the double of the actual note fundamental frequency). Moreover, in the following chapters, medians are preferred to means (reported in the NIME paper cited above) as they are less affected by outliers.

Building the database started by importing the violin and sensor recordings in the audio editor Sonic Visualiser [88].

Pitch data

As this study focuses on participants' pitch accuracy, I considered the portion of each note where the pitch is detectable and is the direct result of participants playing the strings with the bow. I defined this portion of each note by manually placing two-time markers. The first time marker was positioned at the point where the waveform stopped manifesting a chaotic behaviour. Such behaviour is caused by the bow touching and scraping the strings at the beginning of each note execution. The pitch is highly variable and hard to measure in this very short portion of the note, which we could quantify as a few milliseconds. That's why this note segment was excluded from the analysis. I positioned the second marker when the violin bow detaches from the strings and causes the volume to gradually decrease while the violin body and the strings keep resonating by themselves. Its location was identified as the portion of the note's amplitude envelope where the volume steadily decreases. This phase is generally described as the note's *release*. In this note's segment, the resulting pitch tends to adjust according to the violin's acoustic body rather than the player's agency. This portion of the note was excluded from analysis because it is not the result of the violin being directly played.

Figure 3.5 shows a violin note performance which is time-marked as described.

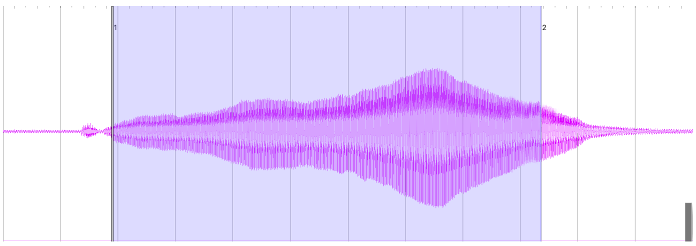


Figure 3.5: a note performance with time markers. Markers are the boundaries of the light blue area. This is the portion of the note considered for data analysis.

I selected each note's region standing between each pair of markers to calculate its pitch frequency envelope. The envelope is calculated using the monophonic pitch tracking algorithm pYin [89] [6]. The algorithm was set to operate at 512 samples per block with a 256 hop size. The resulting data were exported as CSV files. Each file identified a note and consisted of a list of frequencies paired with timestamps at which they happened.

I imported the CSV files in R Studio and checked for frequency bins erroneously generated by the Pyin algorithm. Frequency bins with half or double the precedent or subsequent frequency bin value or with a frequency value below the violin pitch range. Erroneous values were corrected to reflect the actually performed frequency.

By looking at the frequency timestamps column, I then computed the median

[6] The pYin algorithm is available as a plugin in Sonic Visualiser

frequency of each note's start mean and end segments considering the following time windows:

- start time window: 100 ms following the attack phase
- mean time window: the time interval following the start of the note up to the 100 ms before the end of the note
- end time window: the last 100ms before the violin bow detached from the strings of the violin.

A custom script written in C++ computed the calculations and stored the results in a CSV file.

Time

The sensor signal and the violin audio recording were imported into Sonic Visualiser to evaluate participants' time response to the notation stimuli. Figure 3.6 shows a region of a sensor signal and the related violin recording generated during a study session with a participant.

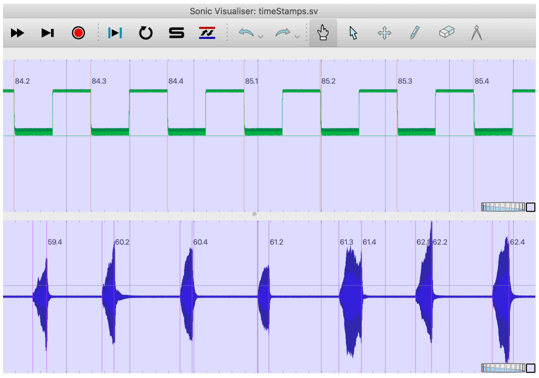


Figure 3.6: a portion of the violin recording in blue and of the sensor signal in green.

I added a time marker each time the sensor signal jumped to 0, meaning a new stimulus was shown on the screen. Then I added a time marker at the beginning of each note attack. Finally, I exported the audio and sensor signals timestamps as CSV files.

Resulting database

The resulting CSV files containing pitch and time information were combined in R Studio to generate a database that includes each note:

- stimulus timestamp
- stimulus frequency
- stimulus performance start timestamp
- performed start frequency
- performed mean frequency
- performed end frequency
- study section
- participant numeric ID

Pitch deviation calculation

Pitch deviation at each note's start, mean and end was calculated as follows. First, the expected frequency for each stimulus $f_e(n)$ was computed as:

$$f_e(n) = m_0 2^{d_n/12}$$

where m_0 is the frequency of a reference note[7], and d_n is the difference in semitones between the stimulus MIDI note number and the reference MIDI note number.

I then calculated the error in cents for the start aS_n, mean aM_n, and end aE_n notes's frequencies as follows:

$$aS_n = abs(1200 * log_2(mS_n/fE_n))$$
$$aM_n = abs(1200 * log_2(mM_n/fE_n))$$
$$aE_n = abs(1200 * log_2(mE_n/fE_n))$$

The following analysis partially focuses on how much performers deviated from the stimuli rather than how much they played above or below the expected frequency. Thus I took the absolute value rather than the raw value of each result. Values produced by the calculations were added as new columns in the data frame.

Time response

Time response for semitones and quarter tones was computed by subtracting the start time of a performed note from the timestamp identifying the moment the stimulus was displayed.

[7] I took the lowest note playable on the violin, which is G3, MIDI note number 55, 196 Hz

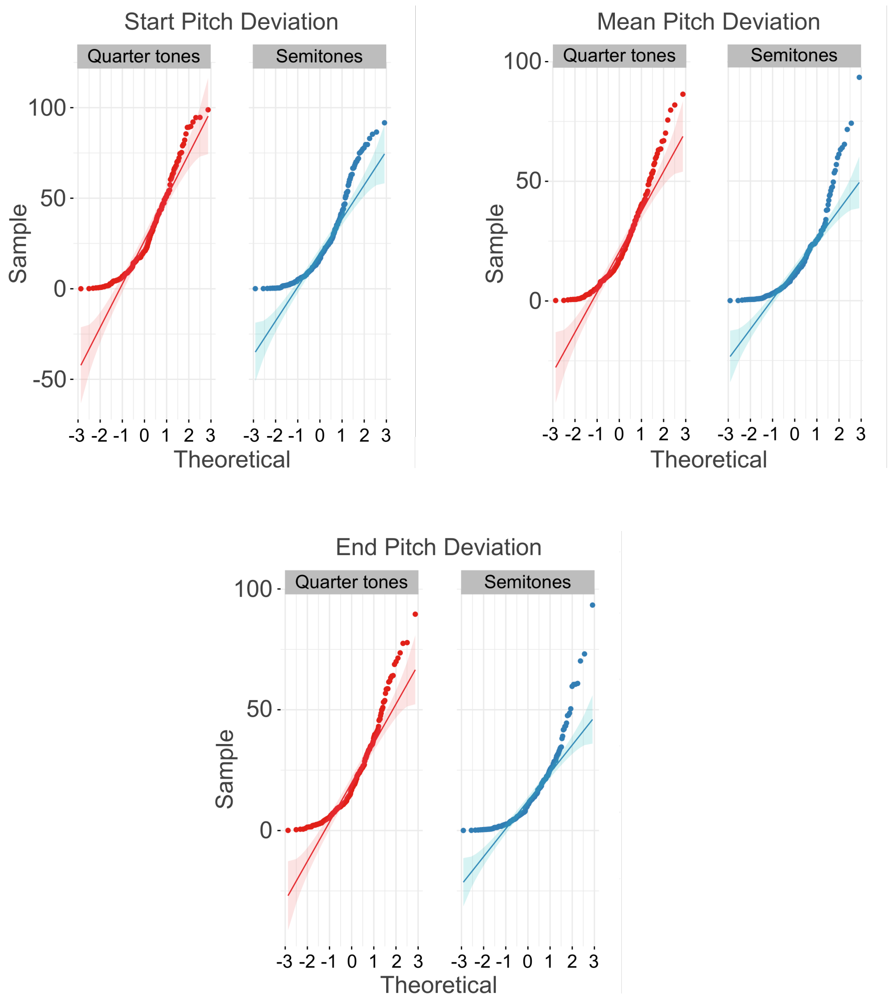
Statistical tests

Tables in the following sections report the semitones data group named “S” and the quarter tones data group “Q”.

Preliminary evaluation of the data showed a non-normal distribution in each data sample considered for the data analysis. The same type of distribution for the pitch deviation data was found in the individual sections of the study.

Pitch deviations for the semitones and the quarter tones groups across the study are shown in Figure 3.7. Several data points do not fall along each group’s (45-degree) reference line. So we can assume the non-normality of the data.

Figure 3.7: Quantile plots about semitones and quarter tones start, mean, and end pitch deviation data.



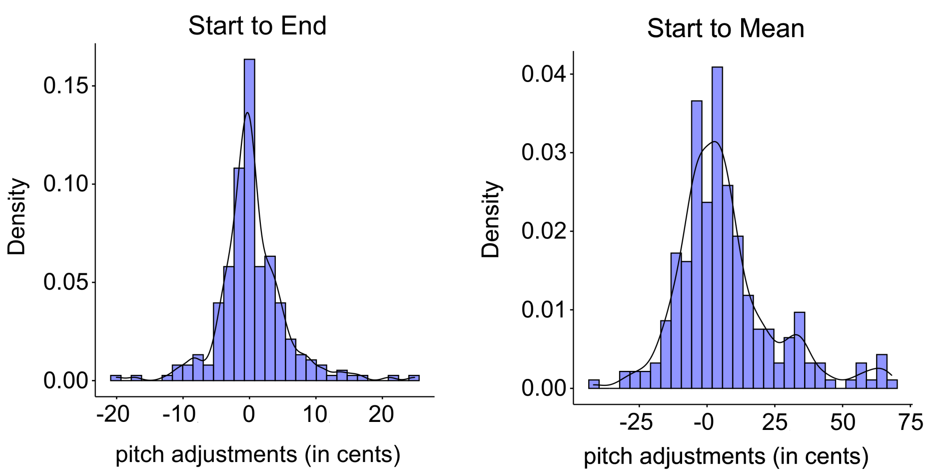
The non-normal distribution of the data led to the choice of a non-parametric statistical test to evaluate the comparisons between unpaired and paired data.

The comparison between unpaired data was evaluated using the Wilcoxon rank sum test. The test is used for comparing two independent groups of samples in a situation where the data are not normally distributed. An example of such a comparison is *pitch start deviation* for semitone versus *pitch start deviation* for quarter tones.

Further tests on paired data showed that their difference is approximately distributed symmetrically around the median. As the paired samples considered for the analysis are not normally distributed, and their difference is distributed symmetrically around the median, their comparison was assessed with Wilcoxon signed rank test on paired samples. The test is useful for comparing two paired groups of samples when the data are not normally distributed. The test assumes that differences between paired samples should be distributed symmetrically around the median.

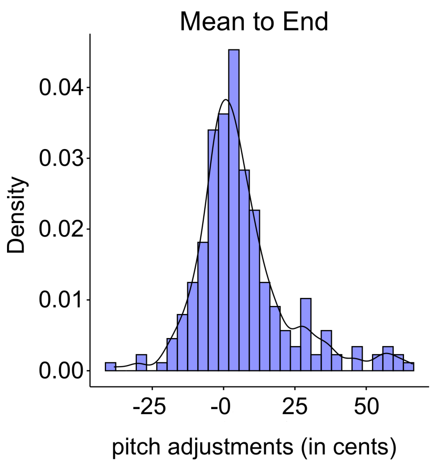
The histogram plots in Figure 3.8 show that the pitch adjustment data (in cents) for quarter tones are approximately symmetrical for quarter tones performances. The same type of distribution was found by looking at semitones across the study and at semitones and quarter tones in each section of the study.

Figure 3.8: Data distribution around the median for pitch adjustments between the start-end, between the start-mean, and between the mean-end of quarter tones.



Statistical tests considered semitones and quarter tones performances in:

- the overall study data
- the audio section data
- the overall notation data (the sum of notation$_a$ plus notation$_b$ data)



- the comparison between audio and notation data
- $notation_a$ section data
- $notation_b$ section data
- a comparison between $notation_a$ and $notation_b$ data

Outliers

Over 576 stimuli were performed across the study:

- thirteen quarter tone stimuli and eleven semitones were performed with a *pitch start deviation* major of 100 cents or minor than -100 cents
- one stimulus was performed with a *mean pitch deviation* greater than hundred cents.

Thirteen of these performances (six semitones and seven quarter tones) were identified as extreme outliers. Outliers and their assessment were accomplished in R Studio using the function *is outlier*. As found in the R Studio help section of the software, values above the third quartile + 1.5xIQR[8] or below the lower quartile - 1.5xIQR are considered as outliers. Values above the third quartile + 3xIQR or below the first quartile - 3xIQR are considered extreme points (in other words, extreme outliers).

Participants were expert violinists and therefore had the ability to precisely perceive pitch and accurately position their fingers on the violin's fretboard. A quarter tone is half a tone which is fifty cents above or below a certain note. So it's necessarily located at some point in between two regular notes which in

[8]Interquartile Range

the equal temperament are one hundred cents distant from one another. Given the participant's skills, even if a player does not know where a quarter tone is located, it would be reasonable to expect them to adjust their tuning in between an interval of 100 cents. Any further deviation could be reasonably explained as a mistake or participant's fatigue.

The influence of the identified exFtreme outliers on the following statistics is trivial (2 cents on pitch deviation medians for pitch start deviations). However, as they probably qualify as mistakes, these performances were not included in the analysed data presented in this document.

Therefore, the following sections describe data analysis over 563 performances, of which 281 were in response to audio stimuli and 282 were in response to notation stimuli. Performances considered in the $\text{Notation}_{\text{a}}$ section are 140, and performances considered in the $\text{Notation}_{\text{b}}$ section are 142.

Pitch deviation results are expressed in cents, while time response results are expressed in seconds.

3.4.2 Pitch analysis results

Overall stimuli performances

Quarter tones had a higher pitch deviation than semitones at the start, mean and end of players' performances. Pitch deviation comparisons between players' execution of quarter tones and semitones were statistically significant with $p < 0.0004$ and small effect size values (between 0.13 and 0.24). Data shows that IQR values are higher for quarter tones compared to semitones. Figure 3.9 shows a visual representation of the data. Full statistics and results are reported in tables 3.1, 3.2.

Table 3.1: pitch deviation statistics for quarter tones and semitones performances across the study.

Note type	Variable	Median	Interquartile range
Q	start	20.786	32.306
Q	mean	17.309	22.672
Q	end	16.841	21.959
S	start	16.579	25.308
S	mean	10.709	16.825
S	end	10.720	15.580

Quarter tones had a higher pitch deviation at the start of the performances compared to the mean and end of quarter tones executions. The median difference between the quarter tones group's start and end pitch deviation is 6.7

Figure 3.9: Quarter tones pitch deviation versus semitones pitch deviation at the beginning, mean and end of executions across the study

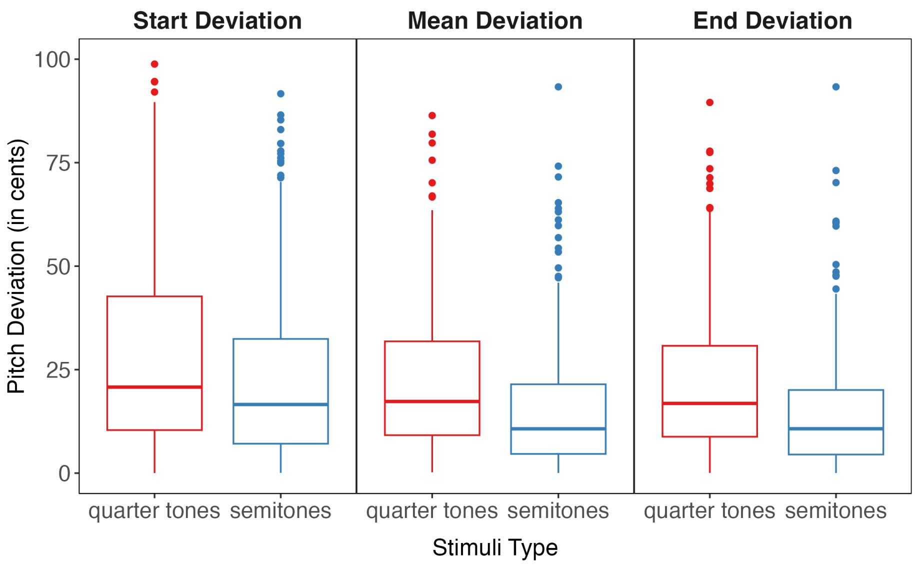


Table 3.2: Quarter tones and semitones pitch deviation comparisons across the study.

Var	Group1	Group2	p	p signif.	effect size	magnitude
start	Q	S	4.03e-03	**	0.13	small
mean	Q	S	1.70e-07	****	0.23	small
end	Q	S	2.23e-08	****	0.24	small

cents. Pitch deviation comparisons between the start and the end of the quarter tones and between the start and the mean of the quarter tones are significant with $p < 0.0001$ and a moderate effect size of 0.31 and 0.34. A further comparison between the mean and the end pitch deviation for the quarter tones group was statistically non-significant (see table 3.4 for details).

Semitones also had a higher pitch deviation at the start of the performances compared to the mean and end of players' executions. The median difference between the semitones group's start and end pitch deviation is 8.9 cents. Pitch deviation comparisons between the start and the end of the quarter tones and between the start and the mean of the quarter tones are significant with $p < 0.0001$ and a large effect size of respectively 0.51 for the start-to-end comparison and 0.53 for the start-to-mean comparison. A further comparison between the mean and the end pitch deviation for the semitones group proved to have small statistical significance (see table 3.4 for details).

IQR values appear to be higher for both groups at the start of the notes and to get lower considering the mean and the end of the stimuli execution. Table 3.3 presents a summary of pitch adjustment data.

Table 3.3: Pitch adjustment statistics for quarter tones and semitones performances across the study.

Note type	Group	Median	Interquartile range
Q	start to end	6.749	18.294
Q	start to mean	6.363	16.378
Q	mean to end	0.386	5.143
S	start to end	8.907	16.576
S	start to mean	8.130	15.142
S	mean to end	0.776	4.366

Table 3.4: pitch adjustment comparisons for quarter tones and semitones performances across the study.

Note type	Group1	Group2	p	p signif.	effect size	magnitude
Q	start	end	1.37e-06	****	0.31	moderate
Q	start	mean	6.88e-08	****	0.34	moderate
Q	mean	end	6.57e-01	ns	0.03	small
S	start	end	6.75e-18	****	0.51	large
S	start	mean	3.91e-19	****	0.53	large
S	mean	end	1.00e-02	**	0.15	small

Audio stimuli performances

Results from the study's audio condition align with the overall results. Quarter tones had a higher pitch deviation than semitones at the start, mean and end of players' performances. Players performed quarter tones stimuli with a pitch deviation of 25.87, 14.47, 14.5 cents, whereas semitones' pitch deviation was equal to 18.89, 9.68, and 8.6 cents. Pitch deviation comparison between quarter tones and semitones at the beginning of their performances were statistically significant, with $p < 0.013$ at the start of the executions and $p < 0.0001$ at the mean and end of players' performances. Results have small effect size values. Figure 3.10 shows a visual representation of the data, and tables 3.5 and 3.6 present a summary of these results.

Figure 3.10: Quarter tones pitch deviation versus semitones pitch deviation at the beginning, mean and end of executions in the audio condition

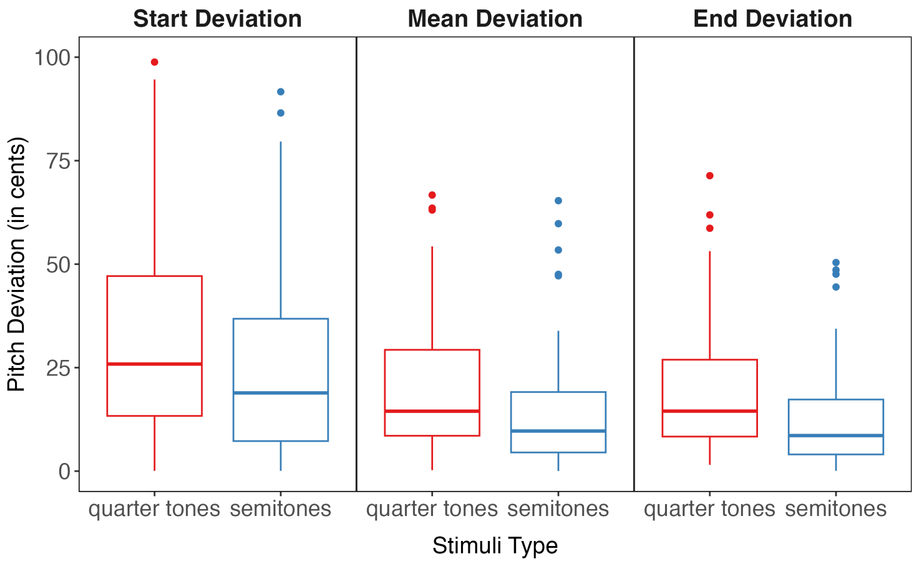


Table 3.5: pitch deviation statistics for quarter tones and semitones performances in the audio section.

Note type	Variable	Median	Interquartile range
Q	start	25.867	33.780
Q	mean	14.474	20.761
Q	end	14.495	18.572
S	start	18.896	29.552
S	mean	9.687	14.599
S	end	8.590	13.267

Table 3.6: Quarter tones and semitones pitch deviation comparisons in the audio section.

Var	Group1	Group2	p	p signif.	effect size	magnitude
start	Q	S	1.32e-02	*	0.16	small
mean	Q	S	2.36e-04	***	0.23	small
end	Q	S	1.03e-05	****	0.28	small

The median difference between the start and the end pitch deviation for the quarter tones group is 12.3 cents, whereas the median difference between the start and the end pitch deviation for the semitones group is 13.35 cents. The two comparisons are highly significant, with $p < 0.0001$ and large effect size values. Tables 3.7 and 3.8 present a summary of the results.

Table 3.7: pitch adjustment statistics for quarter tones and semitones performances in the audio section.

Note type	Group	Median	Interquartile range
Q	start to end	12.947	20.692
Q	start to mean	12.381	18.597
Q	mean to end	0.566	5.547
S	start to end	13.351	19.353
S	start to mean	12.214	17.826
S	mean to end	1.136	5.023

Table 3.8: pitch adjustment comparisons for quarter tones and semitones performances in the audio section.

Note type	Group1	Group2	p	p signif.	effect size	magnitude
Q	start	end	7.92e-09	****	0.54	large
Q	start	mean	1.02e-10	****	0.60	large
Q	mean	end	5.09e-01	ns	0.06	small
S	start	end	3.16e-14	****	0.65	large
S	start	mean	7.59e-15	****	0.66	large
S	mean	end	2.00e-02	*	0.20	small

Notation stimuli performances

Analysis outcomes from the notation stimuli performances align with the overall results. Quarter tones had a higher pitch deviation than semitones at the start, mean and end of players' performances. Players performed quarter tones stimuli with a pitch deviation of 19.34, 19.53, and 19.06 cents, whereas semitones' pitch deviation was equal to 14.67, 11.84, and 12.24 cents. The pitch deviation comparison between quarter tones and semitones at the beginning of their performances was not statistically significant, with $p = 0.1$. However, further comparisons between quarter tones and semitones' mean pitch deviation and between quarter tones and semitones' end pitch deviation resulted in being significant with $p < 0.0003$. Results have small effect size values. Figure 3.11 shows a visual representation of the data, and tables 3.5 and 3.6 provides further insights about these results.

Figure 3.11: Quarter tones pitch deviation versus semitones pitch deviation at the beginning, mean and end of executions in the Notation_a plus Notation_b condition

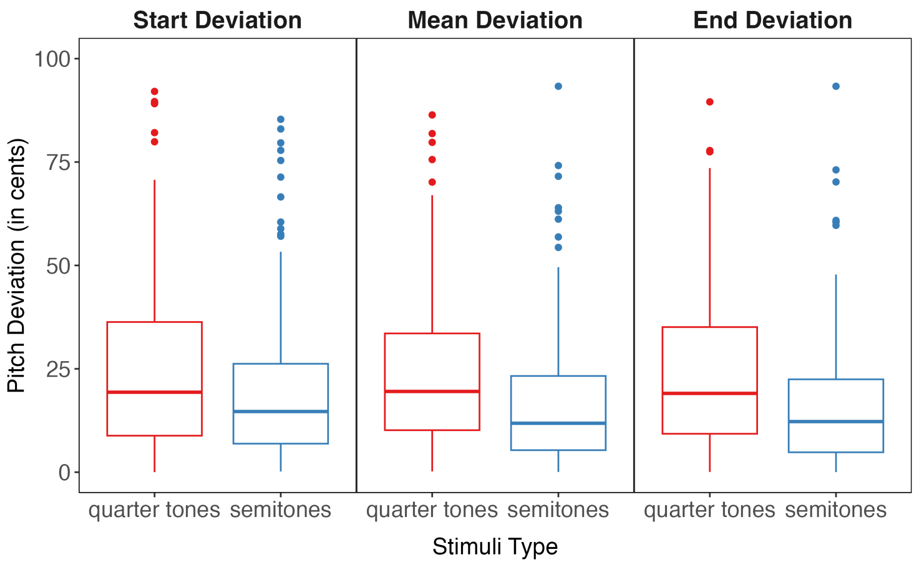


Table 3.9: pitch deviation statistics for quarter tones and semitones performances in the Notation_a section plus Notation_b section.

Note type	Variable	Median	Interquartile range
Q	start	19.343	27.486
Q	mean	19.526	23.396
Q	end	19.058	25.796
S	start	14.672	19.318
S	mean	11.836	17.941
S	end	12.242	17.643

Pitch deviation for the quarter tones group at the beginning, mean, and end

Table 3.10: Quarter tones and semitones pitch deviation comparisons in the Notation_{a} section plus Notation_{b} section.

Var	Group1	Group2	p	p signif.	effect size	magnitude
start	Q	S	0.105000	ns	0.10	small
mean	Q	S	0.000224	***	0.22	small
end	Q	S	0.000338	***	0.22	small

of the executions exhibits close values for the notation stimuli performances. The median difference between the start and the end and the start and the mean pitch deviation for the quarter tones group is respectively 1.2 cents and 0.95 cents. The median difference between the mean and the end pitch deviation for the quarter tones group is 0.22 cents. Pitch deviation comparisons between the start and the end, the start and the mean and the mean and the end pitch deviation are non-significant, with p comprised between 0.4 and 0.9.

Semitones had a higher pitch deviation at the start of the performances compared to the end of players' executions. The median difference between the semitones group's start and end pitch deviation is 4.65 cents. Pitch deviation comparisons between the start and the end and between the start and the mean of performances in the semitones group are significant, with a p-value < 0.0001 and a moderate effect size of respectively 0.36 for the start-to-end comparison and 0.38 for the start-to-mean comparison.

Comparisons between the mean and the end pitch deviation for the semitones group proved to be non-significant, with a p-value $= 0.19$.

Table 3.11 presents a summary of pitch adjustment data, whereas table 3.12 presents a summary of the statistical significance of the comparisons.

Table 3.11: pitch adjustment statistics for quarter tones and semitones performances in the Notation_{a} section plus Notation_{b} section.

Note	Group	Median	Interquartile range
Q	start to end	1.175	13.680
Q	start to mean	0.951	11.759
Q	mean to end	0.224	4.766
S	start to end	4.648	11.992
S	start to mean	4.217	10.706
S	mean to end	0.432	3.609

Table 3.12: pitch adjustment comparisons for quarter tones and semitones performances in the Notation_a section plus Notation_b section.

Note type	Group1	Group2	p	p signif.	effect size	magnitude
Q	start	end	0.426	ns	0.07	small
Q	start	mean	0.466	ns	0.06	small
Q	mean	end	0.935	ns	0.01	small
S	start	end	1.47e-05	****	0.36	moderate
S	start	mean	4.31e-06	****	0.38	moderate
S	mean	end	1.87e-01	ns	0.11	small

Audio stimuli performances versus notation stimuli performances

Overall, pitch adjustment for the quarter tones group is higher in the audio section compared to the notation section. The adjustment happens mostly between the start and the mean of the stimuli executions in both the audio and notation sections. *Between the start and the end of the notes*, the median pitch adjustment for the quarter tones group was higher in the audio section compared to the notation section (12.95 vs 1.82 cents). The comparison is significant with $p < 0.0001$ and small effect size($r = 0.28$). *Between the start and the mean of the notes*, quarter tones also had a higher median pitch adjustment in the audio section (12.381 vs 1.79 cents). The comparison is significant with $p < 0.0001$ and a moderate effect size ($r = 0.34$) *between the mean and the end of the notes.* A further comparison between the median pitch adjustment for the quarter tones group in the audio and notation section resulted in a non-significant p-value = 0.64.

Data describing the semitones group also shows that pitch adjustment is higher in the audio section compared to the notation section. The adjustment for the semitones group in the notation section is higher than that for the quarter tones group (the significance of this comparison is reported in the previous section). *Between the start and the end of the notes*, the median pitch adjustment for the semitones group in the audio section was 13.35 cents, whereas the median for the semitones group in the notation section was 6.13 cents. The comparison is significant with $p < 0.0001$, effect size $r = 0.23$. *Between the start and the mean of the notes*, the median pitch adjustment for the semitones group in the audio section was 12.214 cents, whereas the median for the semitones group in the notation section was 5.889 cents. The comparison is significant with $p < 0.0001$ and a small effect size $r = 0.24$ *between the mean and the end of the notes*, the median pitch adjustment for the semitones group in the audio section was 1.136 cents, whereas the median for the semitones group in the notation section was 0.242 cents. The comparison is non-significant, with $p = 0.34$.

Figure 3.12 shows comparisons between audio and notation sections for quarter tones, and Figure 3.13 shows comparisons between audio and notation sections for semitones. Table 3.13 presents a summary of the statistical significance of the comparisons.

Figure 3.12: Quarter tones pitch deviation in the audio condition versus notation$_a$ plus notation$_b$ condition at the beginning, mean and end of players' executions

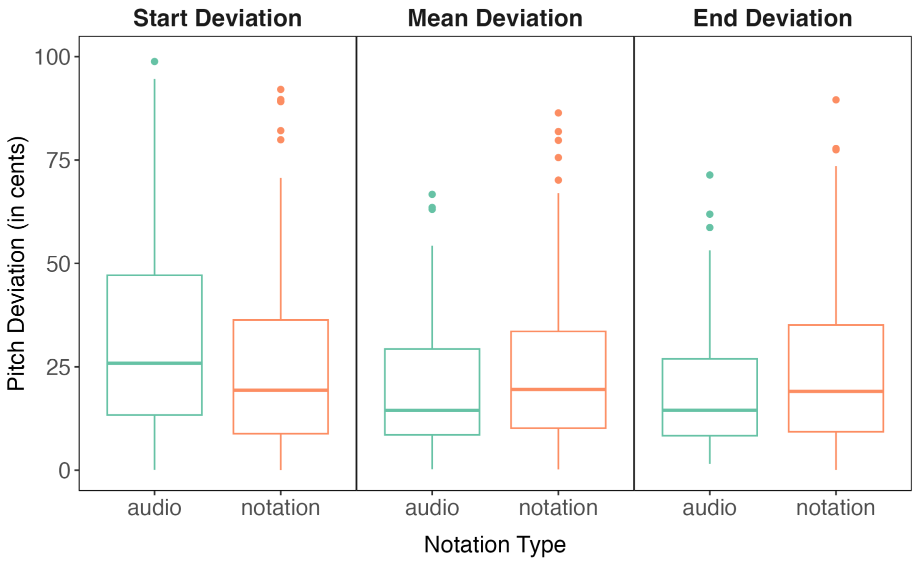


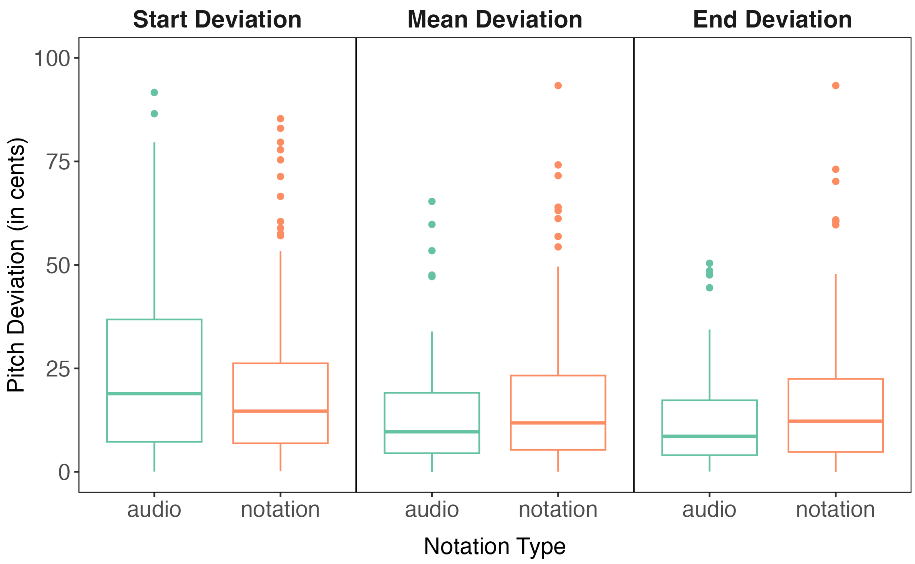
Table 3.13: pitch adjustment comparisons for quarter tones and semitones performances in the audio section versus Notation$_a$ plus Notation$_b$ section.

Note type	Variable	p	p signif.	effect size	magnitude
Q	start to end	1.07e-05	****	0.28	small
Q	start to mean	1.17e-07	****	0.34	moderate
Q	mean to end	6.45e-01	ns	0.03	small
S	start to end	1.32e-04	***	0.23	small
S	start to mean	4.09e-05	****	0.24	small
S	mean to end	3.43e-01	ns	0.06	small

Notation$_a$ performances

Quarter tones had a higher pitch deviation than semitones at the start, mean and end of players' performances. Pitch deviation comparisons between quarter tones and semitones at the start of players' performances were not statistically significant, with p = 0.16. Comparisons between the semitones and quarter

Figure 3.13: Semitones pitch deviation in the audio condition versus notation$_a$ plus notation$_b$ condition at the beginning, mean and end of players' executions



tones groups' pitch deviation at the mean and end of the performances had a small significance with p 0.027 and 0.04 and small effect size values (0.19, 0.17). Plots in Figure 3.14 and tables 3.14, 3.15 present a summary of these results and provide further details.

Table 3.14: pitch deviation statistics for quarter tones and semitones performances in the Notation$_a$ section.

Note type	Variable	Median	Interquartile range
Q	start	20.958	26.850
Q	mean	21.639	26.577
Q	end	22.776	27.918
S	start	16.495	20.743
S	mean	14.870	21.014
S	end	15.109	22.418

The median difference between the start and the end pitch deviation for the quarter tones group is 0.6 cents, whereas the median difference between the start and the end pitch deviation for the semitones group is 3.13 cents. The first comparison is non-significant, with p = 0.8. The second comparison is significant with p = 0.008 and a moderate effect size = 0.31.

Table 3.16 presents a summary of pitch adjustment data, and table 3.17 presents a summary of the significance of the comparisons.

Figure 3.14: Quarter tones pitch deviation versus semitones pitch deviation at the beginning, mean and end of executions in the Notation$_a$ section.

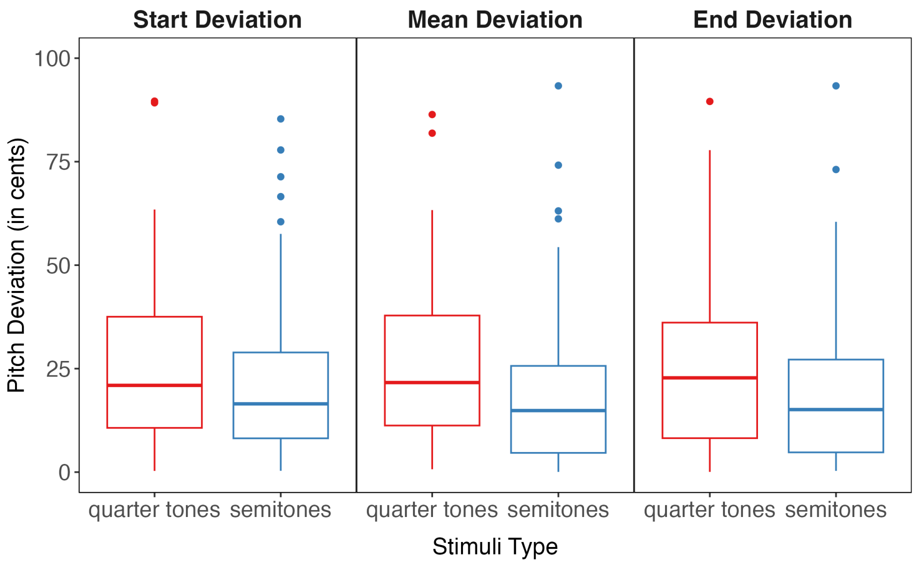


Table 3.15: Quarter tones and semitones pitch deviation comparisons in the Notation$_a$ section.

Var	Group1	Group2	p	p signif.	effect size	magnitude
start	Q	S	0.1650	ns	0.12	small
mean	Q	S	0.0267	*	0.19	small
end	Q	S	0.0414	*	0.17	small

Table 3.16: pitch adjustment statistics for quarter tones and semitones performances in the Notation$_a$ section.

Note type	Group	Median	Interquartile range
Q	start to end	0.563	13.014
Q	start to mean	0.146	11.242
Q	mean to end	0.416	4.432
S	start to end	3.124	9.810
S	start to mean	2.497	8.765
S	mean to end	0.627	3.261

Table 3.17: pitch adjustment comparisons for quarter tones and semitones performances in the Notation_a section.

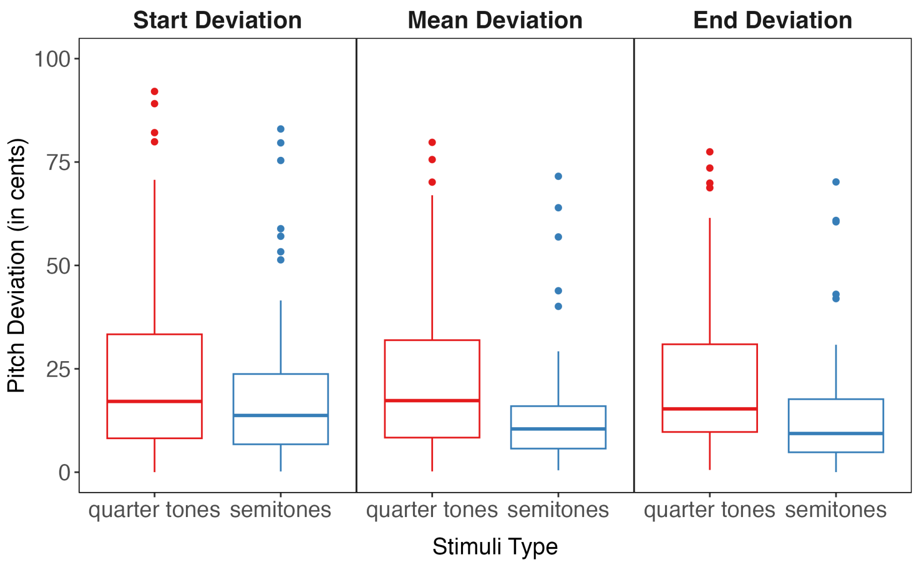
Note type	Group1	Group2	p	p signif.	effect size	magnitude
Q	start	end	0.769	ns	0.04	small
Q	start	mean	0.898	ns	0.02	small
Q	mean	end	0.606	ns	0.06	small
S	start	end	0.008	*	0.31	moderate
S	start	mean	0.014	*	0.29	small
S	mean	end	0.057	ns	0.23	small

Notation_b performances

Notation_b section presents similar results compared to Notation_a section.

Quarter tones had a higher pitch deviation than semitones at the start, mean and end of players' performances. Pitch deviation comparisons between quarter tones and semitones at the start of players' performances were not statistically significant, with p = 0.43. Comparisons between the semitones and quarter tones groups' pitch deviation at the mean and end of the performances had a small significance with p 0.005 and 0.002 and small effect size values of 0.24 and 0.27. Plots in Figures 3.15 and tables 3.18, 3.19 present a summary of these results.

Figure 3.15: Quarter tones pitch deviation versus semitones pitch deviation at the beginning, mean, and end of executions in the Notation_b section.



The median difference between the start and the end pitch deviation for the

Table 3.18: pitch deviation statistics for quarter tones and semitones performances in the Notation_B section.

Note type	Variable	Median	Interquartile range
Q	start	17.117	25.146
Q	mean	17.309	23.558
Q	end	15.311	21.205
S	start	13.722	16.994
S	mean	10.465	10.267
S	end	9.359	12.843

Table 3.19: Quarter tones and semitones pitch deviation comparisons in the Notation_b section.

Variable	Group1	Group2	p	p signif.	effect size	magnitude
start	Q	S	0.43000	ns	0.07	small
mean	Q	S	0.00450	**	0.24	small
end	Q	S	0.00183	**	0.27	small

quarter tones group is 1.82 cents, whereas the median difference between the start and the end pitch deviation for the semitones group is 6.13 cents. The first comparison is non-significant, with p = 0.8. The second comparison is significant with p = 0.008 and a moderate effect size = 0.31.

Table 3.20 presents a summary of pitch adjustment data, and table 3.21 presents a summary of the significance of the comparisons.

Table 3.20: pitch adjustment statistics for quarter tones and semitones performances in the Notation_b section.

Note type	Group	Median	Interquartile range
Q	start to end	1.816	14.423
Q	start to mean	1.793	12.311
Q	mean to end	0.023	5.121
S	start to end	6.131	13.694
S	start to mean	5.889	12.132
S	mean to end	0.242	3.932

$\textbf{Notation}_\textbf{a}$ performances and $\textbf{notation}_\textbf{b}$ performances

In the previous sections of this document, data show that comparisons between semitones and quarter tones in notation_a and notation_b sections are mostly non-

Table 3.21: pitch adjustment comparisons for quarter tones and semitones performances in the Notation_b section.

Note type	Group1	Group2	p	p signif.	effect size	magnitude
Q	start	end	0.433	ns	0.10	small
Q	start	mean	0.254	ns	0.14	small
Q	mean	end	0.586	ns	0.07	small
S	start	end	4.52e-04	***	0.41	moderate
S	start	mean	5.46e-05	***	0.47	moderate
S	mean	end	8.86e-01	ns	0.02	small

statistically significant. Equally, comparisons between notation_a and notation_b are not significant as shown in table 3.22. Non-significance within and between these two study sections may depend on the small size of the sub-datasets (142 and 140 performances over 563 executions across the study).

Table 3.22: pitch adjustment comparisons for quarter tones and semitones performances in the Notation_a section versus Notation_b section.

Note type	Variable	p	p signif.	effect size	magnitude
Q	start to end	0.715	ns	0.03	small
Q	start to mean	0.357	ns	0.08	small
Q	mean to end	0.416	ns	0.07	small
S	start to end	0.399	ns	0.07	small
S	start to mean	0.182	ns	0.11	small
S	mean to end	0.326	ns	0.08	small

3.4.3 Semitones versus quarter tones time response

Semitones performances had a quicker response than quarter tones when the stimuli were presented as notation.

Across the two notation sections, the median *time deviation* in the quarter tones group was 1.8 seconds, whereas the median in the semitones group was 1.42 seconds. The Wilcoxon test showed a significant difference ($p < 0.0001$, effect size $r = 0.407$). In the notation_a section, the median *time deviation* in the quarter tones group was 1.86, whereas the median in the semitones group was 1.56 seconds. The Wilcoxon test showed a significant difference ($p < 0.0001$, effect size $r = 0.40$). In the notation_b section, the median *time deviation* in the quarter tones group was 1.657 seconds, whereas the median in the semitones group was 1.33 seconds. The Wilcoxon test showed a significant difference ($p < 0.0001$, effect size $r = 0.491$). A summary of these statistics is displayed in table 3.23.

Table 3.23: median time response statistics for quarter tones and semitones performances in the Notation_a plus Notation_b sections, in the Notation_a section, and in the Notation_b section

.

Note type	StudySection	Median	Interquartile range
Q	$\text{Notation}_\text{ab}$	1.802	0.622
S	$\text{Notation}_\text{ab}$	1.424	0.525
Q	Notation_a	1.860	0.654
S	Notation_a	1.556	0.572
Q	Notation_b	1.657	0.585
S	Notation_b	1.331	0.309

3.5 Discussion

This study examined musicians' pitch accuracy when they played semitones and quarter tones. It was expected that the performers would be more proficient with semitones, given that they are familiar and common in Western music. Hence, the accurate rendition of semitones compared to quarter tones isn't surprising, but it served as a benchmark for comparison, allowing to contrast musicians' performances when handling familiar elements (semitones) against those when encountering less familiar ones (quarter tones) under differing conditions (audio and notation).

A distinctive finding of this study is that even expert players, possessing the necessary motor and kinaesthetic skills, are unable to accurately perform unfamiliar sounds if they lack an internal representation of the sonic result. In other words, their accuracy in execution is highly compromised in the absence of pertinent auditory imagery. This phenomenon is observable even if the instrument's interface is familiar to the players and its auditory feedback is available. This finding is based on substantial evidence.

Before conducting the study, it was hypothesized that participants would manage to adjust the intonation of quarter tones with greater accuracy in the audio section because it allowed players to hear quarter tones as auditory stimuli before reproducing them. This, in turn, could enable them to make comparisons and use their motor and kinaesthetic skills to adjust their performance accordingly. If this hypothesis held true, minimal adjustments would be expected in the notation condition due to the absence of auditory references for quarter tones.

By looking at the data, this hypothesis is confirmed with a significance level of $p < 0.0001$. Pitch adjustments performed in the audio section provide compelling evidence that participants possess the necessary motor imagery and kinaesthetic understanding to adjust their position on the fretboard and enhance the pitch accuracy of quarter tones. As a result, they could approximate the fretboard position indicated by the notation and subsequently refine their pitch. However, in the notation condition, the notational stimuli presented appear to have provided little assistance to the players. The motor skills players employed to improve their pitch accuracy for quarter tones in the audio section seemed either ineffective or absent in the notation section.

I propose that, in the audio section of the study, auditory cues likely provided a surrogate for this absent auditory imagery, enabling players to establish an auditory benchmark. This benchmark, in turn, facilitated the comparison of their quarter-tone performance with the intended sound. Although this hypothesis necessitates further investigation, it nonetheless provides a plausible explanation for the observed discrepancy in pitch adjustment between the audio and notation sections. It hints that the capability to construct the desired pitch mentally might be of greater importance than the mental imagery of the physical movements associated with producing the pitch. In essence, sounds that aren't easily imagined may also be challenging to play accurately.

Furthermore, the data provide substantial evidence that if a musician does not have a familiar mental image of the sound intended to play (quarter tones in this case), it is not sufficient for the instrument to be familiar. The violinists played an instrument that they knew very well during the study. Nonetheless, they struggled with pitch accuracy and timely response in playing quarter tones compared to semitones both in the notation and audio sections. Comparisons between semitones and quarter tones were statistically significant, with p values < 0.001 or less across the study.

Finally, it is proposed that the auditory feedback from the instrument does not compensate for the absence of auditory imagery. Auditory imagery appears to be salient in determining what can be played on an instrument besides motor imagery and auditory feedback. I propose that quarter tones were not played as well as semitones as participants lack a coherent sonic imagination coupled with the needed motor programs.

3.6 Study Generalisation

The implications of the research, while primarily focused on violinists, may extend to the design modifications of various other musical instruments.

Considering different instruments such as trumpets, or guitars, the physical mechanism of playing a note may differ, but the importance of auditory imagery is likely consistent. Musicians must internally imagine the sound before it's produced, which can influence accuracy. Brass players, like violinists, may find it challenging to hit microtones that aren't typically used in Western music.

Auditory imagery's relevance probably stretches beyond playing quarter tones and could be pertinent to an array of musical tasks. This includes situations that demand playing an instrument or an instrument modification that affords to play new or unfamiliar sounds.

While these findings are derived from a specific study, they present a broader understanding of the role of auditory imagery in music performance. This understanding could influence the design modifications of a wide range of musical instruments. However, these generalizations should be approached with caution. Further research should be considered in each specific context to validate these inferences.

3.7 Implications for Instruments Design

In this chapter, we explored the challenging nature of playing quarter tones, even on an unmodified, traditional violin. Despite having access to precise mappings and auditory feedback, musicians encountered difficulties due to struggles within their auditory imagery. When the auditory imagery of a musician fails to identify a specific sound event, the corresponding motor program for performing it is likewise unavailable, leading to reduced accuracy in performance.

I propose that a lack of auditory imagery can constitute a mental limitation that should be considered in the design process of instruments' modifications that aim at skills retention. The design lens should therefore move from a technology-focused perspective to a more human-centred one to address this limitation.

This research proposes an increased level of attention that accounts for the sensorimotor link in music performance. This process is an essential element to consider in designing interfaces for producing new sounds. If players cannot imagine the sounds, they cannot play them. This consideration significantly differs from evaluating if an instrument's affordance to play microtones or ergonomically fits a player's hands.

Including the sensorimotor in the design process prompts a new design question: does an intrinsic link exist between the designed interface and the performers' existing imagination and techniques? If not, then the instrument may encounter a fundamental mental, rather than technological, limitation.

3.8 Conclusions

The results of the quartertones study challenge traditional feedback loop performance concepts, suggesting that this loop is not critical for accurate execution. Auditory feedback may still be necessary for correction and refinement.

To support the kind of performance that is guided by auditory imagination, there is a need to translate from the imagery of sound to an action that resonates with the instrument. If musicians lack this connection, they will lack the skills needed to perform with the instrument or its augmentation, thereby opening up room for further research. To what extent can existing auditory imagery be used to play unfamiliar sounds? For which musical aspects other than pitch would this approach be valid?

Design strategies that ensure accurate performance of unfamiliar sounds are necessary and could potentially be based on auditory imagery.

Chapter 4

Transposed Violin Experiment

The experiment described in this chapter has been produced from the paper *Design for Auditory Imagery: altering instruments to explore performer fluency* [16].

Violinists from the quarter tones study described in chapter 3 were asked to play on a transposed violin. This study investigates how modifying the relation between gesture and pitch challenges the auditory expectations of twelve violinists and, therefore, their auditory imagery. Using the re-tuned violin to play a note written on a music score results in the instrument producing a different, unfamiliar, note. The term unfamiliarity reflects the unexpected auditory feedback from the violin.

It is not proposed that instruments' re-tuning should be considered a peculiar form of instrument augmentation to query. The re-tuned violin simply represents a case where a modified instrument produces an unexpected sound because of its modification. Tasks involving the re-tuned violin refer to the research question *to what extent do performers' pitch accuracy and fluency deteriorate in the presence of unfamiliar auditory feedback and mismatched auditory imagery? Does their performance improve if the task allows participants to ignore the violin's sound and focus on their internal representation of sound?*

The hypothesis is that performers need to have coherent auditory imagery of the sound they need to perform. Altered or possibly irrelevant auditory feedback produced by a modified instrument may not significantly impede a player's performance. These results would agree with recent research on skilled sensorimotor control that highlights the value of auditory and motor imagery, which participate in feedforward anticipation in embodied musical performance.

The research questions for the transposed violin study and the quarter tones study are of primary importance for designing new instruments: new sounds that cannot easily be imagined and are not coupled to an existing motor program may not likely be easily performed on a musical instrument. When the goal is modifying an instrument and retaining professionals' skills, a designer may benefit from a design paradigm based on the notion of imagery guiding the playing process. This principle may be preferable to a design paradigm based on auditory feedback.

It should be clarified that this case study does not focus on finding the perfect tuning system for violins. Because of the violin re-tuning, imagining a particular note and playing it corresponded to the sound of a different note. Therefore, *when an instrument is modified in such a way that auditory-imagery and motor programs are decoupled, which circumstances allow faster/more accurate executions? In which conditions can participants retain their performing skills, such as fluency of execution and pitch accuracy on a modified instrument?*.

Participants declared they were unfamiliar with instrument re-tuning. Two of them declared having prior experience with playing an instrument tuned in fourth being self-taught guitar players. The individual performance metrics of these two participants indicate they did not outperform the other players involved in the study.

4.1 Study Description

A violin is usually tuned from the lower string to the upper string: G3, D4, A4, and E5 (concert tuning). The violin used for the study was tuned in fourths: A3, D4, G4, and C5 (transposed tuning). The instrument was re-tuned in advance at the university. Re-tuning participants' violins at the beginning of the study would not have guaranteed a reasonably stable tuning during the experiment. When a violin that is usually tuned in fifths is re-tuned, its pegs will tend to go back to their original configuration over time. It took two days for the new tuning to become similar to the violin used during the study. Hence, one re-tuned violin was prepared in advance and given to participants. See figure 4.1 for a visual reference describing the violin transposition. Participants were asked to play only in the first position[1]. Having participants playing in the first position, it was possible to observe players move between strings which maximised the disruption produced by the re-tuning. This potentially leads to data that describes more clearly the re-tuning effect.

Players were introduced to the study by explaining the violin re-tuning and

[1]Otherwise, players could have partially solved the re-tuning challenge by inferring subsequent pitches' position on the same string.

Figure 4.1: violin tuning in a normal setting (concert tuning) and during the study (transposed tuning).

presenting the tasks. They were asked to use the violin to play twenty-four short musical excerpts. The intent was to present short melodies familiar to the violinist's cultural space but that they would not have seen before. Music passages were selected among unknown compositions from the Western Baroque music repertoire. Being not familiar with the excerpts, players needed to resort to the stimuli presentation (notation, audio playback) rather than using their memory to play them. Notations included instructions on tempo. Each music passage was performed by participants twice. The second execution of each passage was considered for data analysis as it possibly represents the players' most accurate execution.

The study is divided into three different conditions:

- transposed notation
- concert notation
- audio playback.

In the transposed notation condition, participants performed music score excerpts re-written to match the violin transposition. This kind of notation was produced for this particular study. Figure 4.2 presents an example of a stimulus in this condition. Because the notation is transposed, it does not support participants in imagining the "right sounds" they should play or the overall musical excerpt. If participants would read, imagine, and play the sounds written on the notation, the results would differ from the music passage they need to execute. However, due to the alignment between the notation transposition and the re-tuning of the violin, this form of notation facilitates players in envisioning the appropriate gestures required to execute on the re-tuned violin's fretboard. Consequently, this alignment leads to the production of accurate pitches and, consequently, the faithful rendition of the musical excerpts. It is important to note that despite the discrepancy between the written notation and the auditory feedback produced by the violin, players are able to establish a meaningful connection between their actions and the intended musical outcome. To what extent can incoherent auditory feedback produced by a modified instrument impact players' fluency of execution and pitch accuracy? To what extent can it prevent the retention of such skills owned by participants? To what extent may participants stop listening to the auditory feedback from the violin, and if that happens partially, to what extent can they perform adjustments on their playing (like pitch adjustments) using the violin's sound? In other words, are they still listening to the sound produced by the modified instrument, and can they use it to improve their performance, similar to what they would usually do on an unmodified instrument?

Figure 4.2: a stimulus displayed in transposed notation.

In the concert notation condition, participants performed music score excerpts written as they were sourced in music books. It is the kind of notation that players normally use to rehearse the excerpts and play them in a concert. Figure 4.3 presents an example of a stimulus in this condition. *Concert notation allows participants to imagine the "right sounds" to play*, which are the pitches they read on the score, and they could hear performing the notation on a regular violin. Because they have to perform the notation on re-tuned violin, players need to re-imagine the connection between what they read plus their internal representation of the sound they are reading, and the gestures they need to perform the music. In other words, players need to think of a new mapping between their auditory imagery and the motor programs they use to play pitches on a violin. When they succeed, the sound coming from the violin (violin's auditory feedback) matches the music written on the score. To what extent auditory feedback that meets auditory expectations provided by the music scores supports players to retain their fluency and pitch accuracy on a modified instrument? In other words, does having indications of the "right sounds" to perform can be more important than having indications about the "right gestures" to perform (i.e. the spots on the re-tuned violin's fretboard where participants should place their finger)?

Figure 4.3: the stimulus shown in Fig 4.2 displayed in concert notation.

In the audio playback condition, musicians replicated music passages they listened to through speakers. This condition was designed to discriminate the effect of reading notation within the challenge of playing the musical excerpts on the transposed violin. However, during the study, participants struggled to identify and repeat music passages they had just listened to through speakers. Most of the passages were not performed entirely, and the execution duration was randomly shorter or longer according to the difficulty of the passage (which could be quantified according to features like the number of notes to remember and tempo in bpm). Therefore, the data gathered in this section is excluded from the analysis presented in the next sections.

Eight music passages were presented during each condition. An additional musical excerpt was shown at the beginning of each section and was regarded as a trial stimulus. Data produced during trial stimuli executions are not included in the analysis. Trial stimuli helped participants to familiarise themselves with the study's tasks. During the whole study, every music passage was presented only once to each participant to prevent potential learning effects and related biases in the data. Each time the performance of a stimulus was completed, I manually activated the following stimulus (more details in the apparatus section). The order of the sections and stimuli was randomised for each participant to prevent biases coming from a certain sequence of sections and/or stimuli. Participants were instructed to perform the stimuli as soon as they felt ready. Musicians were asked to avoid using vibrato in their executions[2].

The stimuli were manually transcribed from paper to digital files using the software MuseScore [90]. They were then exported as JPEG files to be displayed as a notation on a monitor and as uncompressed audio files synthesised (using MuseScore) to be played back through speakers. The virtual instrument chosen to convert the notation into audio was a piano.

4.1.1 Metrics

A potential outcome of the violin re-tuning is impaired fluency, a state where it is not possible to play something at tempo or with proper rhythm or intonation because it is necessary to pay conscious attention to each action. A lack of fluency may impact the stimuli's execution duration. Performances duration was considered a metric to evaluate how each condition affected players' performances. Performances may also manifest a lack of fluency when players repeat portions of the excerpts to correct performance mistakes. Participants' repetition of a portion of the stimuli would result in extra notes. The number of extra notes found in each section was considered as a further metric to evaluate how

[2]This study partially focuses on pitch analysis of players' performances. Executions characterised by vibrato would have biased the data and influenced analysis results since vibrato is a common playing technique influencing pitch

each condition impacted players' fluency. Pitch accuracy at the start and mean of the notes and pitch adjustment between the start and the means of the notes performances are also compared between sections to evaluate which condition mostly impacted players' ability to perform and correct pitch accurately.

4.1.2 Apparatus

A laptop was used to display the musical notation and audio stimuli. The Laptop ran custom Processing code to select hi-res images of the notation (or to start the playback of the stimuli), display them, and log the stimuli ID, performance ID, participants ID and execution timestamps.

The portion of the study apparatus to record performance data included a DPA microphone to capture the audio coming from the violin, and a Bela[67]. The Bela platform hosted a C++ script to record and store the violin audio signal. The visual configuration of the elements included in the apparatus is identical to the one described for the quarter tones study.

4.2 Data analysis

4.2.1 Building the database for analysis

The following paragraphs describe the extraction of the data from the violin recordings and the organisation of the data in a single data frame. The data published in this chapter is in line with the data and findings discussed in the publication by Guidi et al. [16]. Some of the figures related to pitch deviation may differ because of more accurate analysis. The previous analysis relied on the use of the software Tony [91] to automatically segment players' recordings for each passage into notes and provide frequency estimates. In this document, the pitch analysis is based on frequencies calculated following a more accurate procedure which is discussed in the quarter tone study (section 3.4.1).

The number of extra notes in the music passages differs as it was previously roughly calculated using an automatic music transcription algorithm. In this thesis, the number of performed notes (including possible extra notes) was based on manual annotation in the audio editor Sonic Visualiser.

Pitch

Pitch adjustment is evaluated as the difference between the start pitch deviation and the mean pitch deviation of each note executed by players[3].

[3] In the quarter tone study, we observed that much of the pitch adjustment for familiar auditory feedback (semitones) and unfamiliar auditory feedback (quarter tones) happened between the start and the mean of the notes. Therefore, this portion of players' performances

Pitch deviation data was extracted from the violin recordings using the procedures explained in section 3.4.1 with one exception. In the quarter tones study, the expected frequency was the stimulus frequency presented to participants. This guaranteed a one-to-one coupling of the stimulus with the performance. In the transposed violin study, players played extra notes compared to the notes notated on the score. The extra notes resulted in players repeating certain passages of their execution to achieve the excerpt. The pitch accuracy analysis includes these extra notes. In fact, the pitch analysis does not aim to assess the exact pitch correspondence between the notes on the score and their executions. Rather, it focuses on possible pitch adjustments in disrupting auditory feedback in the transposed notation condition versus the concert notation condition. Pitch deviation at the start and the mean of the notes performed by players were calculated using a custom R Studio script. The script takes a frequency in Hz and returns the closest chromatic pitch in a number of semitones above (positive) or below (negative) A4 (440Hz). The result is added to the reference MIDI note 69, and the resulting midi note is then converted into a frequency value which is used to calculate the expected frequency value. Finally, the start pitch deviation and the mean pitch deviation, as well as their difference, which we refer to as *pitch adjustment*), are computed and stored in two separate columns in a CSV file (see section 3.4.1 for more information on the formulae).

Performance duration ratio and extra notes

The execution duration ratio for each performance was calculated as the ratio between the performance duration and its expected duration. Each performance duration was retrieved by segmenting participants' violin recordings in Sonic Visualiser and exporting them as separate audio files. The length of each audio file in seconds represents the duration of the stimulus execution. Each performance was named using an ID identifying the execution, as well as the section within it, was performed and the participant who performed it. Each performance expected duration was retrieved by looking at stimuli MIDI transcriptions. Each music passage (including tempo information) was transcribed with the software MuseScore, synthesised with a piano plugin included in the software, and exported as an audio file. The length in seconds of each audio file represents the expected duration of the stimulus execution. Files were named using an ID identifying the related musical excerpt. A custom C++ script calculated the ratios between each performance duration and the expected duration. The script looked at the length of each performance's audio file, then it matched its

is seen as the subject of primary interest for assessing players' ability to retain and adjust their pitch accuracy in this study as well.

ID with the corresponding stimuli ID, and then retrieved the expected stimulus duration to calculate the ratio between performance and stimulus. The result was added in a new column to the study CSV dataset created at the end of the pitch calculations (see section 4.2.1).

The extra notes for each performance were calculated as the difference between the number of notes in the performance and the number of notes in the related stimulus. A negative number indicates that a participant did not perform all the notes that form the musical excerpt. A number equal to zero indicates that a participant played the same number of notes notated on the score. A value above zero indicates that a player performed additional notes compared to the notation (i.e. because they performed a bar twice). The number of notes in the performance was retrieved by looking at the number of start (or mean) frequency deviation values in the study dataset for that performance. The number of notes in the related stimulus was calculated by looking at the number of notes on messages in the stimulus MIDI file. A custom C++ script was used to retrieve the number of rows in the study dataset for each performance. The number of rows represents the number of notes performed. The script then matched the performance ID with the stimulus MIDI file ID and retrieved the number of expected notes for the musical excerpt. Then the script subtracted the two values to obtain the number of extra notes. The resulting values were added to the study database by creating a new column.

Resulting database

The resulting dataset contains the following columns:

- participant numeric ID
- section ID
- stimulus ID
- performance ID
- start frequency deviation
- mean frequency deviation
- duration ratio
- extra notes.

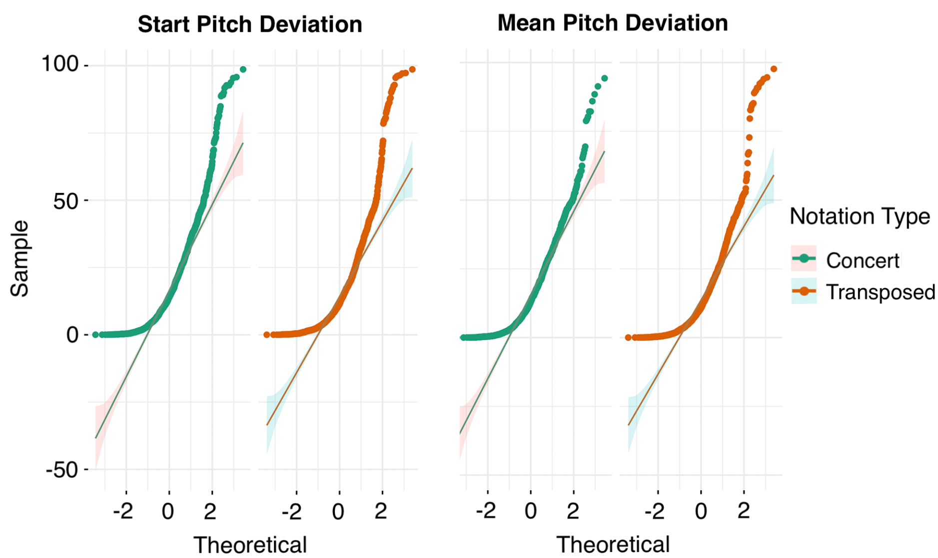
Statistical tests

In the following sections and figures, the concert notation group of data is occasionally named "C", and the transposed notation group of data "T".

Preliminary evaluation of the data showed a non-normal distribution in each data sample considered for the data analysis. The non-normal distribution of the data led to the choice of a non-parametric statistical test to evaluate the comparisons between unpaired and paired data. The comparison between unpaired data was evaluated using the Wilcoxon rank sum test. The test is used for comparing two independent groups of samples in a situation where the data are not normally distributed. An example of such a comparison is the pitch start deviation in transposed notation section versus the pitch starts deviation in the concert notation section.

Data distribution for pitch deviations (start, mean of the notes), as well as the duration ratio of the performances, are shown in Figures 4.4, and 4.5. In each graph, several data points do not fall along each group's (45-degree) reference line. So we can assume the non-normality of the data.

Figure 4.4: Concert and transposed notation start and mean pitch deviation data distribution.



Further tests on paired data showed that their difference is distributed symmetrically around the median. Therefore, comparisons between the start and mean pitch deviation within the concert notation section and the transposed notation section are evaluated using the Wilcoxon signed rank test on paired samples.

Figure 4.5: Concert and transposed notation executions duration data distribution.

The histogram plots in Figure 4.6 show the differences between the start and mean pitch deviation in the concert notation, and the differences between the start and mean pitch deviation in the transposed notation data are approximately symmetrical.

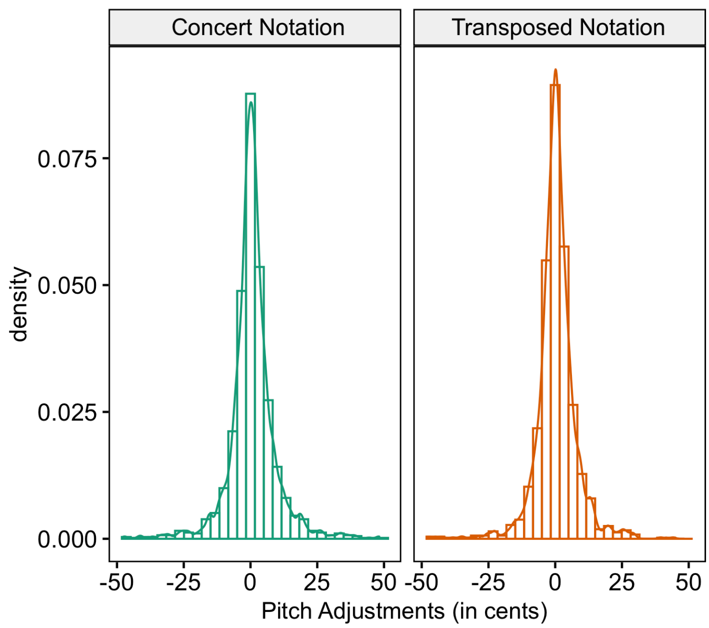
Statistical tests investigated:

- start pitch deviation in concert notation vs transposed notation
- mean pitch deviation in concert notation vs transposed notation
- start pitch deviation vs mean pitch deviation in concert notation
- start pitch deviation vs mean pitch deviation in transposed notation
- pitch adjustment in concert notation vs transposed notation.
- performance duration ratio in concert notation vs transposed notation
- number of extra notes performed in concert notation vs transposed notation

Data samples and outliers

Pitch data statistics are calculated over 1719 notes played in concert notation, and 1437 notes played in transposed notation. The pitch data presents no extreme outliers. Participants performed in the notation sections 192 musical

Figure 4.6: Data distribution around the median for pitch adjustment between the start and the mean of the excerpts execution in the concert and transposed notation conditions.



excerpts, which are split between 96 musical passages in the transposed notation section and 96 musical passages in the concert notation section. The performance duration data includes eight extreme outliers (two in the transposed notation group and six in the concert notation group), which do not impact the statistical results presented in the next sections. Outliers were assessed using the R function "identify outliers" presented in section 3.4.1.

4.2.2 Performance duration and extra notes results

Performances duration was closer to the duration of music passages in the transposed notation condition than in the concert notation condition with an average of 3.33 vs 5.41 seconds, $p < 0.0001$, and a moderate effect size of 0.4. Fewer extra notes characterised executions of transposed notation compared to concert notation. Specifically, 45 versus 653 notes with $p < 0.0001$, a moderate effect size of 0.3.

Table 4.1: performance duration statistics for C and T performances.

Note type	Median	Interquartile range
T	3.334	1.829
C	5.414	3.998

Table 4.2: C vs T perf duration and extra notes comparisons.

Var	p	p signif.	effect size	magnitude
performance duration	2.78e-16	****	0.4	moderate
extra notes	6.22e-10	****	0.3	moderate

4.2.3 Pitch deviation and pitch adjustment results

Notes performed in the concert notation condition had a higher pitch deviation than those performed in the transposed notation condition at the start and the mean of players' performances. Pitch deviation comparisons between players' execution of concert notation and transposed notation were statistically significant with $p < 0.001$ and small effect size values. Data shows that inter-quartile range values result are higher for concert notation compared to transposed notation. Figure 4.7 shows a visual representation of the data. Full statistics and results are reported in tables 4.3, 4.4.

Concert notation have a slightly higher pitch deviation at the start of the performances compared to the mean of concert notation executions. The median

Figure 4.7: Concert notation pitch deviation versus transposed notation absolute pitch deviation at the beginning, and mean of executions.

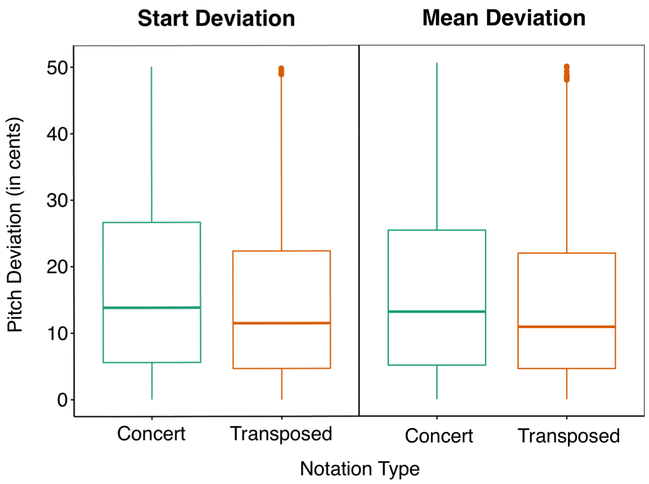


Table 4.3: pitch deviation statistics for C and T performances across the study.

Note type	Variable	Median	Interquartile range
C	start dev	13.833 cents	21.067 cents
C	mean dev	13.246 cents	20.408 cents
T	start dev	11.479 cents	17.669 cents
T	mean dev	10.955 cents	17.443 cents
C	dev adjustment	0.411	6.728
T	dev adjustment	0.320	6.338

Table 4.4: C vs T pitch deviation comparisons across the study.

Var	p	p signif.	effect size	magnitude
Start dev	7.16e-05	****	0.07	small
Mean dev	1.00e-03	***	0.06	small
Dev adjustment	2.96e-01	ns	0.02	small

difference between the start and the mean pitch deviation in the C group is 0.41 cents. Pitch deviation comparisons between the start and the mean of the notes performed in concert notation are significant with $p < 0.0001$ and a small effect size of 0.1 (see table 4.5 for details on the test significance).

Transposed notation have also a slightly higher pitch deviation at the start of the performances compared to the mean and end of players' executions. The median difference between the start and the end pitch deviation for the S group is 0.32 cents. Pitch deviation comparisons between the start and the mean of the transposed notation are significant with $p = 0.0096$ and a small effect size of 0.07 (see table 4.5 for details on the test significance).

Table 4.5: pitch adjustment comparisons for C and T performances (start to mean)

Note type	Var	p	p signif.	effect size	magnitude
T	deviation	9.63e-03	**	0.07	small
C	deviation	3.94e-05	****	0.10	small

4.3 Discussion

Results suggest that the transposed notation led to more fluent executions. Music passages had duration and a number of notes more coherent with the notation's indications than concert notation performances. It could be argued that participants tackled the transposed notation by using existing sight transposition skills. In *sight transposition*, a performer mentally transposes a note written on a score, then plays the note transposed. However, in this research, both the notation and the violin are transposed. The violin matches the notation transposition. No sight transposition is requested to perform the notation correctly. During the study, participants only needed to execute the gestures indicated in the notation. Participants reported to have isolated themselves from the violin's sound and focusing on the gestures to perform. Some of them reported having focused on their internal representation of the sound notated on the score and on the corresponding gesture to play it on the violin, trying to ignore the mismatched sonic result. These reports seem to contrast with the hypothesis that they managed to form newly learned auditory feedback quickly.

It is proposed that transposed notation worked so well because it let musicians use their familiar auditory imagery, which, in turn, is connected to the actions needed to perform the notation. Even if the resulting sound achieved by playing each note was not what they imagined it to be, the fluency of the performance improved. As a result, the transposed notation enables meaningful auditory feedback to obtain the desired melody. However, that is true only considering the overall result of the performance. Each note within a given performance has unexpected auditory feedback as both the violin and the notation are transposed. In this scenario, the imagery was deliberately incorrect; the notation did not produce the specified sounds. Nonetheless, it worked better to use a notation that *specified the correct motor actions* and that *produced* a disruptive auditory result (transposed notation) than to have a notation *specifying* an auditory feedback matching the expectations provided by the concert notation.

Finally, it is uncertain whether players totally ignored the auditory feedback from the violin in the transposed notation section. The amount of pitch adjustment performed by players was almost identical regardless of the type of auditory feedback available in the C and T conditions. This seems to suggest that despite producing disruptive auditory feedback, the transposed notation did not impact players' ability to adjust their pitch between the start and the mean of the notes. It could be that participants could still use some features of the disruptive auditory feedback to achieve the same degree of adjustment they had in the presence of coherent auditory feedback vs auditory imagery vs notation coupling (concert notation condition). However, it could equally be that the slight pitch corrections executed in both the C and T conditions were simply the result of kinaesthetic processes ingrained in participants' playing practice. Further research is needed to assess the cause of pitch corrective actions during players' performances in the T condition.

4.4 Study Generalisation

The specific nature of the context in which the study was conducted - working primarily with professional violin players and a retuned violin - might raise questions about the wider applicability of the findings. It is worth considering that this valid critique may not diminish the broader implications of the study. Often in scientific research, initial explorations start within specific, controlled parameters that then stimulate wider theoretical and practical implications.

What emerged from this study, was the unique adaptability of musicians, rooted not just in their sensorimotor skills but in their capacity to dissociate from actual auditory feedback, aligning more closely with their mental representation of sound. This demonstrates the depth and resilience of sensorimotor imagery. Such adaptability may not be exclusive to the professional violinists studied here.

More broadly, these findings suggest a potential hierarchy or prioritization in the cognitive processes of musicians. Even when auditory feedback contradicts expectations, the ingrained sensorimotor representations—shaped by years of practice and performance—guide the musicians towards fluency. The core revelation, then, is not the importance of sensorimotor imagery in isolation but its salience over real-time auditory feedback in guiding performance fluency.

This study's findings are an invitation to revisit our assumptions about how deeply musicians rely on their internal sonic worlds, even in the face of unfamiliar external stimuli. They present a broader understanding of the role of sensorimotor imagery in music performance. This understanding could influence the design modifications of a wide range of musical instruments. However, these generalizations should be approached with caution. Further research should be considered in each specific context to validate these inferences.

4.5 Implications for instrument design

In the quarter tones study described in chapter 3, we saw that playing quarter tones was challenging even with a traditional violin with no modifications. Correct mapping precision and auditory feedback did not support musicians in this condition. When the auditory imagery of a musician struggles to identify a particular sound event, then the motor program to perform it is also unavailable. The resulting accuracy of the performance diminishes.

In the transposed violin study, when the relation between the musical instrument and the auditory feedback is unfamiliar, but musicians can access their auditory imagery and its related motor programs, the accuracy of the performance (in terms of fluency) is better preserved (transposed notation). Additional examples that demonstrate these principles are prepared piano and MIDI keyboards. A piece for these instruments can still be notated as a piano piece, and players will not have trouble when playing it, even in the presence of unfamiliar auditory feedback. Possibly because they can imagine the traditional music space where the keyboard and the notation sit, and they can play the instrument regardless of the sound it produces.

In instrument design discussions, there is often a focus on the ability to produce any given sound using mapping strategies and technological solutions. I propose that a lack of auditory imagery constitutes a mental limitation which should be accounted for in the design process. The design lens should shift from a technology-focused view to include a more human-based perspective to address such a limitation. Taking a technocentric approach could only account for questions like mapping, precision, and degree of freedom. However, it would be insufficient to consider human-based aspects like the ergonomics of the interface and helpful feedback (audio, tactile, visual). In this research, it is proposed a different level of attention, which accounts for the sensorimotor link in music performance and the feedforward mechanism that describes it. This process is proposed as an essential element to account for in designing an interface that is meant to produce new sounds. Suppose the kind of performance aimed to enable is the one afforded by a traditional instrument (where musicians have an embodied relationship with it). In that case, it almost does not matter what the interface is. If players cannot imagine the sounds, they can not play them.

This approach poses a fundamental difference from asking if an instrument can afford to play microtones or if the instrument is ergonomically sensible to the hands of the player. Including the feedforward process in the design process leads to the following design question: is there an intrinsic link between a particular interface that is being designed and the existing imagination of performers that will play the interface, plus the execution techniques they already own? If the answer is no, then the instrument would run into a fundamental limitation which is mental rather than technological. If the goal of a designer is to let people play unfamiliar sounds, then this human factor needs to be taken into account.

4.6 Conclusions

The results of the study challenge the notion of playing as a feedback loop where performers think about what they want to do, play it, and evaluate the result based on the feedback to correct the performance. This performer study suggests that this feedback loop is not critical for the fluent execution of a performance. Feedback may nonetheless be needed for correction and refinement (i.e. the ends of each note in the first study were typically more in tune than the beginning, which we wouldn't expect to happen without feedback). Musicians can still play when the imagery is not entirely aligned with the auditory feedback (as in the transposed notation condition) as long as their anticipation leads them to the right motor program. Musicians can potentially substitute unfamiliar imagery if they can leverage a notation system connected to their existing sensorimotor imagery.

The study thus emphasizes the significance of human-based considerations alongside technocentric ones, ultimately encouraging a richer dialogue in the realm of musical instrument design and performance. To support the kind of performance that is guided by auditory imagination (i.e. execution with music scores), there is a need to translate from the imagery of sound to an action that is harmonious and appropriate to the instrument. If musicians do not have that connection (i.e. because the instrument is unfamiliar), they will also lack the skills needed to perform with the instrument or its augmentation. Establishing that connection is a goal that provides space for more research questions like *to what extent is it possible to use existing auditory imagery to play unfamiliar sounds? For which musical aspects other than pitch would this approach be valid?*

As shown with concert notation, while thinking and reasoning, the quality of musicians' performance deteriorated. Design strategies that assure that performance remains automatic are needed and could rely on auditory imagery.

Future studies could aim to replicate your experiments with different instruments, populations, and cultural contexts. This would help to confirm the conclusions drawn and their potential applicability to instrument design and performance in general.

Chapter 5

Augmented Instruments Evaluation

The material presented in this chapter, particularly the following experiment, has been produced from the paper *Quantitative Evaluation of Aspects of Embodiment in New Digital Musical Instruments* [92]. Secionts of the description of the augmented plectrum used in the study, which is provided in section 5.2.1, have been produced from the paper *Magpick: an Augmented Guitar Pick for Nuanced Control*[82].

5.1 Introduction

This chapter presents a quantitative method to evaluate whether an expert player can execute skilled actions on an unfamiliar interface while focusing their performance on the musical outcome rather than on the technology itself. The method is presented through a case study. Twelve professional electric guitar players used an augmented plectrum to replicate prerecorded timbre variations in a set of musical excerpts. The task was undertaken in two experimental conditions: a reference condition and a subtle, gradual change in the augmented plectrum's sensitivity, designed to affect the guitarist's performance without making them consciously aware of its effect. We propose that players' subconscious response to the disruption of changing the sensitivity and their overall ability to replicate the stimuli may indicate the strength of the relationship they developed with the new interface. The case study presented in this chapter highlights the strengths and limitations of this method.

New digital musical instruments face many barriers to adoption, both technical and human. Skill acquisition poses a particularly vexing problem: skills that performers acquire over an extended time on traditional instruments do not necessarily transfer to new instruments, with the result that expert-level performances on new instruments remain relatively rare [1, 2, 3].

It is appealing to seek technical solutions to problems of human sensorimotor learning by seeking to leverage existing skills in new designs [93]. Examples of skill transference can be found in commercial instrument design, including the electric guitar and the inclusion of the familiar piano-style keyboard on Moog synthesisers. However, it is far from obvious how to build on existing skills in the general case.

Before answering such a question, we should first ask how we can even evaluate whether a new instrument uses a performer's existing skill. How do we know how far existing sensorimotor skills can transfer? When a performer is confronted with a modified or unfamiliar instrument, how do we learn to what extent their performance uses existing training? This chapter presents a quantitative method for analysing the encounter between a performer and a partially familiar instrument. To what extent do performers adapt their playing to achieve specific sonic outcomes on the new instrument versus simply continuing with existing motor programs from their familiar technique, mainly ignoring the difference in sound produced by the new instrument?

5.2 Performer Study

Professional guitar players were asked to use an augmented guitar pick able to modify electric guitar sound to replicate a set of musical excerpts. Players' musical background (rock-blues music) was central in defining the sound modification's aesthetic and the type of musical excerpts. The augmented plectrum modifies the guitar sound in such a way that it resembles a wah-wah effect (often found in the rock blues repertoire). The kind of stimuli proposed during the study is blues licks[1].

Matching the aesthetic between players' repertoire, sonic augmentation and musical excerpts, I aimed to engage participants in an extension of their performance practice using the augmented plectrum rather than engaging them in a completely unfamiliar activity that could have entirely disrupted their embodiment with the plectrum making embodiment evaluation pointless.

The study stimuli are characterised by timbral modifications achieved using the augmented pick during their recordings. Participants were asked to replicate the licks with respect to pitch and timbre. The first part of the study addresses the research question: how well did participants replicate timbral modifications

[1] A blues lick is a short melody. It is approximately two bars long.

of an electric guitar sound using a modified plectrum for which they do not have an established sensorimotor program?

In the second part of the study, a very slow low-frequency oscillator (LFO) gradually changes the augmented plectrum effect. The guitar sound becomes darker if participants do not adjust their playing to compensate for that change (more details about the LFO in section 5.2.4).

The research questions for this second section are: to what extent are participants listening to the timbre variations produced by the augmented plectrum? And provided that they perceive the sonic result of the sensitivity change in the augmented plectrum behaviour, do they adapt their playing? I did not brief participants about the LFO effect. If participants act to compensate for the LFO effect, we might conclude that they are listening and responding to the sonic modification produced by the augmented plectrum. Suppose they were adjusting their playing without awareness of its disruption (thus subconsciously). In that case, we may infer that they could use the augmented plectrum without losing focus on the external musical environment, which may signify they were able to use their existing motor skills to play the instrument and reach a state of unconscious competence [11].

5.2.1 The Magpick

The augmented plectrum used in this case study has been presented at the NIME 2019 conference [82] and is called Magpick. The Magpick (Fig. 5.1) consists of a hollow, custom-designed pick with several loops of wire embedded within it. When the pick moves within a magnetic field, such as that created by the magnets in an electric guitar pickup, a voltage is induced in the coil, which is proportional to the rate of change in magnetic flux.

Figure 5.1: The Magpick.

In the context of guitar playing, this signal is related to the speed of movement, the angle of the pick with respect to the guitar body, and the proximity of the pick to the pickups. As a consequence, the voltage generated in the wire embedded in the pick provides information about the gestures of the plucking hand. The wire is connected to a small preamplifier embedded in a box that can be worn on the wrist. The preamplifier's output is connected to a Bela audio processor [86], which combines with the guitar's signal to produce a new output signal that modifies or extends the guitar's sound. This solution accurately responds to the speed, location, and intensity of the pick movements in the pickup area with a wide dynamic range.

Therefore, the Magpick can be described as a guitar pick integrating a sensor which detects a combination of the quantity of movement in the picking gestures and proximity to the electric guitar pickup. Playing louder or closer to the pickups produces a stronger signal. The resulting signal can be applied to control an audio effect that modifies the guitar timbre.

In this study, the Magpick controls a resonant bandpass filter whose cutoff frequency follows the amplitude envelope of the Magpick signal. As discussed in the previous section, the way the Magpick modifies the guitar sound resembles the effect of a wah-wah pedal. However, it offers players a faster and more nuanced way of controlling the guitar sound. In fact, players can use picking skills developed over years of training to modify the filter's cutoff position.

The following link gives an audio-visual description of how the Magpick works and of how it modifies the electric guitar sound in this study: `https://www.youtube.com/embed/dz9isJfjf4U?feature=oembed`.

5.2.2 Participants

Twelve professional electric guitar players, working either as university tutors or as session musicians, were invited through an open call sent to music schools. Each guitar player filled out a questionnaire before the start of the study indicating: their demographics, years of training, and repertoire. Participants' demographics resulted in an average age of 41 years old. 38 % of participants identified as female, while 62 % identified as male. The average years of study of the instrument among participants were 25.

During the study, some participants played their own instrument, while others played a guitar offered on-site (a Fender Squier Stratocaster).

Players were paid an hour at a professional rate (50 GBP). The study took place in a music studio within the Queen Mary University of London (Mile End campus) and lasted one hour. The study was approved by the Queen Mary Ethics of Research Committee.

Participants were familiar with using traditional plectrums with electric gui-

```
|---------------------------|
|---------------------------|
|-------9---------12--------|
|-9g11-----11---------------|
|---------------------------|
|---------------------------|
```

Figure 5.2: the tablature for a stimulus played during the study.

tars, had played musical repertoire which mainly included rock-blues music, and were unfamiliar with the augmented plectrum: the Magpick [82].

5.2.3 Stimuli

A total of 48 stimuli were recorded using the Magpick in the days prior to the study. Each stimulus was a blues lick. Fig. 5.2 shows the tablature for a stimulus. The following link gives an example of a stimulus used in this study: `https://youtu.be/nIadS_MLTko`.

Each lick was two bars long and played in a 4/4 time signature. Sixteen stimuli were recorded while gradually increasing or decreasing the brightness of the guitar sound (i.e. increasing or decreasing the quantity of movement applied to plucking gestures over time, thereby changing the amount of sweep of the filter frequency). Sixteen stimuli were recorded, keeping the timbre constantly brighter or duller (i.e. plucking the strings with the Magpick with a constant quantity of movement (strength) for each picking gesture). Sixteen stimuli were recorded, making the guitar sound brighter or softer for specific notes of the excerpts (i.e. applying a higher or lower quantity of movement in plucking some of the notes of the passages).

5.2.4 Procedure

. Before beginning the study, participants were briefed about how the Magpick works and could try it by playing two blues licks provided as trial stimuli. Participants reproduced sixteen guitar licks in the first section of the study and sixteen in the second section. The order of the sections and the order of the stimuli were randomised for each participant. Stimuli were selected randomly from a list of recorded licks. Licks were shown one at a time on a monitor as tablature and played back using speakers. Players were allowed three attempts to reproduce each lick. Only the last attempt for each lick is used for analysis as it possibly represents the moment of maximum familiarity with the stimulus and, therefore, the best performance. In the first part of the study, the Magpick sensitivity was not affected. In the second part of the study, the Magpick sensitivity was subtly disrupted over time by a slow LFO to evaluate players'

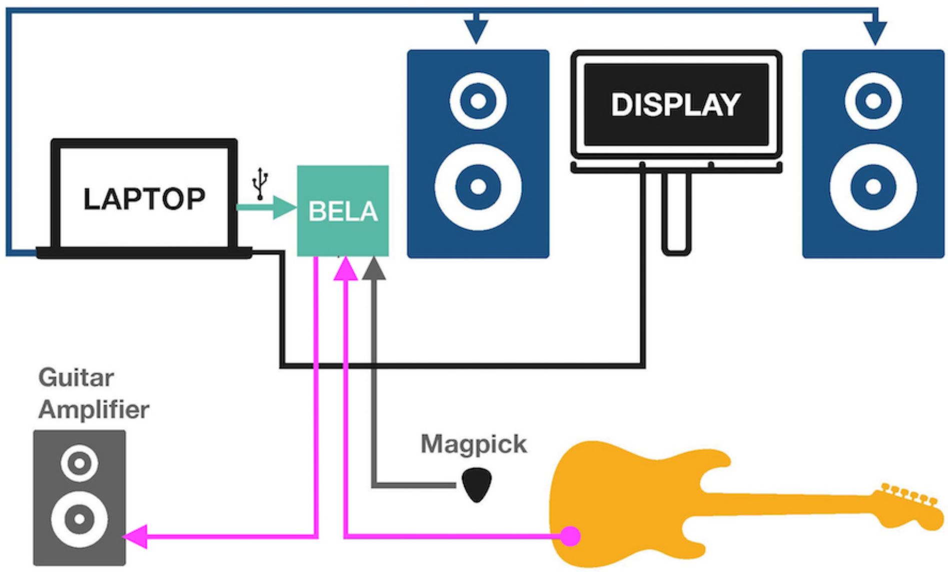


Figure 5.3: study apparatus.

ability to use their skills to compensate for the disruption. The LFO has a triangular shape and decreases the Magpick effect by 33 % over one minute. When the value of the LFO increases, the sensitivity of the Magpick decreases and guitar players need to pluck the strings with more strength to open the cutoff filter and achieve a bright sound. The guitar sound becomes darker if participants do not adjust their playing to compensate. The LFO takes the Magpick sensitivity back to its normal state over one minute.

5.3 Apparatus

The study apparatus can be divided into devices and software to generate and display the stimuli, devices and software for players' performance, and devices and software to collect participants' performance data. Fig. 5.3 gives an overview of the physical devices included in the study apparatus and their physical configuration.

The portion of the study apparatus to present the stimuli included: a display to show the stimuli tablatures and audio speakers to present the related stimuli audio recordings. Both the display and the speakers were connected to a laptop which served as a way to display the graphic user interface to operate the study (play-stop the stimuli, record data) on the researcher's side. The graphic user interface was stored and run on a Bela connected to the laptop. Thanks to a custom p5.js script and the Bela integrated developed environment (IDE), it was displayed on the laptop display. Before the study took place, the interface allowed to calibrate during a pilot study[2]:

- the guitar volume
- the cutoff range of the filter and its resonance value
- the volume of the stimuli playback

.

The portion of the study apparatus that allowed players to perform included: a Fender Squier Stratocaster (which players could choose to use instead of their guitar), a guitar amplifier, and the Magpick. The Magpick signal (whose value ranges from 0 to 1) was fed into a Bela and processed through an envelope follower filter effect written in C++. The code takes the envelope of the Magpick signal to control a resonant filter. An absence of interaction with the Magpick (i.e. zero Magpick signal) results in the filter cutoff being set at 164 Hz (which corresponds to the musical note E3). By contrast, the maximum interaction with the Magpick (i.e. the hardest possible playing) results in the filter cutoff being set at 5274 Hz (corresponds to the musical note E8). The filter Q is set to 8, a distinct resonance that emphasises the filter's sweep controlled by the Magpick. The attack time interval for the envelope follower engine is set to 1 ms so that a sudden picking gesture immediately opens the filter and has a release time of 300 ms to allow for the filter sweep sonic effect to be perceived over time.

In part 2, the code also starts an LFO that affects the sensitivity of the Magpick. The performance envelope is calculated as the Magpick envelope multiplied by one minus the LFO value. A direct effect of the LFO signal increasing is a diminished Magpick sensitivity, leading to a more filtered sound. As an effect of the filtering, the guitar sound manifests a lower amplitude. Fig. 5.4 details how the LFO signal affects the Magpick signal and influences the guitar amplitude.

The portion of the study apparatus to collect participants' performance data and to generate data logs included a custom C++ code running on Bela. The code stored the guitar audio signal, the Magpick signal, and the LFO signal as an interleaved audio channel (at 44.1 kHz) to synchronise the three signals. The

[2]The pilot study took place in the same space used for the following experiment and involved a guitar player doing research at Queen Mary University.

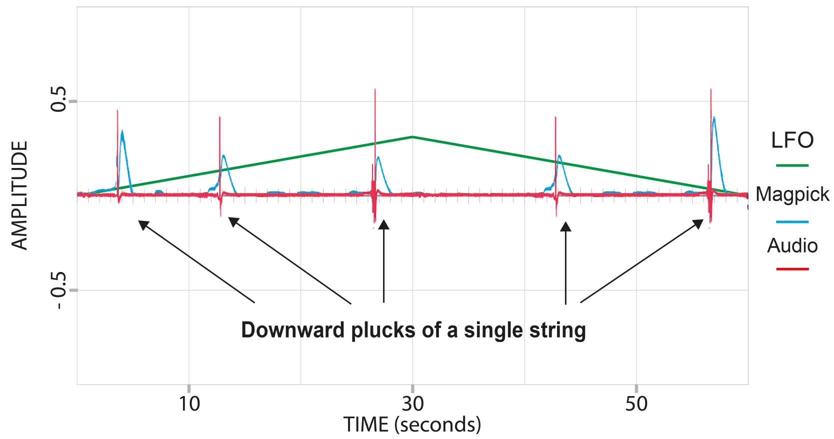


Figure 5.4: Effect of the LFO on Magpick signal and the guitar amplitude.

file name for the resulting file included a numeric participant ID. The code also stored the generated list of stimuli for each study session in a text file.

5.4 Preparing the Data for Analysis

The following paragraphs describe the extraction of the data from the recorded audio and sensor signals and the organisation of the data in a single data frame.

The Magpick reference envelope (generated while recording each stimulus), the corresponding performance envelope (generated during participants' performance), and the LFO generated in part 2 to modulate the Magpick sensitivity were recorded on Bela as 44.1 kHz, 16-bit signals.

The audio files were then imported into the Audacity audio editor [94] on a laptop. The envelope signals were filtered with a 4th-order (24dB) low-pass filter set to a 1Hz cutoff to retain the large-scale shapes of the envelopes while de-emphasising short transient events, which might occur at slightly different times between stimulus and performance. To compensate for the group delay introduced by the filter, I reversed and filtered the signals again using the same settings for a 48dB/octave total slope.

The twelve participants' filtered performance envelopes were concatenated in a single audio file. Likewise, all the filtered reference envelopes of the twelve participants were concatenated in an audio file. The start of each reference envelope, the onset of every performance envelope and, for part 2, the corresponding LFO segment were aligned manually using Audacity to allow for correlations and comparisons. The signals' end was truncated so that correlation tests did not involve portions of the files that displayed silence. The signals were imported into Sonic Visualiser [88] and exported as CSV files (listing one amplitude value for each sample). The resulting CSV files were merged into a single database and imported into R Studio to [87] evaluate their relationship. All the statistical analyses presented in this study were conducted in R Studio.

5.5 Statistical Tests

This section provides the list of statistical tests adopted to analyse the data produced during the study. Each test considered the pairs of variables listed in table 5.1. The following paragraph briefly lists the tests considered to analyse the data. Then, a short description of each test is provided.

Data analysis started by checking the presence of influential outliers using Cook's distance lines. Then, the Breusch-Pagan test and linear regression provided insights into the relationship between each pair of variables. Finally, correlation tests determined the strength of their relationship.

The presence of influential outliers was tested using Cook's distance lines. The Cook's distance is considered high if it is greater than 0.5 and extreme if it is greater than 1. Breusch-Pagan tests offered insights into the type of relationship between each pair of variables[3].

The test indicates a linear relationship when it returns a p-value < 0.05. Linear regression indicated the best linear fit between the variables, and the linear regression coefficients offered insights into the mathematical relationship between each independent variable and the dependent variable. Specifically, the regression coefficient, called slope value, describes the change of the dependent variable (the performance envelope) for a one-unit increase of the independent variable (the reference envelope or the LFO envelope).

The sign of the slope value determines the direction of the relationship. If the sign is positive, both variables increase or decrease in the same direction. If the sign is negative, the dependent variable decreases when the independent variable increases and vice-versa.

The residual standard error and the square of the correlation coefficient, called "R-squared" R^2, give an indication of how well a model describes the relationship between the variables. The residual standard error measures the standard deviation of the residuals in a regression model. The square of the correlation coefficient measures the "fit" of the regression line to the data and tells how well the linear regression model fits the data. R^2 values have a range between 0 and 1. A high R^2 value suggests a linear relation. The p values

[3]The Breusch-Pagan test evaluates whether heteroscedasticity is present. Heteroscedasticity is a condition that suggests a non-linear association between the variables.

for the coefficients indicate whether the relationships between the variables are statistically significant.

Pearson's test determined the significance and the direction of the relationship between each pair of variables with an index that comprises between 1 and -1. 1 means that the variables evolve in the same direction, 0 means that they are independent (hence, they have no relation), and -1 implies that the variables grow in opposite directions.

5.6 Statistical Results

Table 5.1 offers a summary of the main tests' outcomes and the pair of variables considered for statistics. More details are presented in the following paragraphs in this section. The results are discussed in section 5.7.

Variables	Slope Value	Residual Error	Pearson Test
Part 1: Perf Env vs. Ref Env	0.56	0.17	0.51
Part 2: Perf Env vs. Ref Env	0.45	0.15	0.52
Part 2: Perf Env vs. LFO Env	0.44	0.19	0.21

Table 5.1: summary of the main outcomes from the data analysis.

Computing Cook's distance lines did not show any influential outlier for each pair of variables while the Breusch-Pagan test suggests a linear or partially linear relationship for each considered case with $p < 2.2e\text{-}16$.

Linear regression for each pair of variables returned positive slope values, suggesting a relationship between each pair of envelopes. The sign of the slope coefficients is positive. As discussed in section 5.5, this means that variables evolve in the same direction. In part 1, the linear regression model presents a slope value of 0.56, indicating that a change of 1 unit in the reference envelope yields a change of 0.56 in the performance envelope. In part 2, the linear regression model between the same pair of variables presents a slope value of 0.45 while the linear regression model between the performance envelope and the LFO envelope presents a slope value of 0.44. The relationship between each pair of variables is statistically significant, with $p < 2e\text{-}16$.

Models' ability to predict the variables' relationship is characterised by small residual standard errors of respectively 0.17 (performance envelope vs reference envelope in part 1), 0.15 (performance envelope vs reference envelope in part 2), and 0.19 (performance envelope vs LFO envelope). The R^2 values for each relationship are possibly small and respectively 0.26, 0.25, and 0.19. These results are statistically significant with $p < 2e\text{-}16$. The implications of small R values for this study's results and its limitations are discussed in section 5.7.1.

Pearson's product-moment correlation tests confirmed the existence of a significant relationship between each pair of variables and that the variables evolve in the same direction. The test between the performance envelope and the reference envelope in part 1 returned a correlation coefficient of 0.51 with a 95 % confidence interval between 0.5 and 0.52. The test between the same pair of variables in part 2 returned a correlation coefficient of 0.52 with a 95 % confidence interval between 0.51 and 0.53. The test between the performance envelope and the LFO envelope in part 2 returned a correlation coefficient of 0.21 with a 95 % confidence interval between 0.2 and 0.22. The results of the three tests are statistically significant, with $p < 2.2e\text{-}16$.

These tests, therefore, show a correlation between the reference envelopes that were in the stimuli and the performance envelopes that the performer played in response to that, suggesting that overall the performers were able to execute the study tasks with almost a modest degree of accuracy. This will be further discussed in section 5.7.

5.7 Discussion

Data gathered in this study evaluates a certain kind of skilled action: the picking gesture. This skilled action was correlated with the amount of motion produced by recording the stimuli. Participants' feeling was not measured. I rather focused on measuring their action and the quantity of movement in their plucking the strings to get a more quantitative picture of whether a performer is managing to retain their ability to perform with a modified plectrum.

On a scale ranging from -1 (inverse correlation) to +1 (positive correlation), participants could match the timbre stimuli with a correlation coefficient of 0.5 in both the first and second parts of the study. For every change of one unit in the reference signal, the performance envelope changed in the same direction by 0.56 in Part 1 and 0.45 in Part 2. The resulting correlations show a reasonable degree of correspondence between stimulus and performance, which is not present in correlation analyses between deliberately unrelated signals (see Limitations in section 5.7.1). Thus it gives us the confidence that performers were executing the task of replicating the stimuli with at least a modest degree of accuracy. In other words, the positive correlation values and the positive slope values suggest that participants could play the Magpick to open and close the cutoff frequency of the filter applied to the guitar sound as it was recorded while generating the stimuli.

In Part 2, data shows that participants adapted their playing to the LFO effect with a positive correlation of 0.2 and a positive slope value of 0.44. In other words, when the LFO value was increasing, making the Magpick signal less sensitive, participants were also increasing the magnitude of their interaction with the Magpick (i.e. picking the strings with more strength and-or closer to the pickups). As discussed, the effect of the LFO on the Magpick sensitivity is audible as a changing filter cutoff. Since there is no other way for a performer to discover the effect of the LFO, the correlation analysis suggests that participants must be listening to the guitar sound modified by the Magpick, noticing its change either consciously or subconsciously. In turn, they adapted their playing to partially (though not fully) compensate for the effect of the LFO disruption.

Participants were not briefed about the LFO disruption before or during the study. At the end of the experiment, they were asked if they had noticed any change in the magpick behaviour. None of them reported having experienced a difference in the Magpick sensitivity. I thus speculate that participants were not only listening to the sonic augmentation and reacted to the LFO but also that their reaction was unconscious as they did not report its effect. Adapting their playing by adjusting their picking gestures became an automatic subconscious action, possibly similar to the act of placing their finger on the fretboard or plucking the strings. Their interaction focused on the musical task (replicating the guitar lick's timbre) rather than shifting toward the technology (the change in the Magpick sensitivity produced by the LFO).

A learning process is generally required to build skills like this. Professional players spend a lot of time building skills on one interface. A designer then either changes some aspect of the interface and may try to build on the same skills. The method presented in this chapter tells how well the design does with that change of the interface (the augmentation of a plectrum). Can people adapt their existing skills, acquired using a normal pick, without further training, or are they set back in their ability to play? The fact that participants never saw the Magpick motivates the study. Can somebody achieve the desired outcome without resorting to a high cost of conscious attention? Not any subconscious action performed during execution is indeed a result of maintaining their ability to focus on the sonic outcome of performance rather than on the instrument's functioning or gestures. For example, there are ancillary gestures that are not such an indicator. However, in the study, I address a specific type of gesture directly related to the performance. Specifically, the picking gesture.

The correlation values discussed in this chapter could undoubtedly have been higher. Eight participants out of twelve stated that they had to pick the guitar strings stronger than they used to replicate certain stimuli. Being required sometimes to pick the strings stronger than usual may have affected their ability to match the stimuli and, in the second section, adjust for the LFO.

5.7.1 Limitations

The meaningfulness of the correlation analyses was checked against baseline correlations and linear regression models performed on unrelated variables. For example, I computed a linear regression model with the reference envelope from part 1 and the LFO signal from part 2. The test returned a slope < 0.00 with R^2 7.494e-06 and $p = 0.617$. A further linear regression test conducted between the reference envelope from part 1 and the performance envelope from part 2 returned similar non-significant results. However, it also returned a significant p-value. We may conclude that the p values are not always reliable in the linear regression tests applied to this dataset. The s slope and the R^2 values describe the relationship of the variables more accurately. The R-squared results show low values, suggesting low predictability for the model. In other words, the slope values may not be perfectly representative of the numeric relationship between stimuli and responses. The study is based on human-based tasks, possibly leading to uncertainty in the data.

The human-based nature of the study may also have led to a partially non-linear relationship between the variables that, in turn, affects the predictability of the calculated regression models. However, the models computed in this research are not meant to precisely predict the performance values based on the stimuli values. Instead, they help get insights into the data (i.e. whether participants increase their picking strength to achieve brighter sounds when the stimuli sound is brighter). Future research may adopt different statistical tests to measure the correlation between the envelope signals. Primarily tests that are meant to assess partially linear relationships between variables. I relied on participant self-reports to determine whether participants were consciously or unconsciously reacting to the LFO. Choosing whether the players' response to the disruption is conscious or unconscious is critical in determining the stage of instrument motor learning experienced by players. For this reason, additional research strategies are needed to reinforce the hypothesis that performers not only responded to the disruption and adapted their playing but also did it unconsciously as a result of responding to the auditory feedback of the augmentation.

The subjective feeling of participants was not measured; I measured their action and the quantity of movement in their picking gesture. Evaluating to what extent players are experiencing a subconscious response to the LFO may be the subject of future work. In this study, I instead tried to bring an external view to whether somebody can execute skilled actions on an unfamiliar interface. The goal is not to privilege an objective method against subjective methodologies but rather to complement existing methods with something that is outwardly observable and repeatable.

Having different electric guitars in the study (participants were allowed to use their own guitars) has possibly introduced a source of variability and unfamiliarity in the system's behaviour. It might have been good if players had a more ample opportunity to play on that guitar before introducing the Magpick. However, guitarists usually adapt to switching guitars, so it may not have affected their ability.

5.8 Conclusions

This chapter proposed an evaluation method to examine expert players' repurposing of motor skills for new digital or augmented instruments. The evaluation method is quantitative, based on simple correlations based on the replication of target stimuli and slow changes to action-sound mappings. This study is one instantiation of a method that can be used more widely to understand a musician's relationship to their instrument. The method will be most beneficial for instruments that are intended to repurpose existing sensorimotor skills and be characterised by predictable and repeatable forms of interaction. Evaluating new musical interfaces in such a context can be challenging as it requires observing activities that happen subconsciously and cannot be easily queried. The results from this case study appear to show at least a modest subconscious response to changes in augmentation behaviour, and the principles introduced in this research could be adapted to other scenarios in new instrument research. More information on this methodology and its possible contributions to evaluating new or modified musical instruments that aim to transfer players' skills will be discussed in chapter 7.

Chapter 6

Streams of execution

6.1 Introduction

In this study, players are asked to execute musical excerpts using a digital piano with an extra foot pedal compared to what people normally use. The pedal presents two different behaviours given to participants in two different conditions. The study aims to compare the effect of the two behaviours on players' performances. Which of the two designs allows the pedal to better integrate into the overall execution? The answer to this question may inform the design of musical instrument modifications or augmentations that aim to be easily integrated into existing performative practices.

The study does not aim to evaluate the pedal as a specific interface. The action of tapping a pedal during a piano performance is regarded as a good example of a performing process that superimposes the playing of a familiar instrument (the piano) with another action (pressing the pedal) which might not be familiar or may not have familiar results (e.g. triggering a sound or transposing the piano keyboard). Do participants integrate the gestures they need to perform on the pedal into the same stream of imagery they use to perform the piano, or do the actions they perform on the piano and pedal remain entirely separate streams of execution in the player's mind?

The underlying hypothesis is that a design that better integrates into the overall performance should allow participants to conceive the pedal execution as part of the musical outcome produced on the keyboard. I speculate that players should be able to imagine the overall performance (including the piano and the pedal) as a single process rather than two parallel mental processes to retain their fluency and tempo accuracy. Otherwise, players would experience a state of impaired fluency in which the tempo of the performance may fluctuate according to the difficulty experienced by performers in integrating the modified instrument. This challenge would also affect the timing and accuracy of the execution.

6.1.1 Research Questions

We compare the effect of two designs on players' performance. Players' performance is examined in two conditions: a sound condition where the sustain pedal triggers the sound of a cymbal, and a modification condition where the pedal transposes the piano up by 1 octave.

which of the two conditions better supports participants in integrating the pedal in the overall musical executions?. Will direct auditory feedback in the sound condition allow participants to increase their timing accuracy of the pedal events? In the octave condition, will the lack of immediate sonic transformation (transposition) in response to pedal pressure lead to lower accuracy in their pedal timing? Which condition will better integrate into the performance model where participants subconsciously imagine the sound they want to achieve, act to achieve it and then compare the result with their initial expectations?

I designed a study with expert pianists to answer these questions and analysed the resulting performances. Data are analysed concerning the tempo stability of the player's execution on the piano and the time accuracy of the pedal execution.

6.2 Context

During the transposed violin study study 4, players were asked to use a re-tuned violin to perform musical excerpts. Players could play the re-tuned violin in the *transposed notation* by focusing on the notation and ignoring the auditory feedback. To a certain extent and with a very reduced fluency, they could also perform in the *concert notation* condition by taking extra time to think about the fingers' position on the re-tuned strings. The same process was observed by Morreale et al. [82]. Guitar players could use the unmodified aspect of an augmented[1] plectrum (shape, ergonomics) to perform on a guitar. However, when they tried to control the augmented aspect of the plectrum that could modify the guitar sound, they needed extra time to engage in a trial-error process to integrate the pick-modified behaviour into their playing.

Participants in both studies seem to operate the instrument and its modification as two separate entities. The operations devoted to the traditional side of the instrument (holding the pick, and plucking the strings) were performed

[1]The plectrum was able to sense the strength of their picking gestures and influence the guitar's timbre or dynamic.

as usual. The operations needed to control the instrument modification (i.e. positioning the pick closer or farther from the guitar pickups) appeared to constitute a separate set of gestures which required more time to "think" about how to control it.

I speculate that these two distinguished sets of actions are, to some extent, the result of two streams of thought. Equally, I speculate that the modified instrument is conceived by players, to a certain extent, as two separate instruments, resulting in two separate yet concurrent performances. The performance with the traditional aspects of the instrument retains a certain fluency. The performance with the modified aspects of the instrument does not.

It has to be noted that I am not claiming that such a separation constantly exists and is in the same shape throughout a whole performance. Still, it seems that the instrument modification and the underlying aspects of the original instrument can each have a different impact on the performance.

Engaging in two co-occurring activities can be challenging as it requires, trivially speaking, doing two things simultaneously instead of one. A specific design may or may not help players integrate (physically and/or mentally) the two sets of gestures needed to perform on the modified instrument and its modification into one task. We may also speculate that a particular design modification may produce different effects if we accept this idea.

I also speculate that a design that achieves such a result may differentiate itself from a design that does not reach it. In the first case, participants would have to focus on one overall task instead of two concurrent tasks resulting in a more fluent execution.

6.3 Performer Study

Twelve professional keyboard players with ABRSM Grade 7 qualification[2] or higher were invited to join the study. Players were familiar with the western classical music repertoire and playing sustain pedals. They had an average age of 31; the gender distribution was 58 per cent female and 42 per cent male. The average years of study were 22. They were paid an hour at a professional rate (40 GBP). The study took place in a music studio within the Queen Mary University of London (Mile End campus) and lasted one hour. The study was approved by the Queen Mary Ethics of Research Committee.

During the study, performers played ten notated musical excerpts using a digital piano and a momentary foot pedal[3] (mechanically equivalent to a digital

[2]ABRSM (Associated Board of the Royal Schools of Music) is an accredited board awarding exams and diploma qualifications in music within the UK

[3]Clavia Nord Single Pedal. Model number:

piano sustain pedal) in two different conditions. In the first condition (sound condition), the pedal pressure triggers a short-ride cymbal (with a sharp attack and one-second decay). In this condition, the pedal does not modify the sound of the piano and triggers a sound when pressed. I chose a cymbal sound because it is close to a broad western cultural context of piano executions involving drums.

In the second condition (octave condition), pushing the pedal transposes up the keyboard register by one octave. In other words, all the notes played after the pedal tap are transposed up by one octave. A subsequent pedal tap brings back the piano register to the original pitch. This latter modification modifies the sound of the piano *after* the pedal is pressed.

Excerpts were presented as paper scores, taken from the music book Mikrokosmos by Bela Bartók. Fig 6.1 shows the first page of a musical excerpt used during the study. The order of the musical excerpts is randomised for each participant, so it is the order of the conditions. Each excerpt is one or two pages long. Excerpts two pages long were shown with one page next to the other, so participants did not need to interrupt their playing during the execution to turn the page. A version of the music passages without pedal markings was sent to the participants four days in advance to familiarise themselves with the piano notation. They were otherwise unfamiliar with the musical material. During the study, participants played the notation, including the pedal markings. Pedal markings were not originally part of the notations. I added them using a third staff.

In the sound condition, we might expect less accurate performances because players have to keep track of two different types of sound simultaneously. On the other hand, this condition has immediate sonic feedback. So it may facilitate the integration of the pedal in the performance. The octave condition, by contrast, does not provide immediate sonic feedback to the pedal taps. This may also cause less accurate performances because players have no sonic feedback on whether the pedal tap was early or late. However, players may be able to play the pedal more accurately because the pedal transposes the piano by one octave preserving the musical sense and coherence of the excerpts in terms of melody, harmony and timbre. The pedal could be better integrated into a single stream of performative gestures related to the piano execution.

Pedal markings are initially positioned as sparse material, and then they gradually increase in density and complexity. I refer to the word density as the number of pedal events in a bar. I refer to the term complexity as their rhythmic position concerning the piano execution (e.g. on a strong beat or syncopated). The complexity of the musical material increases over time. Increasing the difficulty of the notation, we may observe that the timing of the pedal's execution may be more precise in one of the two conditions. As the difficulty of the music

Figure 6.1: the first page of a musical excerpt.

material increases, players are pushed into a musical space where they do not have the time to think [4]. In the same way, if someone cannot drive a motorcycle well, they could drive slowly and carefully. But if they want to race on a track, they need expertise. The more complex the challenge of playing musical excerpts, the more we may observe whether players internalise one of the two pedal modifications. If one of the conditions is particularly difficult, we may observe some pianists stop operating the pedal (or exhibit abysmal timing) and focus on the rest of the performance. They may need to stop thinking about the pedal and focus on the notes they need to play.

6.4 Apparatus

The keyboard used for this experiment is a Nord Stage Piano. The sustain pedal is a standard NORD Sustain Pedal and is connected to the keyboard. The control change messages produced by the pedal (with values of either 0 or 127) and MIDI notes made by the keyboard are sent from the keyboard to a laptop using a USB cable. The computer runs an instance of Ableton Live, which maps and records the incoming MIDI data. MIDI notes messages are routed to a MIDI track that hosts an instance of the Waves Grand Rhapsody Piano by Ableton. The Nord Stage Piano internal sound module is not used. MIDI control change messages are mapped according to the experimental condition using custom Max for Live devices.

The DAW is set up to declare an overall latency of twelve milliseconds. A study pilot with an expert pianist and researcher from the Centre for Digital Music (Queen Mary University of London) did not identify such latency as an issue.

6.4.1 Pedal functioning in the two experimental conditions

In the first condition, each control change message generated when pressing the pedal triggers a ride cymbal audio file. The file is loaded in a virtual sampler hosted on a track on Ableton Live. Control change messages received by Ableton Live are routed to a Max For Live custom device to make this possible. Fig 6.2 shows the flow of MIDI control messages in the first condition. The sampler is set to retrigger the sample each time it receives a note on message.

[4] *Failing* by Tom Johnson's [95] is an excellent example of a music piece that asks double bass players to extend their performing practice and scales upon complexity up to a point to reduce their performing accuracy. Players are requested to play and talk at the same time. The playing and the conversation get more complicated and overlap in the playing. It just becomes impossible to play and talk as the piece prescribes. The piece causes players to split the performance into two musical streams as the difficulty increases.

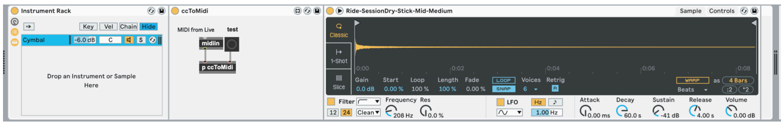


Figure 6.2: flow of MIDI control messages in the sound condition.

The Max For Live device receives the cc message and checks that it has a value different from 0 (0 means that the pedal is not pressed). If the condition is met, then the script triggers a middle C note on the message and, after 100 milliseconds, a middle C note-off message. The note-on and note-off messages are routed to the sampler to trigger the sound (note-on message) and reset the sampler pointer to the beginning of the audio file (note-off message). The routing of the MIDI note messages between the Max For Live device and the sampler is provided by Ableton Live. Fig 6.3 shows the Max For Live Script that maps control messages with control value 127 to a middle C MIDI note.

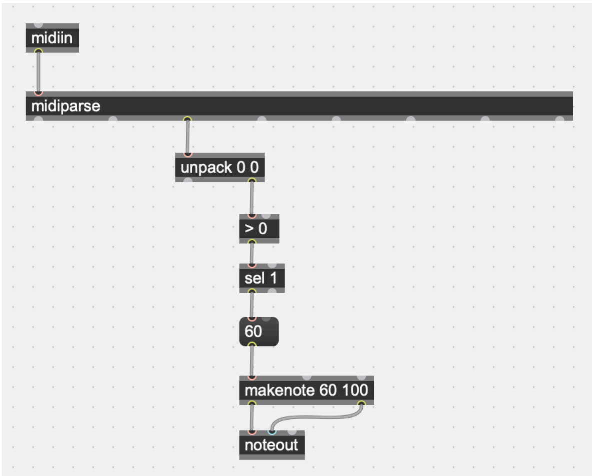


Figure 6.3: custom Max For Live script that maps MIDI cc messages to a middle C MIDI note message.

In the second condition, a control change message generated when pressing the pedal sets a MIDI pitch transposition device to +12 semitones. The device receives the incoming MIDI notes and transposes them before they are sent to the piano VST. A subsequent pedal tap sets the MIDI effect to 0, meaning that the MIDI notes are no longer transposed by one-octave up. A further pedal tap sets the device again to +12 semitones and so on. Control change messages Ableton Live receives are routed to a Max For Live custom device. Fig 6.4 shows the flow of MIDI control messages in the second condition. The Max For Live device shown in Fig 6.5 receives the cc message, and it checks that it has a

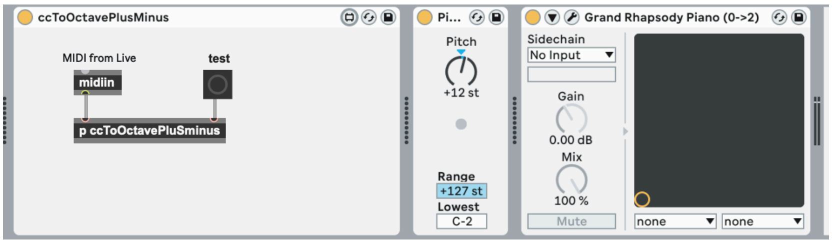


Figure 6.4: flow of MIDI control messages in the second condition.

value different from 0 (0 means that the pedal is not pressed). If the condition is met, the script sets the transposition value of the MIDI pitch effect after a 50 ms delay. Without delay, simultaneous pedal and key presses could produce an unpredictable result depending on which MIDI message arrived first. Because of the delay, only the notes after the pedal tap are transposed.

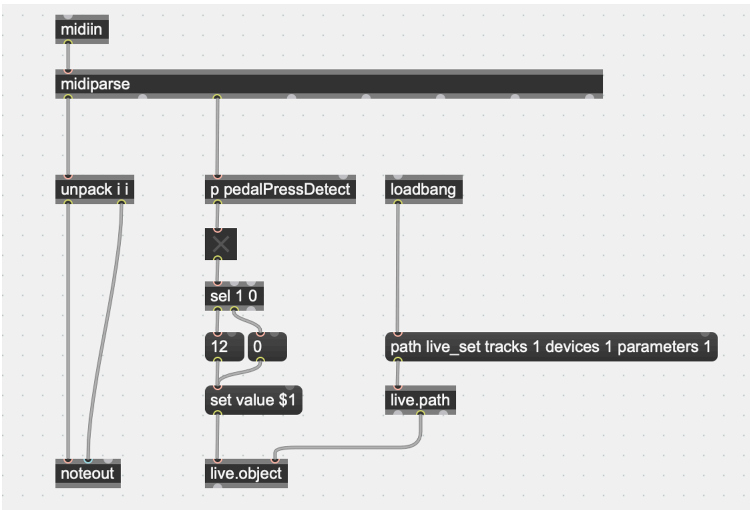


Figure 6.5: custom Max For Live script that maps MIDI cc messages to transposition values on the pitch MIDI effect of Ableton Live.

6.5 Data preparation for analysis

Preparing the data for the following analysis required filtering some of the MIDI information recorded during each performance. MIDI data was recorded using Ableton Live. Ableton Live visualises MIDI messages in a Piano Roll style view. Fig 6.6 shows the MIDI note visualisation in Ableton Live of a performance of a musical excerpt by a participant. Ableton also allows a more complex view where control change messages are visualised on top of the MIDI notes. Both the midi note views and the midi notes plus CC messages views allow to filter or edit of the respective MIDI information. Once the filtering operations were concluded, I exported the MIDI file for each performance. I wrote the following information in each file name: the number of the music excerpts performed, an alias name for the participant who performed it, and the type of condition.

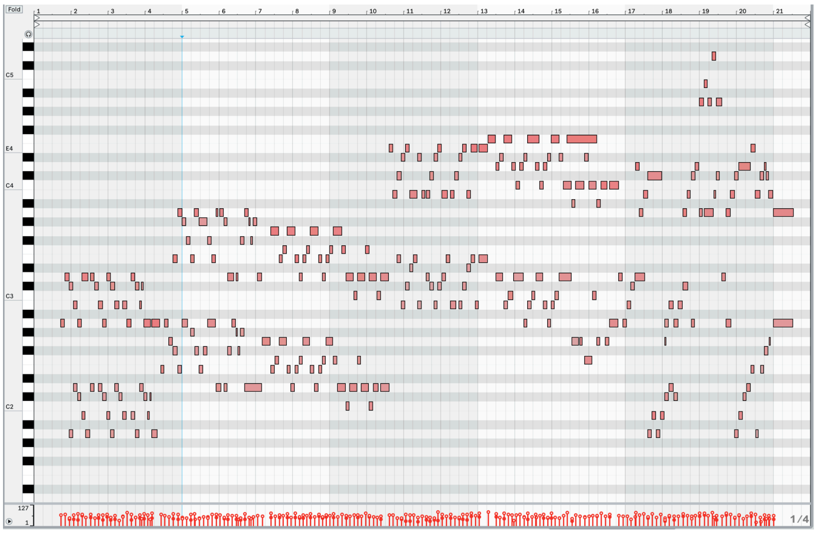


Figure 6.6: MIDI note visualisation in Ableton Live of a performance of a musical excerpt.

6.5.1 MIDI notes data preparation

To calculate tempo variations within each music performance, I selected the MIDI notes falling on specific beats in each measure (more information in section 6.6). I manually annotated their beat number in a text file. I identified these notes by comparing the recordings with the related notation. I selected them by filtering the remaining notes (muting MIDI notes in the piano roll is sufficient to filter them). Fig 6.7 shows a musical excerpt with filtered notes (in white) and selected notes (in red).

For the 3/4 time signature, I kept all the notes falling on each beat. A beat number list may look: [1,2,3,4,...]. For the 4/4 metre signature, I kept the notes falling on each first and third beat in each measure. A beat number list, in this case, may look: [1,3,5,7,...]. For the 2/2 and 2/4 metre signatures, I kept the first and second beats. For the 6/8 metre signature, I kept the notes on each measure's first and fourth beat. A beat number list, in this case, maybe: [1,4,7,10,...]. A missing beat in a measure (i.e. not performed or absent in the original score) was omitted. If two or more notes were overlapping on the same beat, I considered the note with the highest pitch. Fig 6.8 shows the beat number annotation for a few bars of a musical excerpt.

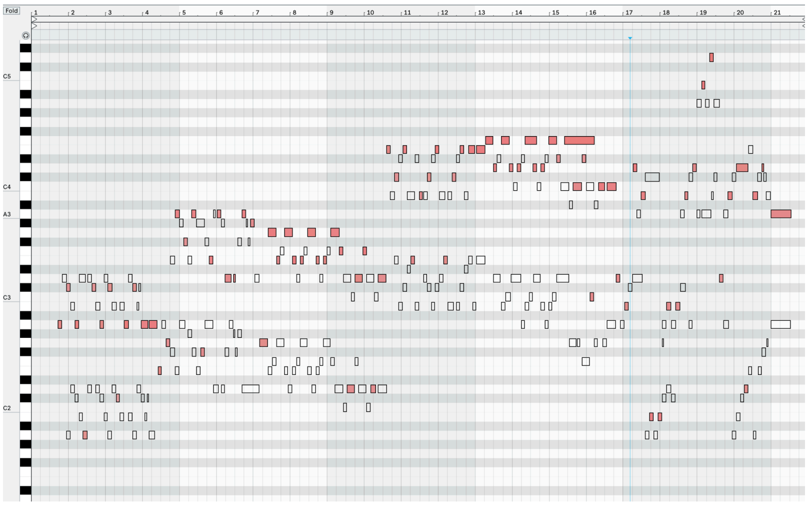


Figure 6.7: MIDI notes filtered are white coloured.

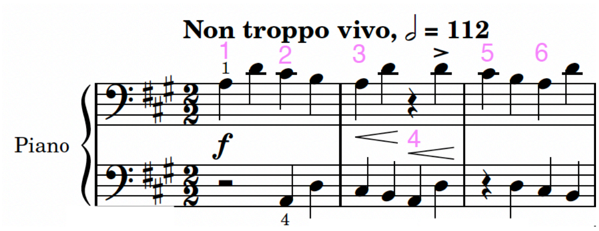


Figure 6.8: beat number annotation on a few measures of a musical excerpt.

6.5.2 Control change messages data preparation

Control change events are visualised in Ableton Live as a blue line with a value alternating between 0 (no pedal press detected) and 127 (pedal is pressed). The cc visualisation is plotted on top of the MIDI notes of the performance, which helps to collocate the pedal taps in the music performance. Fig 6.9 shows the visualisation of MIDI control change messages in Ableton Live for one of the study performances.

I started by identifying and counting how many pedal events notated in the music scores were not performed (missed pedal taps) by participants and how many pedal events were instead performed even in the absence of indication (extra pedal taps). To do so, I manually scanned all the recorded pedal events and compared the pedal executions with the notations. When multiple taps were recorded within a tiny time window (below 50 ms), I deduced they resulted

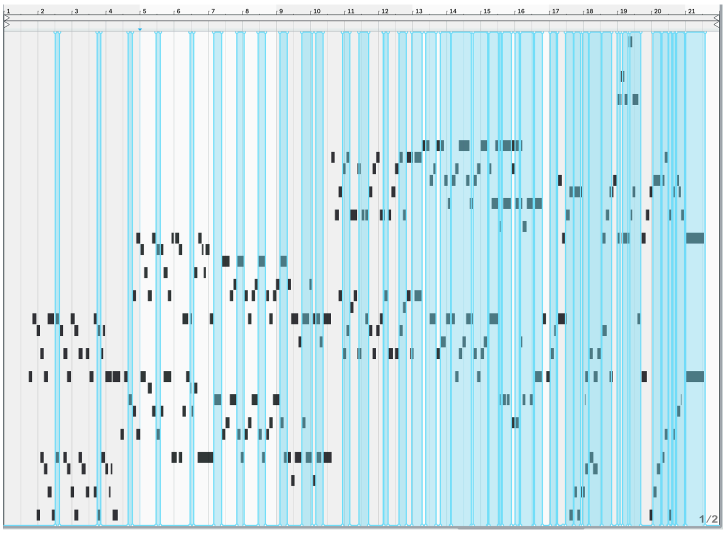


Figure 6.9: MIDI control change messages visualisation in Ableton Live of a performance of a musical excerpt.

from the electrical circuit inside the pedal bouncing. As the first control change message is the one performed by the participant, and the following cc messages are the result of the behaviour of the electric circuit, I kept the first pedal tap and manually deleted the following ones falling within a 50 ms time window. When the notation indicated a single pedal tap, but the participant performed multiple taps, I considered the additional taps as extra taps caused by the participants' errors. I counted these extra taps and took note of their amount. Fig 6.10 shows two a case where the performer played two pedal events instead of the single pedal event written in the music score.

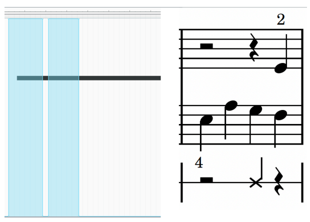


Figure 6.10: two pedal events in rapid succession instead of one.

I then selected the cc messages corresponding to the pedal taps played according to the notation. Ableton Live does not allow filtering cc messages. So I copied the MIDI recordings and deleted all the extra cc messages. I then exported a MIDI file for each performance, including only the cc messages corresponding to the player's performing a pedal tap notated on the score. This information was used to evaluate participants' time accuracy in the pedal execution (more information in section 6.6).

Finally, I manually scanned through the notation of each musical excerpt and made a list of the best number for each pedal event in a text file. Fig 6.11 shows the pedal beat numbers on the first line of a musical excerpt (highlighted in blue).

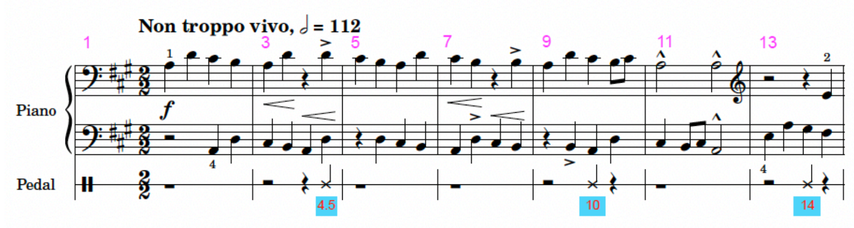


Figure 6.11: pedal beat number annotation, highlighted in blue, on the first line of a musical excerpt.

6.5.3 MIDI data parsing and construction of a dataframe

I coded a custom python script which uses the pretty MIDI library[5] [96] to open the MIDI files exported from Ableton.

The script opens the files, reads the events, and identifies the type of event (note, control message) and the associated data (i.e. note number, velocity or Control number, control value). It extracts the timestamp at which it occurred. Then it reads the file's name and extracts the number of the musical excerpt for the current event, the type of condition in which it was played, and the alias name of the participant who played it. The script then associates the number of beats corresponding to each event by cross-referencing the number of the stimulus, the type of event and the related list of beat numbers created in the previous steps (see MIDI notes and control change data preparation sections). Once these data are available and stored in Python local variables, the script calculates the tempo in BPM between each pair of consequent MIDI notes, and it calculates the time error between the expected timestamp and the actual timestamp of each pedal event (more information about this calculation in section 6.6). All this information is finally stored in a CSV file for analysis.

[5]Pretty MIDI has functions for handling MIDI data so that it's in a format that is easy to modify and extract information from.

6.6 Data analysis

6.6.1 Tempo fluctuation calculation and statistics

The tempo was calculated twice (for 4/4, 2/2, 2/4 and 6/8 metrics) or three times (for 3/4 metric) for every bar of the performance but the last two beats. The last two beats were excluded as I expected performers to slow down toward the end of the musical excerpt for matters of expressive playing. The resulting values calculated across each performance describe its tempo fluctuations.

The tempo in BPM was calculated using the following formula:

$$\frac{\frac{60}{s} \cdot b}{L}$$

In the formula s denotes the start time difference in seconds between two consecutive beat timestamps, and b is the difference between the expected beat number and the actual beat number played by the musician. The result is then normalised by L which represents the length of a beat in the performance (3 for triple metres, 1 for 2 or 4 metres).

I then calculated the standard deviation for each performance as a metric of tempo change using their related tempo lists. I also calculated the interquartile range on the same data as it benefits from excluding outliers' influence. It is the median of the lower and upper half of the data.

I calculated the mean of the standard deviation values and the interquartile range values in each condition. Then I evaluated the statistical significance of the results in the two conditions with a t-test and a Mann–Whitney U test. The Mann–Whitney U test was considered as I do not assume the normality of the data for tempo fluctuation.

It has to be noted that tempo could also fluctuate for expressive/interpretive reasons (e.g. slowing down near cadences). However, a rigorous analysis that tries to unpick all the different sources of tempo fluctuation is beyond the scope of this thesis.

6.6.2 Pedal time error calculation and statistics

I first calculated the expected time of each pedal tap based on the notation and the piano performance. The expected timestamp was calculated in seconds as follows:

$$t_A + \frac{(t_B - t_A) * (b_P - b_A)}{(b_B - b_A)}$$

t_A is the timestamp of the note on the beat preceding a pedal tap, while t_B is the timestamp of the note on the beat right after the pedal tap. b_P, b_A and b_B are, respectively, the beat number of the pedal tap, the note on the beat preceding the pedal tap, and the note right after the pedal tap. The latter part of the formula normalises the pedal event to a fraction of the way between beats A and B. The terms involving t scale that fraction to a range of times in milliseconds.

I subtracted the expected time stamp for each pedal event from the actual time of the pedal execution. The result is the amount of time by which the pedal press is either early or late.

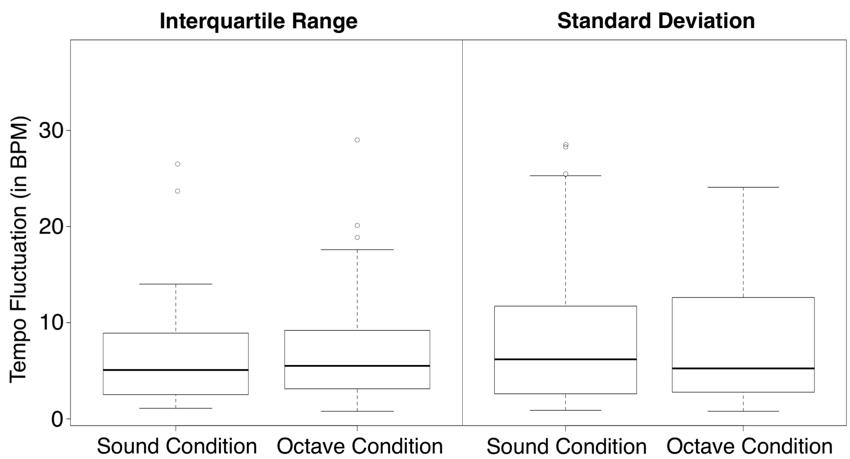
Then I calculated the mean and the median of the pedal time errors in each condition. I ran these statistics on the raw pedal time errors and their absolute values. Then I evaluated the statistical significance of the results in the two conditions using a t-test and a Mann–Whitney U test. The Mann–Whitney U test could be particularly indicated on the mean and median of the absolute values as we cannot expect them to be normally distributed.

6.6.3 Results

Participants had an overall tempo fluctuation across their performances of 7.12 BPM in the sound condition and 7.27 BPM in the octave condition.

A subsequent t-test comparing the interquartile range of each performance in the two conditions showed $p = 0.61$ with a 95 per cent confidence interval. A Mann-Whitney U test returned $p = 0.86$.

Figure 6.12: standard deviation values in BPM for tempo fluctuation in each condition.



Looking at figures 6.15, and 6.16, pedal time errors seem to be normally distributed.

The average Performers' pedal time error calculated on absolute values pre-

Figure 6.13: tempo fluctuation in the octave condition: data distribution.

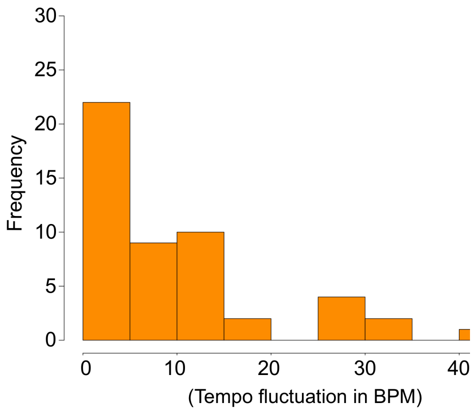


Figure 6.14: tempo fluctuation in the sound condition: data distribution.

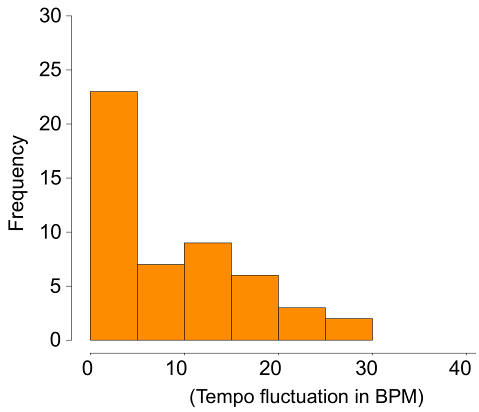


Table 6.1: tempo change statistics.

Condition	**Mean of Standard Deviation Values**	**Mean of Interquartile Range Values**
Sound condition	9.73 BPM	7.12 BPM
Octave condition	7.23 BPM	7.27 BPM
Test on Sound Condition, Octave Condition	**P Value on Mean of Standard Deviation Values**	**P Value on Mean of Interquartile Range values**
T Test	0.9	0.61
Mann–Whitney U test	0.51	0.86

Figure 6.15: pedal time error distribution in octave condition.

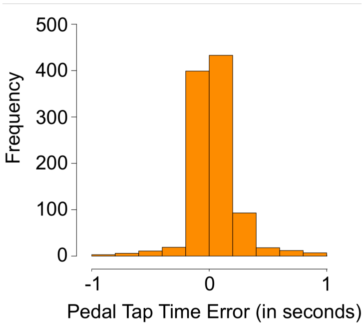


Figure 6.16: pedal time error distribution in sound condition.

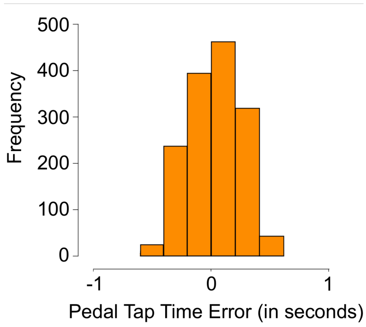


Table 6.2: pedal time error statistic tests.

Condition	MEAN		MEDIAN	
	On Original Values	**On Absolute Values**	**On Original Values**	**On Absolute Values**
Sound Condition	0.048	0.099	0.007	0.026
Octave Condition	0.04	0.118	0.005	0.044

Table 6.3: pedal time error mean and median.

Test on Sound vs Octave condition	P Value on Original Values	P Value on Absolute Values
T Test	0.423	0.026
Mann–Whitney U test	0.401	0.002

sented in the sound condition an absolute deviation mean of 0.099 sec vs 0.118 cents in the octave condition with p = 0.026 (t-test) with a 95 per cent confidence interval and p = 0.002 (Mann-Whitney U test).

The number of expected pedal taps across the study is 2450. Although absent in the notation, the number of extra pedal taps performed is 45 in the sound condition and 54 in the octave condition. The number of taps indicated in the notation but were not performed by participants is 263 in the sound condition and 261 in the octave condition. Table 6.4 summarises these data. The mean missed pedal taps were 3.8 in the sound condition and 4.5 in the octave condition. The mean extra pedal taps were 22 in the sound condition and 21.75 in the octave condition.

Table 6.4: extra and missing pedal taps.

	Total Taps	Extra Taps	Missing Taps
Expected Taps Across the Study	2450		
Sound Condition		45	263
Octave Condition		54	261

6.7 Discussion

Data presented in table 6.1 shows that the tempo variations in the sound and octave conditions do not present significant differences. That means that in terms of the keyboard performance, performers appear to have played roughly the same regardless of what sound the pedal made.

It is possible that a more extensive study would lead to seeing some effect between these two conditions. But at least to a first approximation, keyboard performances appear unaffected by the pedal auditory feedback.

Participants did not play one-fourth of the pedal events notated on the score. It is possible that performers prioritize their keyboard performance over the pedal but the statistical result cannot confirm this. However, some of the participants stated in the interviews to have neglected the pedal and prioritised the piano execution. The following qualitative comments each belong to a different participant: "I left out quite a lot of the pedal marking as I directed my attention to my piano playing", "I had to decide whether to press the pedal or not when I was late with one event, and there was a second close one", "I missed some of the pedal marking to maintain the fluency. I gave priority to the timing of the piano execution."

Data presented in table 6.3 shows that participants performed the pedal with a lower timing accuracy in the octave condition. The mean of the absolute values in the octave condition is lower by about 20 ms. Table 6.2 shows that these results are statistically significant but with a perhaps high p-value of $\approx$ 0.026.

I speculate that players were less accurate in the octave condition because they could easily ignore the sound resulting from the pedal interaction. The octave condition only affects future sounds, so there is no penalty for being inaccurate. In fact, being a few milliseconds earlier or late in the pedal execution won't be perceived in the performance. The sonic effect will happen for the note that is performed *following* the pedal tap. On the other hand, the immediate auditory feedback resulting from a pedal tap in the sound condition could be harder to ignore. As a result, participants may have used the cymbal sound to better line up the timing of their executions.

I also speculate that perceiving a sound (the ride cymbal) that is extraneous to the main instrument (the piano) may have polarised players' attention to the pedal, which in turn would have led to more accurate executions. By contrast, shifting the piano keyboard by one octave up or down keeps the resulting timbre somewhat coherent with the piano execution. Ignoring the piano transposition could be easier than ignoring the cymbal sound.

Still, such a weak statistical p-value suggests that there is not a massive difference in people's ability to play the pedal regardless of the condition that has been investigated. But it could also indicate that the octave condition impacted participants' time accuracy to a modest extent.

6.8 Conclusion

In this study, we looked at how well participants can retain their tempo skills in two conditions. The study is designed to challenge participants' ability to play the piano and the pedal in two experimental conditions. The evaluation is conducted by looking at tempo metrics.

Based on the available data, there does not appear to be any clear difference between the two conditions on these various metrics. Therefore it is concluded that the type of sonic feedback in the two conditions had a minimal effect on the performances. To play the piano in this study, it appears that it does not matter what kind of sonic feedback the pedal produces.

It is possible that people are prioritising the piano performance and what they do with the pedal is secondary. In fact, focusing on the familiar instrument to the detriment of new behaviours is also what the literature shows with other augmented instruments.

A more extensive study may lead to different conclusions.

Chapter 7

Conclusions

Mastering a musical instrument is a long-term commitment, often spanning several years. Professionals, having invested significant time in honing their skills, naturally aspire to operate at a high level, even when faced with unfamiliar or modified versions of familiar instruments. This research casts a spotlight on these encounters, aiming to uncover strategies that enable the retention of skills while also proposing methods to evaluate player responses in a systematic manner.

Each case study within the research was based on the deliberate choice to create artificial situations giving people new or unfamiliar instruments to play traditional repertoire. Corrective effects and mistakes were considered as data to test a series of the discussed research hypotheses. Each study aimed to create a laboratory to observe performers' behaviour using a modified or augmented musical instrument in the outlined musical context.

This research focused on the kind of situation where expert musicians play musical events notated on a pre-defined score controlling aspects such as timing, volume, timbre, accents, and articulation on a micro-scale and micro-time level. I am aware that a different set of musical values exists in communities like NIME. Those values include uncertainty, exploration, and the instrument as a co-creative agent [5, 2, 7]. I do not advocate that skills transfer should be a goal for all kinds of music contexts or practices.

The research questions in this thesis are summarised as follows:

- RQ1: How can designers leverage performers' existing skills to play something new or unfamiliar? How does the ability to imagine the target sound affect accuracy and response time compared to playing familiar sounds?

- RQ2: To what extent does unfamiliar auditory feedback affect performers' pitch accuracy and fluency?

- RQ3: Does performance improve when participants focus on their internal representation of sound instead of the instrument's auditory feedback?
- RQ4: How can we assess the transferability of existing motor skills to a modified or extended musical instrument? To what extent do players utilize their existing training?
- RQ5: Does transferring existing musical skills to a new musical interface require conscious attention, and does it impact other performance aspects?
- RQ6: How easily can a gesture producing an unfamiliar sound in real-time integrate into a music performance compared to a gesture modifying the sound in the near future?
- RQ7: How can we evaluate the integration of instrument modifications into players' execution?

The primary conclusions derived from this research are consolidated and presented in Table 7.1. The subsequent sections of this chapter delve deeper into these findings, offering a more comprehensive exploration of the results in relation to the research questions considered in this thesis.

Table 7.1: Main Findings

Finding	**Research Question**	**Outcome**	**Evidence**
Sensorimotor Imagery Salience for Skills Retention	RQ 1	Partially Unexpected	Substantial
Auditory Imagery Salience over Motor Imagery	RQ 1, 3	Unexpected	Substantial
Auditory Feedback Contribution to Performance	RQ 2	Expected	Substantial
New Quantitative Evaluation Method for Skills Transfer	RQ 4	NA	Strong

7.1 Shifting the design lens from auditory feedback to sensorimotor imagery

The research shows the significance of sensorimotor imagery in performing unfamiliar (or new) sounds featuring microtones material accurately. Findings from studies presented in chapters 3 and 4 have substantial evidence and suggest that imagining the sound and the associated motor actions to produce it could be beneficial for preserving professional players' pitch accuracy.

This insight questions conventional wisdom which typically leans towards a design strategy for skill transfer that is centred on technical elements or sound design in instrument creation. Thus, this thesis suggests a shift from a purely technology-focused design approach to one that is more human-centred. It advocates for designers to consider the sensorimotor link integral to musical performance, underscoring the potential advantage of a design perspective that values human aspects of musical interaction.

While critiques may emphasize the need for additional research to generalise its outcomes, they do not necessarily invalidate the study's broader applicability. Indeed, they emphasize the importance of context when interpreting the findings, underscoring the potential limitations of their universal applicability across all musical contexts.

Future studies should aim to further develop this kind of inquiry, examining how these findings translate to other instruments, genres, and populations, including left-handed musicians or those familiar with alternate tunings. Such research could refine the findings of this study, and potentially reveal further insights into the role of auditory imagery in music performance.

Even though the exact conclusions of the study may not be universally applicable, the broader notions it proposes about the significance of auditory imagery and its relationship with motor actions in music performance are considerably important. These insights contribute to a deeper understanding of the human elements involved in music performance and instrument design and can inform more effective and nuanced strategies for designing musical instruments. The study underscores the need to consider human-centric factors alongside technological ones, fostering a richer conversation in the domain of musical instrument design and performance. Although the research primarily focuses on violinists, its implications may extend to the design modifications of a diverse range of musical instruments. It's vital to acknowledge that despite the variation in physical mechanics across different instruments, the role of auditory imagery likely remains consistent.

Considering instruments such as trumpets or guitars, the physical act of playing a note may differ, but the role of auditory imagery remains salient. Musicians are required to internally envision the sound before producing it, which can significantly impact accuracy. Brass players, akin to violinists, may face challenges when aiming for microtones that are not typically utilized in Western music.

The importance of auditory imagery likely extends beyond playing quarter tones and may be relevant to a variety of musical tasks. This includes scenarios that require playing an instrument or a modified instrument that facilitates the production of new or unfamiliar sounds.

The findings presented here offer a wider perspective on the role of auditory imagery in music performance. This understanding could inform design modifications across a spectrum of musical instruments.

In both the quarter tones study 3 and the transposed violin study 4, we examined the degree to which appropriate auditory feedback and mapping could facilitate the performance of unfamiliar sounds on known or modified interfaces.

The results of these two studies align with initial expectations, confirming that merely using a familiar instrument equipped with accurate auditory feedback and mapping does not guarantee the preservation of pitch accuracy. The challenge became evident when performers were asked to produce quarter tones on their violins. They exhibited a reduced ability to adjust their pitch compared to their performance with semitones, particularly under notation conditions that lacked audio stimuli. These findings confirm the argument that an absence of sensorimotor imagery, encompassing auditory imagery, is a significant barrier when it comes to producing unfamiliar sounds on a familiar instrument, even for professional players.

In the transposed violin study 4, proper auditory feedback failed to assist violinists in maintaining their fluency. Even more intriguing was the finding that performers achieved smoother performances under the transposed notation condition, where the auditory feedback could be considered disruptive. These results imply that access to sensorimotor imagery significantly contributes to performance accuracy, even when playing a modified instrument. This principle seems applicable to performers on prepared piano: despite unfamiliar auditory feedback, a pianist may still perform a piece if it is notated as a piano composition.

The collective findings of these studies hint that the sound produced by the instrument, or even the interface's familiarity, is not the crucial factor in determining whether players can retain their skills. Instead, the ability to transfer skills to a modified version of an instrument seems largely contingent upon the design's facilitation of a connection to the player's sensorimotor imagery. That is, performers can maintain their pitch accuracy and fluency if they can mentally envision the sound they wish to produce and then activate the appropriate motor program to achieve it.

It is also proposed that auditory imagery might be more impactful than motor imagery in accurately playing new sounds. This infers that the ability to mentally project the desired auditory output could possibly be a more influential factor than the ability to envision the required motor actions. This finding is unexpected and based on speculations built upon the results from the quarter tones study 3. It needs further research to provide strong evidence.

Nonetheless, deficiency in sensorimotor imagery signifies a mental barrier. If players are unable to imagine the sounds they need to perform, they cannot link to an effective motor program to produce them. This outcome challenges the traditional view of musical performance as a feedback loop in which performers contemplate their desired action, execute it, and adjust based on the feedback. As shown in the transposed notation condition, musicians can still perform even when their sensorimotor imagery does not align precisely with the auditory feedback, provided their anticipation guides them to the correct motor program.

Therefore, to design instrument modifications that enable the performance of novel sounds using pre-existing skills, I propose a shift from a technology-centric viewpoint to a more human-based one. Beyond addressing the technological aspects like mappings, sensor precision, interface ergonomics, and various forms of assistive feedback, designers could focus on the sensorimotor link integral to musical performance. This link constitutes part of a feedforward mechanism, where players anticipate the sound they aim to produce and how to execute it mentally.

This approach deviates from mere questions of whether an instrument can produce microtones or if it is ergonomically suited to the player's hands. Incorporating the feedforward process into the design process leads to a critical question: does an intrinsic connection exist between a specific interface being designed and the existing mental imagery of performers who will use the interface, along with the techniques they already possess? If the answer is no, then the instrument confronts a fundamental mental obstacle, rather than a technological one. Thus, if a designer's objective is to enable performers to produce unfamiliar sounds, this human element needs to be taken into account.

The findings from these two studies extend our understanding of auditory imagery's role in music performance. They could inform the process of modifying or augmenting musical instruments.

7.2 Auditory feedback in the design of new or modified instruments

This research acknowledges the potential contribution of auditory feedback to the performer's experience of the instrument, suggesting that it could play a role in refining a musical performance. This is an expected finding that has substantial evidence. While sensorimotor imagery is highlighted, the research also suggests a comprehensive design approach that encompasses auditory feedback and ergonomics.

Auditory feedback has indeed importance in the design of new or modified instruments. Modifying a traditional instrument can lead to a new design that produces novel sounds. **RQ2** asked *to what extent does performers' pitch accuracy and fluency deteriorate in the presence of unfamiliar auditory feedback?* **RQ3** asked *does players' performance improve if the task allows participants to ignore the sound coming from the instrument and to focus on their internal representation of sound (i.e. their expectations based on the sound the instrument used to reproduce)?* In the first and second pitch studies, we saw that auditory feedback is not salient in determining whether players' can retain certain aspects of their performance on a modified violin. What seems to be important is their ability to imagine the sound they need to perform and the presence of pre-existing motor programs that can be used to play it. Players confronted with an unfamiliar interface (the transposed violin) and mismatching auditory feedback (transposed notation condition) managed to retain their fluency by ignoring the auditory feedback produced by the instrument.

However, I do not claim that auditory feedback should be disregarded in the design process. Auditory feedback can matter for reasons related to how performers experience the instrument. This has been demonstrated in research studies and publications [97, 98]. Enactive research generally discusses how it is possible to develop a methodology for sound control based on commonalities between objects being played and the consequent sound and the incorporation of the instrument into body schemas [99, 100]. Perceptually guided action defines the 'feel' and playability of a musical instrument [99]. Cadoz argues there is a natural relationship between the dynamic of the input and the instrument output. What makes an instrument playable is a transfer from action to sound [101]. Being a property of sound and instrument design, auditory feedback does matter. Moreover, auditory feedback produced by the instrument could also be useful for adjustments during players' performances. In the Magpick study, we saw unity in what performers imagined and played. When the strength of the effect gradually changed, players acted to compensate for it. I speculate that this happened as they were listening at some level to the Magpick sound and adapting their playing to it. Auditory feedback may be needed for correction and refinement during a musical performance and should be considered for aesthetic and cultural reasons.

Nevertheless, to support a performance guided by sensorimotor imagination (i.e. execution with music scores) would entail designing an instrument that allows a translation from the imagery of a sound to an action that is harmonious and appropriate to the instrument. If musicians do not have that connection (transposed violin study, concert notation), they will also lack the skills to perform with the instrument or its modification.

7.3 Evaluating skills transfer in new or modified musical instruments

The case study presented in chapter 5 introduces a new quantitative evaluation method aimed at investigating skill transfer on modified or augmented instruments. While more tests may be needed, this method provides a potential direction to understand how musicians may adapt to new or modified instrument designs by actively listening and adjusting their play style based on the feedback. The case study presented in chapter 6 shows the first steps toward a further quantitative methodology that aims to evaluate music instrument modifications designed for skills transfer.

RQ4 asked *how do we evaluate how far existing motor skills can transfer to a modified or extended musical instrument? How do we know to what extent players make use of existing training?* The methodology presented in the Magpick study addresses this question study by proposing a quantitative method to query players' responses in the interaction with a modified or new instrument. The first contribution of the methodology presented in the study is that it provides a numerical specificity that describes how a modified or new instrument design allows players to use their skills to control it successfully.

Evaluating whether players actively embrace and control a design can be difficult. It may be especially challenging to reply to the question are players listening to the instrument modification? The method uses simple correlations to query whether players actively listen to the auditory result of interacting with the modified plectrum and adjusting their playing accordingly or whether they are simply replicating motor programs they already own by using the plectrum as they would normally do if it was unmodified.

The method uses mapping and mapping disruptions to query whether players are listening to the auditory response to their control of the instrument modification. It does so in a novel way by introducing subtle changes to the response of the modification. Mapping is sometimes valued as a way to free instruments from the constraints of acoustic systems [102]. I propose that mapping has an additional useful function for studying performers' interactions. Changes are periodic and designed to be unnoticed by players so that players do not consciously know that any change in the modification response to their playing is taking place. Perturbing the relation between action and sound allows us to look for a compensating effect in the playing actions that control the modified instrument. Such compensating effects signal whether players are listening to the instrument modification response to their playing gestures and actively controlling the modification to achieve certain sonic results.

The method may also be beneficial for studies taking place in a musical context in the wild. The fact that it is possible to change the mapping in real-time while execution is taking place and to change it without the performer knowing that it is changing means that researchers have a way to study the relation between action and sound by observing players' reactions. Because the perturbation is designed to be latent and subtle, it does not become the focus of people's attention. Being designed not consciously to distract players from their primary activity: playing music, the disruption becomes possibly compatible with performing in a realistic context such as a rehearsal room or a concert. A latent perturbation of the system possibly allows players to stay in their embodied motor thinking as they would normally do in a naturalistic musical context.

RQ5 asked *provided that players succeed in transferring some of their existing skills to play a new musical interface, does the skills transfer happen on a conscious basis (involving deliberate attention) to the cost of neglecting other aspects of their performance (i.e. articulation, fluency)?* The results from the case study appear to show at least a modest subconscious response to changes in augmentation behaviour. At the end of the experiment, players were asked if they had noticed any change in the Magpick behaviour. None of them reported having experienced a difference in the augmented pick sensitivity. I speculate that participants were not only listening to the sonic augmentation and reacted to the LFO but also that their reaction was unconscious as they did not report its effect. Adapting their playing by adjusting their picking gestures became an automatic subconscious action, possibly similar to the act of placing their finger on the fretboard or plucking the strings.

7.4 Evaluating the integration of modified instruments in players' performance

A further research goal tackled in this thesis is evaluating the integration of modified instruments in players' performance. A designer may decide to extend an existing musical instrument to *instantly* produce *a new sound* in response to a gesture. Otherwise, they may decide that a gesture affects the *subsequent sounds* produced by the *underlying original instrument.* **RQ6** asked: *does a gesture producing an unfamiliar sound in real-time easily integrate into a music performance compared to a gesture modifying the sound of the instrument in the near future?* The question was addressed in the fourth study presented in this thesis and did not lead to any significant findings. However, I propose the study could be the first step toward a methodology to address **RQ7**: *how to evaluate whether the modification of an instrument can be integrated into players'* *execution.* The evaluation is achieved by observing the degree of separation between the stream of interaction on the instrument modification and the stream of interaction on the underlying original instrument. The degree of separation is identified by looking at measurable results in players' performances, such as timing accuracy. The method queries whether different events distributed between the two aspects of the instrument can enter and integrate into the sonic imagination of the performer. The method needs further work and perhaps a different approach that finds a condition that would highlight the integration (or the lack of integration) of the gestures performed on the instrument modification and the gestures performed on the underlying traditional instrument.

7.5 Future challenges

While providing practical guidance for the design of modified instruments was beyond the scope of this thesis, the findings from the quarter tones and transposed violin studies offer valuable principles to inform practical design endeavours. These principles underscored the significance of performers' ability to imagine the desired sound and activate corresponding motor programs to retain performance fluency and accuracy. While this thesis does not cover practical design guidelines, it provides practical guidance for the evaluation of modified or augmented instruments. The evaluation method introduced in the Magpick study serves as a practical tool to assess the alignment of a design with these principles and evaluate its effectiveness in enabling players to utilize their skills within the musical context outlined in this research. Exploring the translation of these principles into practical tasks represents a direction for future research.

Outcomes from the case studies presented in this thesis offer additional questions that could represent future directions for this research. These questions address two focal points of the research: the influence of auditory feedback and instrument modification, and the evaluation method presented in chapter 5.8.

Results from the first and second studies suggest that players overcame a mismatch between sensorimotor imagery and auditory feedback by ignoring the sound coming from the instrument. However, were they really ignoring the sound of their performance to its full extent (pitch, timing, articulation)? Were they ignoring certain aspects (i.e. pitch) while subconsciously considering others? Which of these aspects possibly helped integrate the modified violin into their performances? Additional research may focus on evaluating the influence of auditory feedback even when players decide to ignore the sound of their performance[1]. Further research may as well confirm the generalisation of the two

[1]More information on this subject was provided in section 4.3

studies' findings for different instruments.

The evaluation method presented in the Magpick study, section 5, relied on participant self-reports to determine whether participants consciously or unconsciously reacted to the LFO. Choosing whether the players' response to the disruption is conscious or unconscious is critical in determining the stage of instrument motor learning experienced by players. For this reason, additional research strategies are needed to improve the methodology and to reinforce the hypothesis that performers not only respond to the disruption and adapt their playing but also that they do so unconsciously as a result of responding to the auditory feedback of the augmentation.

The work presented in this thesis offers a possible guide to designers who want to make further instruments which build on skills that already exist. Successful examples of instrument designs that transfer existing skills, like the Moog synthetic keyboard or the electric guitar, suggest that this goal is achievable. Designers who want to transfer players' skills to a modified interface may shift the design lens primarily to sensorimotor imagery. The evaluation method described in section 5 offers a quantitative approach to design evaluation in the musical context considered in this thesis. Such methodology may integrate existing qualitative methods and could inform whether a designer is on the right track while modifying or creating a new interface that aims to extend professional players' skills.

Bibliography

[1] Fabio Morreale and Andrew McPherson. Design for longevity: ongoing use of instruments from nime 2010-14. In Cumhur Erkut, editor, *17th International Conference on New Interfaces for Musical Expression, NIME 2017, Aalborg University, Copenhagen, Denmark, May 15-18, 2017*, pages 192–197. nime.org, 2017.

[2] Fabio Morreale, Andrew McPherson, and Marcelo M. Wanderley. NIME identity from the performer's perspective. In *18th International Conference on New Interfaces for Musical Expression, NIME 2018, Blacksburg, VA, USA, June 3-6, 2018*, pages 168–173. nime.org, 2018.

[3] Daniel J Levitin, Stephen McAdams, and Robert L Adams. Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(2):171–189, 2002.

[4] Tod Machover. Hyperinstrument: Musically intelligent and interactive performance and creativity systems. *Proceedings of 1989 ICMC*, pages 186–190, 1989.

[5] Michael Gurevich and Jeffrey Treviño. Expression and its discontents: toward an ecology of musical creation. In *Proceedings of the 7th international conference on New interfaces for musical expression*, pages 106–111, 2007.

[6] Tom Davis. The feral cello: A philosophically informed approach to an actuated instrument. In *NIME*, 2017.

[7] Halldór Úlfarsson. The halldorophone: The ongoing innovation of a cello-like drone instrument. In *NIME*, pages 269–274, 2018.

[8] Gallwey Green, B. *The inner game of music.* Pan Macmillan, 1987.

[9] Peter E Keller. Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, 1252(1):206–213, 2012.

[10] Richard A Schmidt, Timothy D Lee, Carolee Winstein, Gabriele Wulf, and Howard N Zelaznik. *Motor control and learning: A behavioral emphasis*. Human kinetics, 2018.

[11] David Norton. Not pulling strings by joseph o'conner. new maiden, surrey: Lambent books, 1987.£ 6.95, 168 pp. *British Journal of Music Education*, 6(1):107–108, 1989.

[12] Marrin Nakra. *Inside the Conductor's Jacket: Analysis, interpretation and musical synthesis of expressive gesture*. PhD thesis, Massachusetts Institute of Technology, 2000.

[13] Marc Leman, Micheline Lesaffre, Luc Nijs, and Alexander Deweppe. User-oriented studies in embodied music cognition research. *Musicae Scientiae*, 14(2_suppl):203–223, 2010.

[14] Peter Schwenkreis, Susan El Tom, Patrick Ragert, Burkhard Pleger, Martin Tegenthoff, and Hubert R Dinse. Assessment of sensorimotor cortical representation asymmetries and motor skills in violin players. *European Journal of Neuroscience*, 26(11):3291–3302, 2007.

[15] Fabio Morreale, Jack Armitage, and Andrew McPherson. Effect of instrument structure alterations on violin performance. *Frontiers in psychology*, 9:2436, 2018.

[16] Andrea Guidi, Fabio Morreale, and Andrew McPherson. Design for auditory imagery: altering instruments to explore performer fluency. In *20th International Conference on New Interfaces for Musical Expression, NIME 2020, Royal Birmingham Conservatoire, UK, 2020*, 2020.

[17] Dan Overholt. The overtone violin. In *Proceedings of the 2005 conference on New interfaces for musical expression*, pages 34–37. Citeseer, 2005.

[18] Andrew P. McPherson. Touchkeys: Capacitive multi-touch sensing on a physical keyboard. In Georg Essl, R. Brent Gillespie, Michael Gurevich, and Sile O'Modhrain, editors, *12th International Conference on New Interfaces for Musical Expression, NIME 2012, Ann Arbor, Michigan, USA, May 21-23, 2012*. nime.org, 2012.

[19] Joseph Thibodeau and Marcelo M Wanderley. Trumpet augmentation and technological symbiosis. *Computer Music Journal*, 37(3):12–25, 2013.

[20] Daniel Overholt. The overtone fiddle: an actuated acoustic instrument. In *New Interfaces for Musical Expression*, pages 4–7. University of Oslo, 2011.

[21] Robert J Zatorre, Joyce L Chen, and Virginia B Penhune. When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience*, 8(7):547–558, 2007.

[22] Eckart Altenmüller. Neurology of musical performance. *Clinical medicine*, 8(4):410, 2008.

[23] Sarah-Indriyati Hardjowirogo. Instrumentality. on the construction of instrumental identity. In *Musical instruments in the 21st century*, pages 9–24. Springer, 2017.

[24] Maurice Merleau-Pointy. *Phenoménologie de la perception*. Gallimard, Reading, Massachusetts, 1976.

[25] Edmund Husserl. *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy: Second book studies in the phenomenology of constitution*, volume 3. Springer Science & Business Media, 1989.

[26] Silvano Zipoli Caiani. The ecological meaning of embodiment. *Phenomenology and mind*, (1):132–138, 2011.

[27] Shaun Gallagher. Dynamic models of body schematic processes. *Body image and body schema*, pages 233–250, 2005.

[28] Thor Magnusson. Of epistemic tools: Musical instruments as cognitive extensions. *Organised Sound*, 14(2):168–176, 2009.

[29] Don Ihde. *Technology and the lifeworld: From garden to earth*. PhilPapers, 1990.

[30] Don Ihde. *Technics and praxis: A philosophy of technology*, volume 24. Springer Science Business Media, 2012.

[31] Luc Nijs. The merging of musician and musical instrument: Incorporation, presence, and levels of embodiment. In *The Routledge companion to embodied music interaction*, pages 49–57. Routledge, 2017.

[32] Martin Heidegger. *Being and Time*. Blackwell, 1962.

[33] Shaun Gallagher and Dan Zahavi. *The phenomenological mind*. Routledge, 2020.

[34] Jessica K Witt. Action's effect on perception. *Current Directions in Psychological Science*, 20(3):201–206, 2011.

[35] Lawrence Shapiro. *Embodied cognition*. Routledge, 2010.

[36] Thomas Fuchs. Intercorporality–how embodied interaction shapes mind and brain. *Sabine Flach, Daniel Margulies und Jan Söffner: Habitus in Habitat I–Emotion and Motion, Bern ua, Peter Lang*, pages 55–67, 2010.

[37] Marc Leman. *Embodied music cognition and mediation technology.* MIT press, 2007.

[38] Marc Leman and Pieter-Jan Maes. The role of embodiment in the perception of music. *Empirical Musicology Review*, 9(3-4):236–246, 2015.

[39] Marc Leman, Luc Nijs, and Nicola Di Stefano. On the role of the hand in the expression of music. In *The Hand*, pages 175–192. Springer, 2017.

[40] Luc Nijs. Incorporation, presence, and levels of embodiment. *The Routledge companion to embodied music interaction*, page 49, 2017.

[41] Marc Leman, Pieter-Jan Maes, Luc Nijs, and Edith Van Dyck. What is embodied music cognition? In *Springer handbook of systematic musicology*, pages 747–760. Springer, 2018.

[42] Koray Tahıroğlu, Thor Magnusson, Adam Parkinson, Iris Garrelfs, and Atau Tanaka. Digital musical instruments as probes: How computation changes the mode-of-being of musical instruments. *Organised Sound*, 25(1):64–74, 2020.

[43] Don Ihde. *Hermeneutic Phenomenology: The Philosopher of Paul Ricoeur.* Northwestern University Press, 1971.

[44] Giuseppe Riva. Being-in-the-world-with: Presence meets social and cognitive neuroscience. In *From communication to presence: Cognition, emotions, culture towards the ultimate communicative experience*, pages 47–80. IOS Press, 2006.

[45] Thor Magnusson. The migration of musical instruments: On the sociotechnological conditions of musical evolution. *Journal of New Music Research*, 50(2):175–183, 2021.

[46] Andrew Feenberg et al. *Critical theory of technology*, volume 5. Oxford University Press New York, 1991.

[47] Don Ihde and Craig Hanks. A phenomenology of technics. In *Technology and values: Essential readings*, pages 134–155. Wiley-Blackwell Chichester, 2010.

[48] María Herrojo Ruiz, Hans-Christian Jabusch, and Eckart Altenmüller. Fast feedforward error-detection mechanisms in highly skilled music performance. In *Proceedings of the International Symposium on Performance Science*, pages 187–197. Citeseer, 2009.

[49] Daniel M Wolpert, Zoubin Ghahramani, and Michael I Jordan. An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882, 1995.

[50] Rolf Inge Godøy. Gestural imagery in the service of musical imagery. In *International Gesture Workshop*, pages 55–62. Springer, 2003.

[51] Terry Clark, Aaron Williamon, and Aleksandar Aksentijevic. Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. *Musical imaginations: Multidisciplinary perspectives on creativity, performance, and perception*, pages 351–365, 2012.

[52] Allan Paivio. Cognitive and motivational functions of imagery in human performance. *Canadian journal of applied sport sciences. Journal canadien des sciences appliquées au sport*, 10(4):22S–28S, 1985.

[53] Zebulon Highben and Caroline Palmer. Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education*, pages 58–65, 2004.

[54] Rachel M Brown and Caroline Palmer. Auditory–motor learning influences auditory memory for music. *Memory & Cognition*, 40(4):567–578, 2012.

[55] Laura Bishop, Freya Bailes, and Roger T Dean. Musical imagery and the planning of dynamics and articulation during performance. *Music Perception: An Interdisciplinary Journal*, 31(2):97–117, 2012.

[56] Laura Bishop, Freya Bailes, and Roger T Dean. Musical expertise and the ability to imagine loudness. *PloS one*, 8(2):e56052, 2013.

[57] Timothy L Hubbard. Some anticipatory, kinesthetic, and dynamic aspects of auditory imagery. *The Oxford Handbook of Sound and Imagination, Volume 1*, page 149, 2019.

[58] Fernando Martínez-Manrique and Agustín Vicente. The activity view of inner speech. *Frontiers in psychology*, 6:232, 2015.

[59] Andrea Halpern. Dynamic aspects of musical imagery. *Annals of the New York Academy of Sciences*, 1252:200–5, 04 2012.

[60] A Halpern. Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, 2015.

[61] Michel Denis. *Image et cognition*. Presses universitaires de France Paris, 1989.

[62] Robert J Zatorre, Andrea R Halpern, and Marc Bouffard. Mental reversal of imagined melodies: a role for the posterior parietal cortex. *Journal of cognitive neuroscience*, 22(4):775–789, 2010.

[63] James M Jordan. Audiation and sequencing: An approach to score preparation. *The Choral Journal*, 21(8):11–13, 1981.

[64] Ian D Colley, Peter E Keller, and Andrea R Halpern. Working memory and auditory imagery predict sensorimotor synchronisation with expressively timed music. *Quarterly Journal of Experimental Psychology*, 71(8):1781–1796, 2018.

[65] Emma B Greenspon, Peter Q Pfordresher, and Andrea R Halpern. Pitch imitation ability in mental transformations of melodies. *Music Perception: An Interdisciplinary Journal*, 34(5):585–604, 2017.

[66] Patricia Holmes. Imagination in practice: A study of the integrated roles of interpretation, imagery and technique in the learning and memorisation processes of two experienced solo performers. *British Journal of Music Education*, 22(3):217–235, 2005.

[67] Albrecht Schneider and Rolf Inge Godøy. Perspectives and challenges of musical imagery. *Musical imagery*, pages 5–26, 2001.

[68] Emmanuel Mellet, Laurent Petit, Bernard Mazoyer, Michael Denis, and Natalie Tzourio. Reopening the mental imagery debate: lessons from functional anatomy. *Neuroimage*, 8(2):129–139, 1998.

[69] Stephen M Kosslyn, Giorgio Ganis, and William L Thompson. Neural foundations of imagery. *Nature reviews neuroscience*, 2(9):635–642, 2001.

[70] Christopher Connolly and Aaron Williamon. Mental skills training. *Musical excellence: Strategies and techniques to enhance performance*, 5(1):221–245, 2004.

[71] Sibylle C Herholz, Andrea R Halpern, and Robert J Zatorre. Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of cognitive neuroscience*, 24(6):1382–1397, 2012.

[72] María Herrojo Ruiz, Hans-Christian Jabusch, and Eckart Altenmüller. Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, 19(11):2625–2639, 2009.

[73] Veerle L Simoens and Mari Tervaniemi. Auditory short-term memory activation during score reading. *PLoS One*, 8(1):e53691, 2013.

[74] William L Berz. Working memory in music: A theoretical model. *Music Perception*, 12(3):353–364, 1995.

[75] Daniele Schön and Mireille Besson. Visually induced auditory expectancy in music reading: a behavioral and electrophysiological study. *Journal of cognitive neuroscience*, 17(4):694–705, 2005.

[76] David JM Kraemer, C Neil Macrae, Adam E Green, and William M Kelley. Sound of silence activates auditory cortex. *Nature*, 434(7030):158–158, 2005.

[77] Guy Weir, Victoria J Williamson, and Daniel Müllensiefen. Increased involuntary musical mental activity is not associated with more accurate voluntary musical imagery. *Psychomusicology: Music, Mind, and Brain*, 25(1):48, 2015.

[78] Amber M Leaver, Jennifer Van Lare, Brandon Zielinski, Andrea R Halpern, and Josef P Rauschecker. Brain activation during anticipation of sound sequences. *Journal of Neuroscience*, 29(8):2477–2485, 2009.

[79] Xing Tian and David Poeppel. The effect of imagination on stimulation: the functional specificity of efference copies in speech processing. *Journal of cognitive neuroscience*, 25(7):1020–1036, 2013.

[80] Christian Hoppe, Christoph Splittstößer, Klaus Fliessbach, Peter Trautner, Christian E Elger, and Bernd Weber. Silent music reading: Auditory imagery and visuotonal modality transfer in singers and non-singers. *Brain and cognition*, 91:35–44, 2014.

[81] Mihaly Csikszentmihalyi and Jeremy Hunter. Happiness in everyday life: The uses of experience sampling. In *Flow and the foundations of positive psychology*, pages 89–101. Springer, 2014.

[82] Fabio Morreale, Andrea Guidi, and Andrew McPherson. Magpick: an augmented guitar pick for nuanced control. In *19th International Conference on New Interfaces for Musical Expression, NIME 2019, Porto Alegre, Brazil, June 3-6, 2019*, pages 65–70. nime.org, 2019.

[83] Casey Reas and Ben Fry. *Processing: a programming handbook for visual designers and artists*. MIT Press, 2007.

[84] Scott Wilson, David Cottle, and Nick Collins. *The SuperCollider Book*. The MIT Press, 2011.

[85] JP Gather. *Amsterdam Catalog of Csound Computer Instruments*. Buffalo, 1995.

[86] Andrew McPherson and Victor Zappi. An environment for submillisecond-latency audio and sensor processing on BeagleBone Black. In *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[87] Jeffrey S Racine. Rstudio: a platform-independent ide for r and sweave, 2012.

[88] Chris Cannam, Christian Landone, and Mark Sandler. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1467–1468, 2010.

[89] Matthias Mauch and Simon Dixon. pYin: A fundamental frequency estimator using probabilistic threshold distributions. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 659–663. IEEE, 2014.

[90] Maike Watson. Musescore. *Journal of the Musical Arts in Africa*, 15(1-2):143–147, 2018.

[91] Matthias Mauch, Chris Cannam, Rachel Bittner, George Fazekas, Justin Salamon, Jiajie Dai, Juan Bello, and Simon Dixon. Computer-aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency. In *Proceedings of the First International Conference on Technologies for Music Notation and Representation*, 2015.

[92] Andrea Guidi and Andrew McPherson. Quantitative evaluation of aspects of embodiment in new digital musical instruments. In *22th International Conference on New Interfaces for Musical Expression, NIME 2022, The University of Auckland, New Zealand, 2022*, 2022.

[93] Perry R. Cook. Principles for designing computer music controllers. In Ivan Poupyrev, Michael J. Lyons, Sidney S. Fels, and Tina Blaine, editors, *New Interfaces for Musical Expression, NIME-01, Proceedings, Seattle, WA, USA, April 1-2, 2001*, pages 3–6, 2001.

[94] M. Martin. *Audacity 2*. Mediaforma, 2014.

[95] Gracie U Ibemere. *Interactions between Text and Music in Three Pieces*. PhD thesis, Massachusetts Institute of Technology, 2018.

[96] Colin Raffel and Daniel PW Ellis. Intuitive analysis, creation and manipulation of midi data with pretty_midi. In *15th international society for music information retrieval conference late breaking and demo papers*, pages 84–93, 2014.

[97] Indiana Wollman, Claudia Fritz, Jacques Poitevineau, and Stephen McAdams. Investigating the role of auditory and tactile modalities in violin quality evaluation. *PLOS ONE*, 9:1–27, 12 2014.

[98] Charalampos Saitis, Claudia Fritz, and Gary Scavone. Sounds like melted chocolate: How musicians conceptualize violin sound richness. In *International Symposium on Musical Acoustics*, 2019.

[99] Georg Essl and Sile O'modhrain. An enactive approach to the design of new tangible musical instruments. *Organised sound*, 11(3):285–296, 2006.

[100] Francisco J Varela, Evan Thompson, and Eleanor Rosch. *The embodied mind, revised edition: Cognitive science and human experience*. MIT press, 2017.

[101] Claude Cadoz. Instrumental gesture and musical composition. In *ICMC 1988-International Computer Music Conference*, pages 1–12, 1988.

[102] Doug Van Nort and Nicolas Castagné. *Mapping, in digital musical instruments*. Enactive Systems Books, 2007.

Appendix A

Notation Stimuli Questionnaires and Interviews

This appendix contains interview questions, and questionnaires, used in this research. The following link gives the notation stimuli used in each study: `https://www.dropbox.com/sh/8066228w6lz4wd2/AACTtJnj_lprD2JMUODCaAWMa?dl=0s`.

A.1 Interviews

The following sections present the list of pre-defined questions used in the semi-structured interview for each study (chapters 3, 4, 5, 6).

A.1.1 First Study Interview

- What aspect of these tasks did you find the most difficult, and why?
- Was the visual or the auditory tasks more difficult, and why?
- Which strategies did you take to overcome these difficulties?
- When playing a quarter tone from notation, how did you know when you were in tune?
- Did you notice any difficulty difference between the first and second set of notation stimuli?

A.1.2 Second Study Interview

- Which of the three forms of presentation (concert notation, transposed notation, audio stimuli) was more natural and harder and why?
- Did you feel that the last musical excerpts of each section (known music passages) were more accessible and why?
- How did you know where to place your finger on the transposed violin while reading concert notation?

A.1.3 Third Study Interview

- What aspect of these tasks did you find the most difficult, and why?
- Which section of the study was easier and why?
- Did you notice a change in the behaviour of the pick during the study?
- Did you change the way you were playing in the different conditions?
- Which strategies did you take to match the target stimuli?
- When trying to match the stimuli, how did you know you were succeeding?
- How did you know how much energy to put into the picking gestures to match the target stimuli?

A.1.4 Fourth Study Interview

- What aspect of these tasks did you find the most difficult, and why?
- Which section of the study was easier and why?
- Did you change the way you were playing in the different conditions?
- Which strategies did you take to tackle the increase in complexity of the tasks in each section of the study?

A.2 Questionnaires

Fig. A.1 shows the questionnaire used for the first and second studies described in chapters 3 and 4. Each session of two studies occurred on the same day with the same participant. This allowed using one questionnaire for the two experiments. Fig. A.2 shows the questionnaire used for the third study described in sections 5. Fig. A.3 shows the questionnaire used for the fourth study described in sections 6.

Figure A.1: questionnaire for the quarter tones and transposed violin studies; sections 3 and 4.

QUESTIONNAIRE

- Age ________________________

- Gender
  - |_| male
  - |_| female
  - |_| non-binary

- How long have you being studying violin? ________________________

- Do you have
  - |_| perfect pitch
  - |_| relative pitch

- Do you play any other instrument? If so, how many years did you study it?

  ____

- Do you have any experience with playing and/or listening music with quartertones or micro intervals?

  ____

  ____

- Do you have any experience playing with scordatura (altered tunings)?

  ____

  ____

- What style(s) of music do you primarily play, and what are some examples of recent pieces you have played?

  ____

  ____

Figure A.2: questionnaire for the Magpick tones study; section 5

QUESTIONNAIRE

- Age ________________________

- Gender
  - |_| male
  - |_| female
  - |_| non-binary

- How long have you being studying electric guitar?

  ____

- Do you play any other instrument? If so, how many years did you study it?

  ____
  ____
  ____

- Do you have any experience with playing or listening music with guitar synthesizers or heavy modified electric guitar sounds?

  ____
  ____
  ____

- What style(s) of music do you primarily play, and what are some examples of recent pieces you have played?

  ____
  ____
  ____

Figure A.3: questionnaire for the piano-pedal study; section 6

QUESTIONNAIRE

- Age:
- Gender
  - |_| male
  - |_| female
  - |_| non-binary
  - |_| prefer not to say

- How long have you being studying keyboard or piano?

- Do you play any other instrument? If so, how many years did you study it?

- Do you have any experience with pedals?