Article

Musical Control Gestures in Mobile Handheld Devices: Design Guidelines Informed by Daily User Experience

Alexandre Clément [1,*] **, Luciano Moreira** [1,2] **, Miriam Rosa** [3] **and Gilberto Bernardes** [1,4]

1 Faculdade de Engenharia (FEUP), Universidade do Porto, 4200-465 Porto, Portugal; lucianomoreira@fe.up.pt (L.M.); gba@fe.up.pt (G.B.)
2 Centro de Investigação em Química da Universidade do Porto (CIQUP), 4169-007 Porto, Portugal
3 Instituto Universitário de Lisboa (Iscte-IUL), CIS-IUL, 1649-026 Lisboa, Portugal; miriam.rosa@iscte-iul.pt
4 Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência (Inesc-Tec), 4200-465 Porto, Portugal
* Correspondence: a.clement@fe.up.pt

Abstract: Mobile handheld devices, such as smartphones and tablets, have become some of the most prominent ubiquitous terminals within the information and communication technology landscape. Their transformative power within the digital music domain changed the music ecosystem from production to distribution and consumption. Of interest here is the ever-expanding number of mobile music applications. Despite their growing popularity, their design in terms of interaction perception and control is highly arbitrary. It remains poorly addressed in related literature and lacks a clear, systematized approach. In this context, our paper aims to provide the first steps towards defining guidelines for optimal sonic interaction design practices in mobile music applications. Our design approach is informed by user data in appropriating mobile handheld devices. We conducted an experiment to learn links between control gestures and musical parameters, such as pitch, duration, and amplitude. A twofold action—reflection protocol and tool-set for evaluating the aforementioned links—are also proposed. The results collected from the experiment show statistically significant trends in pitch and duration control gesture mappings. On the other hand, amplitude appears to elicit a more diverse mapping approach, showing no definitive trend in this experiment.

Keywords: digital musical instruments; interaction; instrument control; musical gesture; smartphone; human computer interaction

1. Introduction

Mobile phones have been on the radar of digital music since 2002 when the audience members' Nokia phones at an Ars Electronica Festival were used to create a collaborative musical piece [1]. With the emergence of smartphones and other handheld smart-devices, the hardware, processing power, and inbuilt sensor availability have evolved exponentially [2]. Smartphone spread throughout the population is also steadily growing—around 76% of adults own a smartphone in advanced economies, and 45% do in emerging economies, amounting to over 3 billion users worldwide (3.2 billion in 2019 with a projected 3.8 billion in 2021) [3]. Mobile Handheld Device (MHD) portability, availability, and simplicity of operation have given them a quasi-prosthetic role in our lives. The wide availability and portability of MHDs have made them widely adopted interfaces for musical expression. The creation of libraries that port popular audio development engines (e.g., libPD [4] and SuperCollider-Android (https://github.com/glastonbridge/SuperCollider-Android, accessed on 5 January 2021)) onto mobile systems, and the development of tools allowing for integration with already established digital music software (e.g., Apple Garageband for iOS (https://www.apple.com/ios/garageband/, accessed on 2 February 2021) and Steinberg Cubasis (https://new.steinberg.net/cubasis/, accessed on 2 February 2021)), have fostered musical creation on MHDs even further.

MHD musical tools and software exhibit different gesture mapping and interaction methods, anchored in different objectives and application types, as we shall further discuss in Section 2.

Prior experience in developing collaborative (audience–performer), interactive music systems [5,6] has suggested idiosyncratic appropriations of MHD Digital Musical Instruments (DMIs). In the aforementioned project, composers approached the devices' interaction control and feedback capabilities from multiple perspectives. Informal feedback from participants in the audience included difficulties in understanding the interaction control of the MHD musical interface. There was little awareness of the control gestures and actions available or their musical feedback. Regarding the composers, they were faced with designing the sonic interaction design methods for their works from scratch without undergoing a training or explanation phase before the performance. In summary, the lack of a common approach to the MHD DMI control and behavior was the major frustration among audience members. As very incisively stated and explored by the authors in [7], these devices' specificities have seldom been taken into account while developing control mechanisms for MHD musical instruments.

The experiment described in this article is relevant for two major fields of study: gesture-to-sound (specifically for digital musical instruments) and mobile gesture analysis for interaction control. Both have been widely studied independently, but no attempt to evaluate them in conjunction exists. The multitude of mapping approaches from self-contained applications and the open-ended nature of the embeddable libraries and sensor communication applications highlight this work's premise: the need for systematized guidelines concerning musical interaction design for mobile applications. This diversity in interaction and design approaches is very evident in the reviews by Essl [8] and Turchet [9].

Building upon the growing research on gestural affordances, gesture meaning, and device usage in a musical context [10–12], we aimed to denote guidelines for the instinctive and fluid appropriation of MHDs as musical instruments by attempting to establish instinctive relations among interaction control gestures (e.g., touch manipulations, movement manipulations) and musical parameters (e.g., note onset, pitch, and duration). We conducted an experiment to understand how musically proficient users and lay users map interaction control gestures and musical parameters on MHDs. Ultimately, we aimed to understand what gestures are most commonly adopted in controlling sound parameters and define guidelines for the appropriation of MHDs as digital musical instruments. In the absence of a standard for this evaluation, we also propose an easily implementable and expandable protocol with specific tools, test scripts, and analysis methods.

The remainder of this article is structured as follows. In Section 2, we contextualize the work in terms of base premises and definitions. In Section 3, we detail the experiment, the data collection process, and the analytical process. In Section 4, we provide the results stemming from the experimental data. In Section 5, we interpret and discuss the experimental results, explain the reasoning leading to our proposed guidelines, and identify shortcomings and avenues for future work. Finally, in Section 6, we summarize our study's main contributions and the implications drawn from the experimental results. Supplementary material in the Appendices lists in full the collected data (Appendix A.2), the data categorization (Appendix A.1), and the questionnaires (Appendixes B.1 and B.2) to support readers' understanding of the experimental methodology.

2. MHD Musical Interfaces

"Playing music requires the control of a multimodal interface, namely, the music instrument, that mediates the transformation of bio-mechanical energy to sound energy, using feedback loops based on different sensing channels, such as auditory, visual, haptic, and tactile channels." [13]

This quote defines in broad strokes the process behind any musical instrument's operation—traditional or otherwise. At its core, designing DMIs is similar to designing traditional instruments: one has to associate a given action with given feedback. More

specifically, in designing an instrument, one has to determine what action produces the sound or allows the user to modify the produced sound. Playing acoustical instruments consists of control systems bound by physical processes and laws and whose physical manipulation and operation results in sound generation [14]. By altering the method and parameters of control, the generated sound is manipulated and altered. This alteration allows the performer to change the sound production without changing the instrument's physical structure and construction. In terms of DMI creation, there is a need to determine the user interactions from device affordances, particularly its embedded sensor and actuator technology. There is a separation between the interface (how the user controls the system) and the sound engine (what the system provides as sonic feedback) [15]. This mapping stage between the interface and sound engine is intrinsic to DMI creation [16].

Figure 1 shows a diagram of the mapping process [17]. These three mapping layers correspond to the translation of different data types onto another, going from the base system interface to sonic feedback. Arrows represent information flow between layers. Each layer may encompass a non-defined number of parameters whose information is passed and translated between each mapping layer (using several arrows illustrates the possibility of having several controls being mapped). The first mapping layer takes the actual data from sensor input and maps them to perceptual or abstract parameters (e.g., brightness, energy). A second layer takes these parameters and maps them to specific sound characteristics (e.g., cutoff frequency, amplitude). A third layer converts those characteristics into data able to drive the sound engine and provide acoustical feedback. However, this mapping stage is somewhat vague, allowing many approaches to building the controlling data corpus. In the case of our experiment, as the interface is limited to hand- and touch-related operations, this corpus would consist of a so-called control gesture set. Various methods have been employed to analyze those gestures, from machine learning [18] to neural networks [19]. Nevertheless, no defined and formal gesture/meaning corpus exists thus far.
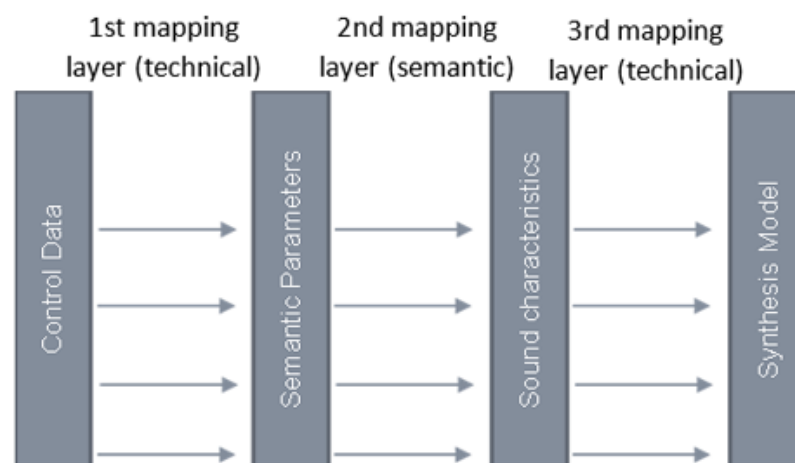
Figure 1. Three-layer mapping. Inspired by [17].

Similar to traditional musical instruments, MHDs have specific interaction modes in the context of daily device operation. The latter have been widely used in a musical context, both as simple physical interfaces and sound generation systems. However, their operation model remains mostly an emulation of the former's interface and usage model. In the following sections, we will go over device specifications and give an overview of common control interfaces in MHD DMI and musical tools to establish a baseline concerning the context of possible and adopted control methods.

2.1. Device Specifications

Digital musical instruments rely on physical actuators to produce feedback for the performer, either acoustical, mechanical, or optical (e.g., speakers, motors, lights). MHDs have very particular sensors and actuators, which are endemic to them and fundamental to their operability. In addition to the many sensors and actuators bundled with the MHDs, there are several projects and prototypes that aim to expand their control methods and sensory feedback with external add-ons (a comprehensive breakdown of augmenting approaches is available in [20]). However, an MHD self-contained DMI must conform to the device specifications and available capabilities. These hardware specificities can be seen as limitations or opportunities (constraint or affordance). On the one hand, these sensors and actuators are readily available without using any additional external device. On the other hand, the control over the sensors and actuators in most MHDs exists as black boxes, whose response behavior can be accessed without any advanced degrees of regulation of the underlying logic. Currently, MHDs are commonly equipped with (at least) the following:

- Two physical sensors (i.e., accelerometer and gyroscope), one optical sensor (i.e., camera), and one acoustical sensor (i.e., microphone);
- One acoustical actuator (i.e., speaker), one mechanical actuator (i.e., vibration motor), and many optical actuators (e.g., status LEDs, edge/rim LEDs, flashlight);
- One hybrid optical sensor/physical actuator (i.e., the Touchscreen).

On top of these physical input and output capabilities, MHDs' computing power is on par with personal computers, even outperforming some of them [21,22].

Two aspects that are paramount in terms of instrument appeal and adoption are its learning curve and potential for virtuosity. Jordà [23] considers the ability to appeal to both beginners and experts as the ultimate goal in designing an instrument. Tanaka describes the smartphone as a "*self-contained and autonomous sound-producing object that enables a musician to perform in a life situation*" [12]. Taking into account their wide availability and pervasiveness in our society, coupled with the aforementioned wide array of possibilities in terms of physical sensing and feedback, as well as processing power, it is easy to see the potential of these devices as DMIs available and appealing to a large public, both novice and expert. The devices themselves are not musical instruments but, much like any other digital music system, allow creating software applications that take advantage of their hardware capabilities to enable music and sound creation. When designing such software, one has to follow the traditional instrument design methods. For example, one must ensure that a novice user can easily play the instrument by exploring interaction methods common to the device (e.g., tying sound production and manipulation to common device interaction methods, such as touch and movement). On the other hand, one should equally consider enough degrees of control for expert proficiency over prolonged use.

2.2. Music Control Metaphors and Methods

The workflow and interface metaphors of existing mobile music software for MHDs are commonly appropriated from other, often older, realms, such as:

- Traditional musical instruments via graphical interface (e.g., Perfect Piano (https://play.google.com/store/apps/details?id=com.gamestar.perfectpiano, accessed on 19 February 2021), Real Guitar (https://play.google.com/store/apps/details?id=br.com.rodrigokolb.realguitar, accessed on 19 February 2021), Real Drum (https://play.google.com/store/apps/details?id=br.com.rodrigokolb.realdrum, accessed on 19 February 2021));
- Traditional musical instruments by mixing graphical interface and physical sensor input (e.g., DrumKnee Drums 3D (https://apps.apple.com/us/app/drumknee-drums-3d-drum-pad/id1241682795, accessed on 19 February 2021), Ocarina Smule [24];
- Digital tools for other platforms (such as Digital Audio Workstations and synths) only using the touch as emulation for mouse-based interaction or physical controller opera-

tion (e.g., Apple Garageband (https://www.apple.com/ios/garageband/, accessed on 2 February 2021), Steinberg Cubasis (https://new.steinberg.net/cubasis/, accessed on 12 February 2021), Korg INC. Apps (https://apps.apple.com/us/developer/korg-inc/id363714046, accessed on 19 February 2021).

The above categories target different users and usages, where MHDs can enhance degrees of usability (e.g., portability, learning curve, physical strain) compared to their modeled counterparts. However, while interaction methods, namely control gestures, are appropriated from the emulated model, a translation to the MHD is needed. Touch-based interaction gains particular highlight, with users usually controlling the tools via visual interface elements.

Other musical tools adopt different approaches, tied to specificities of the devices (e.g., Gyrosynth (https://apps.apple.com/us/app/gyrosynth-for-iphone-4/id386527164, accessed on 19 February 2021), Holon (https://apps.apple.com/app/holon/id1352687747, accessed on 19 February 2021)), using additional accessories to augment the device's regular operation (e.g., The Motion Synth [25]) or adopting an exploratory or abstract control approach [26], designing their control methods based on the specific desired audiovisual end result. Nonetheless, each approaches interaction control in a specific way.

Beyond the above self-embedded controller and sound/music generator applications, there is a growing interest in the bridge between mobile and desktop music environments, either by giving desktop systems access to MHDs sensor data and actuator control through various communication protocols (e.g., SensoDuino (https://play.google.com/store/apps/details?id=com.techbitar.android.sensoduino, accessed on 19 February 2021), PhonePi (https://play.google.com/store/apps/details?id=com.phonepi, accessed on 19 February 2021, Sensors2OSC (https://sensors2.org/osc/, accessed on 19 February 2021), TouchOSC (http://hexler.net/software/touchosc, accessed on 19 February 2021), Sensor Node Free (https://play.google.com/store/apps/details?id=com.mscino.sensornode, accessed on 19 February 2021) or allowing the embedding audio processes and programs directly into native mobile apps (e.g., libPD [4], SupperCollider-Android (https://github.com/glastonbridge/SuperCollider-Android, accessed on 5 January 2021), Csound for Android (https://play.google.com/store/apps/details?id=com.csounds.Csound6, accessed on 19 February 2021), MobMuPlat (http://www.mobmuplat.com/, accessed on 19 February 2021).

3. Materials and Methods

A procedure was designed to study how users approach control gestures without prior instructions on the interaction methods. Two constraints were adopted: using a smartphone as the physical device and restricting participant manipulation to touchscreen operation and physical device movement. Touchscreen operations consisted of the manipulation of the axis-based touch position. Regardless of device position, *X*- and *Y*-axes corresponded to the horizontal and vertical axis, respectively (e.g., whether in portrait or landscape mode, the vertical axis was always considered the *Y*-axis). Physical device movement was considered both in terms of device translation and rotation. Users were prompted to reproduce a series of sound stimuli via the aforementioned smartphone degrees of control. No information on the interaction control or nature of the stimuli was provided before the experiment. This strategy aimed to capture the participants' everyday use of the MHD as an instinctive response. A premise from user experience design is adopted here by assuming that intuition and the practical user knowledge acquired from daily device operation can guide the design on fluid interactions controlling sonic parameters. Guidelines would result from gesture controls that most naturally are associated with interacting with the device.

The experiment was divided into two similar phases, each consisting of two distinct tasks. The first task consisted of sound stimuli reproduction. For each sound stimulus, participants were asked to listen (and were allowed to re-listen to the stimulus once if desired), wait for a visual prompt from the mobile application, and proceed with the

control gesture, which best represented what they heard. Next, they would either vocally or visually indicate gesture completion. This procedure was repeated for each of the sound stimuli. The second task consisted of reviewing a video recording of their performance on the first task and answering an open-ended questionnaire, which can be found in Appendix B.2. At the end of this questionnaire, these two tasks were repeated in phase 2, providing us with a defined rule-set to analyze emerging trends in participant choices.

This twofold design aimed at having participants approach the experiment in two different ways—instinctive and informed. Before the first phase, participants were only informed about the device manipulation constraints and nature of the task but were unaware of the nature of the sound stimuli or musical parameters under study. This approach resulted in the participants reacting instinctively to the varying musical parameters while having little time to internalize a structured gesture rule-set. During the second task of phase 1, participants were given a chance to reflect on their approach and internalize expectations and reactions by reviewing their performance on task one.

In phase 2, the detailed experiment in the first phase was repeated, but this time, participants were not only aware of the expected musical mappings but had gone through an assessment of their choices, approaching this new phase of the experiment with knowledge of what the expectations were in terms of musical parameter mapping—what we call an informed approach. In allowing participants to reflect and evaluate their performance, the second task allowed us to collect data concerning the participants' performance via a questionnaire. Collected results consisted of the three answers for each sound stimuli in each phase. For each of the sound stimuli, we verified firstly whether the participant perceived the variations in musical parameters (i.e., note pitch, duration, and amplitude), what gestures were used to represent the variation (e.g., touch coordinates, physical device movement), and the rationale for the choice of said specific gestures (e.g., trying to mimic an instrument, reproducing a visual interface control). If the particular musical parameter variation was not perceived, participant's choices were disregarded, as any chosen mapping would pertain to some other hypothetically perceived parameter.

The collected data allowed us to compile frequencies for gesture choices, mapping rationale, and musical parameter variation perception. From the frequency data, we further analyzed emerging trends in mappings and established comparisons between participant profiles and the potential impact it might have on their approach to the experiment.

1. Are there predominant gestures that users associate with mapping a given specific musical parameter (i.e., note pitch, note duration, note amplitude)?
   When confronted with a sound stimulus exhibiting a specific musical parameter variation and given no instructions outside of manipulation constraints, each user will have an instinctive choice to represent that variation. We want to determine if any choices are shown to be prevalent and, thus, can be considered as more natural (in this context) than others;

2. What is the most common rationale behind these mappings?
   We also want to understand why users make their specific parameter mapping choices. This would allow us to understand better how to design interaction methods for MHD-based DMIs. Depending on the prevalence (or absence) of trends in mapping rationale, one can approach further parameter mappings (other than the three studied in this experiment) from a similar perspective and approach;

3. Is there a change in mapping between an instinctive and an informed approach?
   Users change their behaviors and control over musical instruments with prolonged use. Their approaches to instrument manipulation change as they assimilate constraints, affordances, and response. We want to try and ascertain the impact of the users' ability to adapt on the mapping choices and if there is perceivable learnability even in such a short cycle of usage;

4. Are these gesture mapping choices, rationale, and changes influenced by musical expertise?

Considering we are dealing with a musical context and musical parameter control, it is essential to ascertain whether musical expertise plays a role in the results. The potential familiarity with other methods of musical control may introduce bias in the process. For instance, one's instinctive choices for musical parameter control might be dictated by previous instrumental practice, or the ability to perceive specific musical details may be hindered by the absence of musical training. This assessment is extremely important for trend analysis and keeping with the premise of developing instruments usable by both novice and expert users.

3.1. Experiment Task 1

The first task of each phase of the experiment expected the participants to reproduce sound stimuli using a provided device. Figure 2 shows a diagram of the disposition of the whole apparatus during this task.
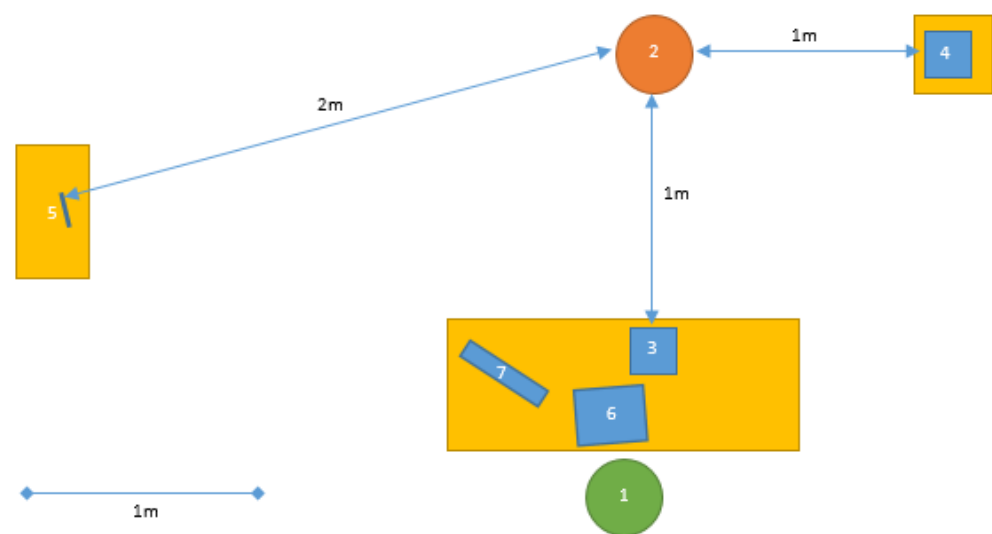
Figure 2. Experiment set-up: 1—Researcher; 2—Participant; 3,4—Speakers; 5—Video recording; 6—Researcher computer; 7—External monitor.

An Android mobile application was developed in Java to run on the device participants used during the experiment. This application ran on a low-range 5.5″ MHD with *Android* OS version 5.1.1 (*Vodafone Smart Ultra 6*) and served to prompt the participants to start device manipulation by showing a visual call-to-action (app screen changed from the app logo to an entirely white background). The application also logged timestamped touchscreen interaction data (i.e., number of touches and coordinates) and raw accelerometer data. Sensor logging data are not yet reflected in this present paper, with its objective further detailed in Section 6.1.

A *Pure Data* patch created using *PD Vanilla 0.50.2* [27] remotely controlled the mobile application. This patch was used to send messages to the app controlling its behavior (i.e., triggering the call-to-action change, data logging start and end) and receiving networked messages from the smartphone, allowing the researcher to know what the status of the mobile app was at all times and ensuring messages were correctly delivered. It was also responsible for playing back the sound stimuli. Figure 3 shows a simplified view of the system structure.

Figure 3. System dataflow structure.

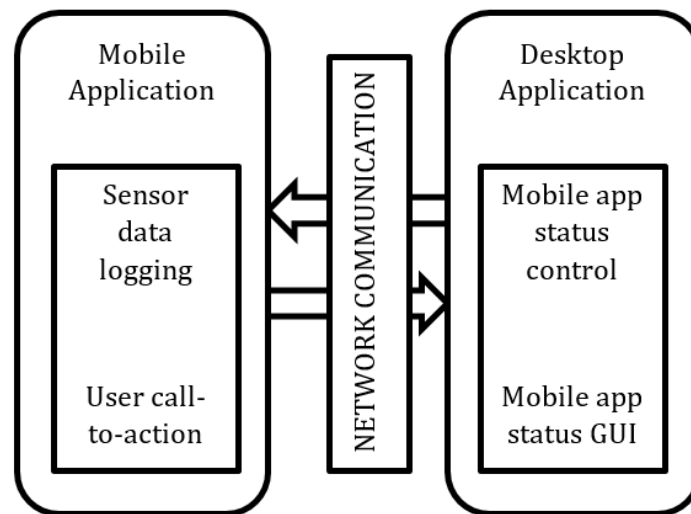Participants' performance was video recorded using a tripod-mounted smartphone (*Sony Xperia*) camera pointed at the participant's arms and hands, transmitted via wireless live video feed to the computer, and recorded in real-time. Recorded videos were subsequently reproduced on the *HP Elitebook* for participant review on an external screen.

3.2. Experiment Task 2: Questionnaire

The second task of each phase aimed at both allowing the researcher to confirm participant's performance and reflecting on their choices. This review consisted of reviewing the video footage of the participant's performance and asking them three questions (detailed in Appendix B.2) about each sound stimuli to help understand musical parameter perception, control gesture choices, and intention behind each choice.

Question 1 served as a baseline to assess the participant's awareness of the stimulus's parameters variations, determining if their mappings could be considered for analysis. For any given stimulus, if the participant could not perceive that stimulus's specific musical parameter variation, their mappings would not reflect the targeted variation but some other arbitrary one. This correct perception did not depend on precise parameter identification (i.e., if the participant was able to define the varying parameter correctly —e.g., *note pitch was changing*), but rather on ascertaining if the participant perceived a change that corresponded to that particular stimulus's variation.

Question 2 aimed to ascertain the intended performed participants' gestures and actions.

Question 3 directed the participants to analyze and reflect on their gestures and assess the underlying rationale and motivation.

3.3. Participants

In this experiment's scope, target users were considered part of the general population with common MHD usage experience. High-level proficiency was not expected, but familiarity with MHD operation was required. Three primary tasks were considered to establish a baseline for this degree of familiarity: e-book/document reading and creation, photo editing, and gaming. Regular execution (at least daily) of either of these tasks was considered enough proficiency in the operation of MHDs.

Considering the potential impact of musical performance training on gestural control of sonic parameters, we adopted Cifter and Dong [28] classification on *Professional-Users* and *Lay-Users*. Participants with current or past regular musical instrument practice (acoustic or digital) or formal musical training were considered to fall into the "musician" profile, encompassing Cifter's definition of both *Professional Users* and *Experienced Users* (basic primary school-level music classes were disregarded in this determination of musical

proficiency). Other participants would be considered *Novice-Users* according to Cifter's classification and fall under the "non-musician" profile. Non-musicians were expected to be regular music listeners to be eligible for the experiment to guarantee they were familiar with the basic musical characteristics to be evaluated. This selection was achieved via the pre-experiment questionnaire (listed in Appendix B.1), establishing participant eligibility, and categorization in musical profile.

Participants (N = 27) were recruited with purposive sampling. We recruited among personal contacts for young (20–30 years old) musician and non-musician participants. Age was restricted to minimize the potential impact on participant profile. Musicians (n = 14) were aged around 23 years old (M = 23.5, SD = 2.67) and included nine males and five females. Non-musicians (n = 13) were aged around 22 years old (M = 22.4, SD = 2.00), with six males and seven females. Differences in age and gender were not statistically significant (*age: U* = 68.5, *p* = 0.280; gender: $\chi^2(1)$ = 0.898, *p* = 0.343).

3.4. Sound Stimuli

Figure 4 shows the five sound stimuli used in the experiment. Stimuli encompass parameter variations in pitch, duration, and amplitude. The three first sound stimuli (a)–(c) introduced gradual variations of the note parameters under consideration. The strategy aimed to increase the participants' perception of each new attribute and become acclimated to their variation.

The controlled and gradual note parameter variations had implications in designing the procedure by implementing randomization to remove any possible parameter learning bias introduced by a particular fixed order. The first three stimuli (where musical parameters vary individually) can be arranged in a total of 6 permutations (*ABC, ACB, BAC, BCA, CAB, CBA*). These possible order permutations were distributed so that each was attributed to the same number of participants of both profiles. The particular attribution to each participant was defined using *random.org*'s (https://www.random.org/, accessed on 20 December 2020) list randomizer.

The fourth and fifth stimuli combine variations across all parameters under study, aiming to understand participants' control gestures using multi-parameter variations.

Figure 4. Sound stimuli (**a**–**e**), represented using musical notation for visualization.

Sound stimuli had variable durations (between roughly 3 and 4 s) to account for notes with different durations. For the sake of simplicity, we adopted 500 ms and 1000 ms to denote different short and long note durations to provide distinguishable parameter values and accommodate both musician and non-musician participants' perceptions. Dynamics represented in the notation (piano, mezzo-forte, forte) are not bound to any specific amplitude and serve as a visual representation of the note volume difference. Amplitude differences between dynamic levels were determined via experimentation based on hardware specificity to avoid inaudible low amplitude levels or distorted high amplitude levels. The three different levels were selected as a compromise in audio reproduction quality (i.e., the absence of distortion or other undesired artifacts) and perceivable differences. Both duration and amplitude parameters were defined in a prior informal pilot experiment (discussed further in Section 5).

Table 1 provides an overview of the musical parameter variations across sound stimuli. The first stimulus (henceforth referred to as stimulus a) introduced the participant to note pitch change while using a short note duration (500 ms) and constant amplitude. The second stimulus (stimulus b) introduced the participant to varying note duration (1000 ms, 500 ms, 1000 ms), with fixed amplitude and pitch. The third stimulus (stimulus c) featured notes with constant pitch and short duration (500 ms) while introducing variation in amplitude, corresponding to piano, mezzo-forte, and forte dynamics (low, medium, and high amplitude). The fourth stimulus (stimulus d) introduced a simultaneous variation of all three characteristics (pitch, duration, amplitude). The fifth and last stimulus (stimulus e) consisted of a "*curve-ball*", so to speak, introducing the new parameter of unexpected polyphonic note reproduction and forcing the participant to reconsider their previous mapping choices. Furthermore, it aimed to provoke a deeper questioning while completing the second task of each experiment phase (i.e., the questionnaire part of the experiment).

Table 1. Musical stimuli musical parameter variations.

	Stimulus				
Musical Parameter	**a**	**b**	**c**	**d**	**e**
Pitch	Changing	Fixed	Fixed	Changing	Changing
Duration	Fixed	Changing	Fixed	Changing	Changing
Amplitude	Fixed	Fixed	Changing	Changing	Fixed
Polyphony	No	No	No	No	Yes

Sound stimuli were generated with a Sawtooth waveform synthesizer and exported as 44.1 kHz/24-bit WAV files. Sawtooth was chosen to have a synthetic sound and avoid bias from instrument sound approximation (preliminary testing revealed that some participants associated Sine wave sounds to flute or recorder sound, resulting in them biasing their device manipulation to emulate those instruments).

Sound stimuli were reproduced on an *HP Elitebook* laptop computer, using an external sound interface and good-quality audio monitors.

3.5. Data Analyses

The statistical analyses were conducted in SPSS version 25 (IBM, 2020). χ^2 testing was used to determine whether there were statistically significant associations between variables.

For the purpose of statistical analyses, we considered the following significance level conventions: Significant result $p < 0.05$, Marginally significant result: $0.05 \leq p < 0.10$, Non significant result: $p \geq 0.10$. Although reporting marginal results is a controversial practice, considering the exploratory nature of this study, it is interesting to be aware of possible tendencies slightly outside of the traditional $p < 0.05$ significance level. We chose to adopt the marginally significant definition [29,30] to represent this.

In addition to the analysis performed on collected data, the control gestures' list was coded into broader categories for additional analysis. This categorization was based on the gestures' nature, grouping them into more generalized categories based on their core characteristics, and aimed at analyzing mapping approaches in a more general sense, attempting to find links between manipulation types and the studied musical/sonic parameters. The final gesture list was analyzed to find common characteristics and achieve this categorization. Considering the interaction constraints for the experiment (touchscreen operation and device movement), we found that touch-based gestures could further be categorized into coordinate-based gestures, by which the participant was mapping variation to a specific position on the touchscreen (e.g., using the vertical or horizontal axis to represent a scale of values) and touch characteristic-based gestures, by which participants mapped variation to a specific characteristic of the touch in itself (e.g., the duration of the touch, the pressure of the touch, the area covered by the finger touching). We also found that participants sometimes combined gestures from any of the three main categories. The resulting categorization consisted, thus, of four broader categories: 2D Plane manipulation, Touch characteristics, Device position, Combination.

Gesture choice rationale answers were equally organized into broader categories to organize the participants' answers. These emergent categories were reached by analyzing the collected data after the experiment and performing a screening based on common characteristics. Considering the open-ended nature of answers, it would be complicated to find trends among answers. Categorization aimed to organize answers into more specific categories and reduce the answers to their main underlying reason. It resulted in the following broader categories: Instrument mimicking, Graphical representation, Intuition, Physical mapping, Musical bias, Exploration, Unsure, User experience, Complementing other mappings, Using previous mappings, Combining previous mappings, Instrument mimicking, and physical mapping.

Collected data categorization for both (gestures and rationale) is detailed in full-length in Appendix A.1.

3.6. Experiment Protocol

In addition to the data collection and analysis supporting guidelines for musical parameter mapping in MHD musical instruments, we developed a protocol to evaluate these mappings, consisting of a complete experiment script, questionnaires, and tools to run the experiment and analyze gathered data. This experiment served to validate the protocol and the developed tools, which we made available for open access (https://zenodo.org/record/4553522, accessed on 20 February 2021). We believe this protocol fills a gap in digital music tools by analyzing and validating MHD musical tool operations.

4. Results

As detailed in Section 3, the experiment was divided into two phases. In the first phase, participants reacted to the stimuli without prior knowledge of the stimuli's nature. In the second phase, participants reacted to the same stimuli after reflection. It is, thus, essential to analyze and compare the results from both phases. We collected frequency distributions for each phase and assessed participant profile's impact across the following variables: musical parameter variation perception, uncategorized mapping choices, categorized mapping choices, mapping rationale, and mapping changes (intra- and inter-phase). These results are based on collected data from the questionnaire, denoting participants' intentions on performed gestures, which we list in full in Appendix A.2.

4.1. Gesture Mapping Frequencies

Figures 5 and 6 show the observed gesture mapping frequencies for phases 1 and 2 for both participant profiles. Gestures with only one participant choice are grouped into the category "others" for better readability.
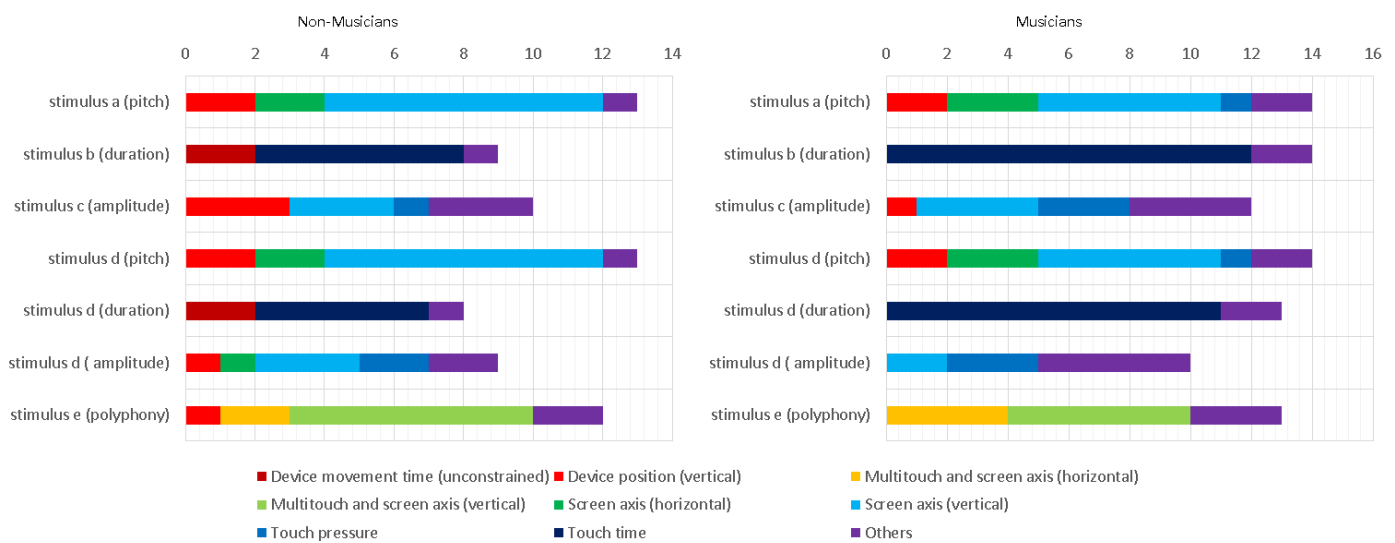
Figure 5. Frequencies of gestures performed by participants for each stimulus in phase 1. **Left**: Non-musician participant frequencies; **Right**: Musician participant frequencies; Full results reference: Tables A3–A9.

Figure 6. Frequencies of gestures performed by non-musician participants for each stimulus in phase 2. **Left**: Non-musician participant frequencies; **Right**: Musician participant frequencies; Full results reference: Tables A10–A16.

One additional piece of information that can be introduced as a by-product of this experiment's results is the mapping of note triggering/onset mapping. This mapping is an integral part of any instrument and is tied to all the considered musical parameters. This mapping was inferred from analyzing the gesture mappings. It was verified to be the same between phases: fourteen musicians and eleven non-musicians used gestures that triggered notes using touch. In contrast, two non-musicians used gestures using device movement as a note trigger, which also denotes a pronounced trend towards touch-based note triggering.

4.2. Mapping Rationale

When asked about the rationale for mapping, participants varied in their responses. We first determined the relevant degrees of comparison in the data to analyze the mapping rationale results. In particular, from the results from stimulus d (combining all three musical parameter variations), we observed that the overwhelming majority of participants justified their mapping choices as an attempt to combine previous mappings. The same was verified in the answers for stimulus e. It is more interesting to consider the mapping

rationale given for the previous stimuli (a, b, c), since they were the de facto bases for subsequent answers.

As seen in Figures 7 and 8, participants reported Instrument mimicking as the main reason in the case of pitch and duration mapping. As for amplitude, we can observe that both Instrument mimicking and Graphical representation came out as the most frequent in phase 1. In phase 2, Instrument mimicking became the most common, although not by much. Rationale answers with only one participant choice were grouped into "others" category for better readability.

4.3. Gesture Mapping Changes

Full results reference: Tables A3–A16.

Participants from both profiles changed mapping gestures for individual parameters (pitch, duration, amplitude) from stimuli a–c to stimulus d in both phases. Figure 9 details the frequencies of these inter-phase mapping changes and of the changes occurring between stimuli across both phases.

Figure 7. Frequencies of rationale evoked by participants for the gestures performed for each stimulus in phase 1. **Left**: Non-musician participant frequencies; **Right**: Musician participant frequencies; Full results reference: Tables A17–A21.
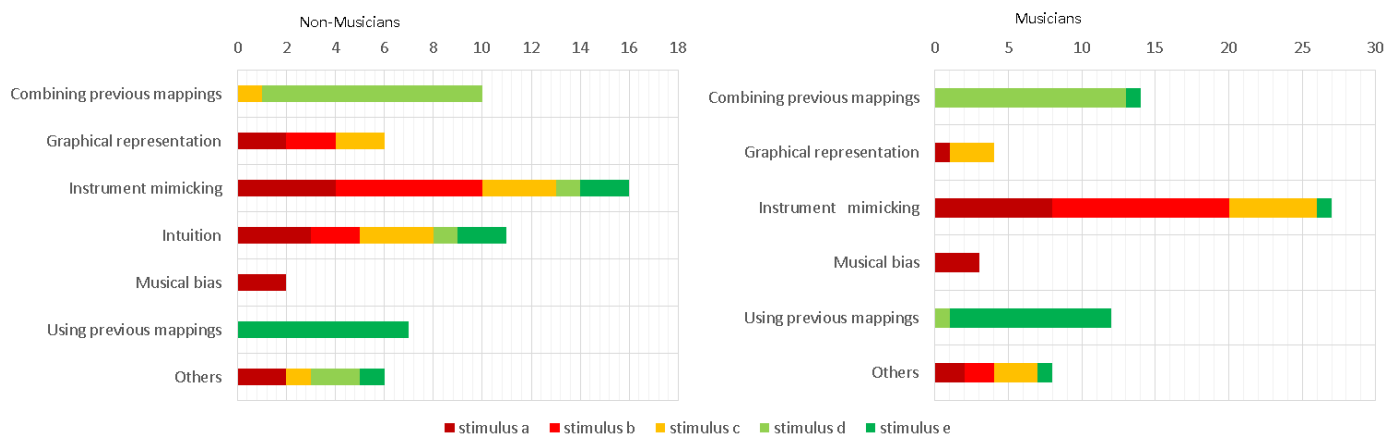
Figure 8. Frequencies of rationale evoked by participants for the gestures performed for each stimulus in phase 2. **Left**: Non-musician participant frequencies; **Right**: Musician participant frequencies; Full results reference: Tables A22–A26.
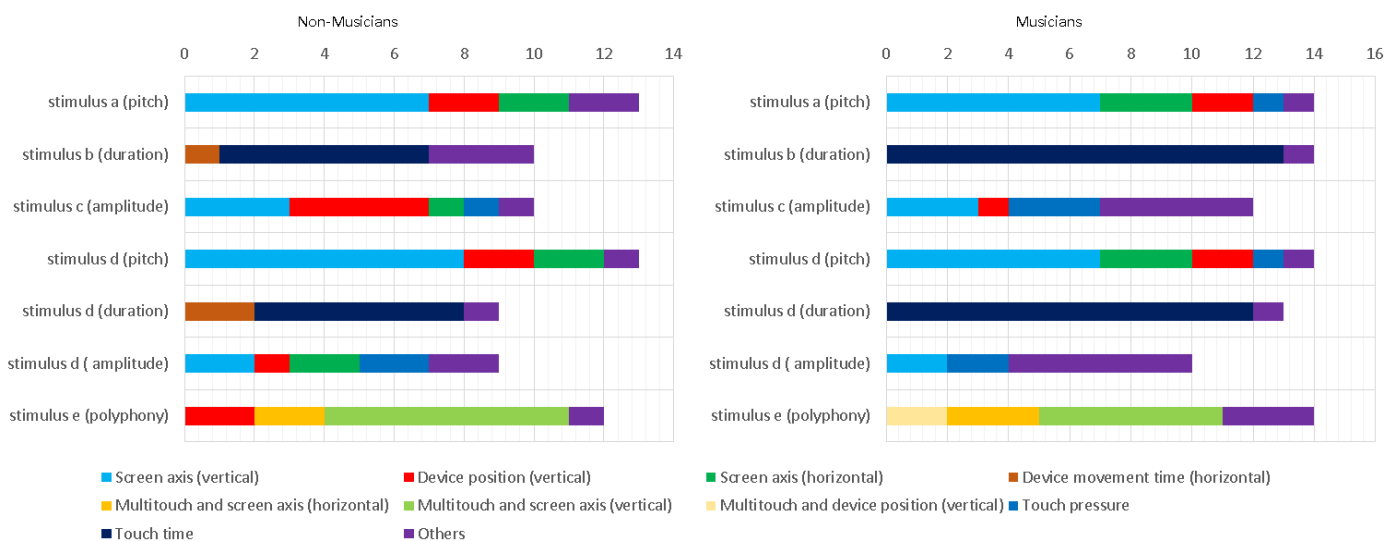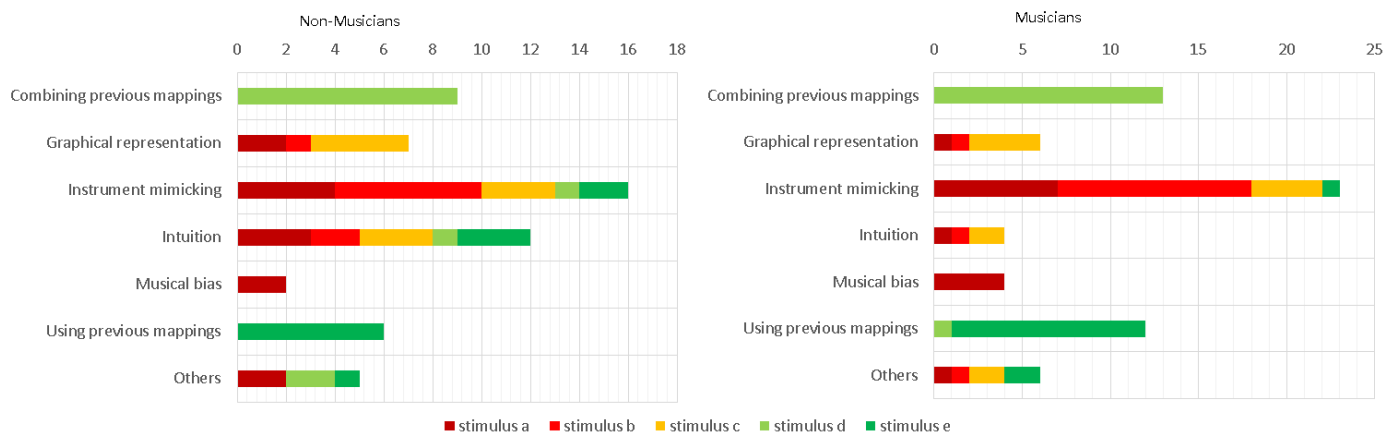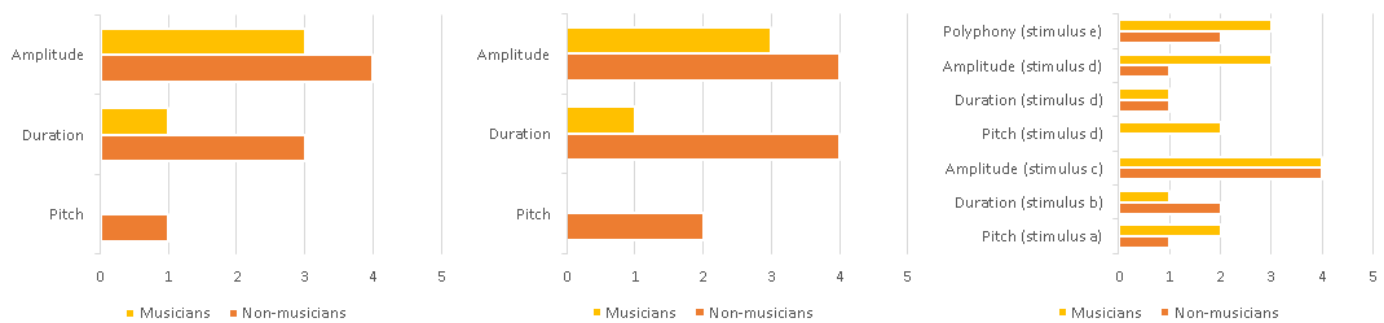
Figure 9. Frequencies of participant gesture mapping change for each phase and between phases 1 and 2; **Left**: Phase 1 intra-phase mapping changes (stimuli a, b, c to stimulus d); **Center**: Phase 2 intra-phase mapping changes (stimuli a, b, c to stimulus d); **Right**: Phase 1 to phase 2 mapping changes.

4.4. Profile Association

Parameter variation perception

In phase 1, duration variation perception in stimulus b was shown to have a strong association with participant profile (significant: $\chi^2(1) = 5.06$, $p = 0.04$). The same parameter variation perception also exhibited a notable but less pronounced association to participant profile in stimulus d (marginally significant: $\chi^2(1) = 3.83$, $p = 0.08$). Other variables were found to exhibit no statistically significant association to participant profile. In phase 2, duration variation perception was the only variable that exhibited any association with participant profile, with its perception showing a marginally significant result ($\chi^2(1) = 3.64$, $p = 0.10$) for stimulus b. Other variables were found to exhibit no statistically significant association to participant profile.

Gesture mapping for stimulus b (isolated duration variation) was the only one showing any degree of association to participant profile, with a marginally significant result ($\chi^2(3) = 4.80$, $p = 0.06$).

There was no statistically relevant association found for mapping rationale or mapping changes.

Full results reference: Tables in Appendix A.3.3.

5. Discussion

Before discussing collected data, it is essential first to analyze parameter variation perception, which directly influenced sample size for each participant profile, as these impacted the number of collected answers for each variable. As explained in Section 3.2, participants were expected to correctly identify parameter variation for their answers and choices to be considered eligible. This perception varied from stimulus to stimulus across both profiles and is shown in Figure 10. As a reminder, $N = 27$, with *n (non-musicians)* = 13 and *n (musicians)* = 14.

Pitch variation perception was shown to be the most universal. All participants were able to identify it in all of the corresponding stimuli, with only one non-musician failing to identify note pitch variation in the polyphonic stimuli. Note duration was well perceived by musician participants, with only one failing to perceive it in the combined stimulus (d) on both phases. Some non-musicians, on the other hand, did struggle to perceive it, both on the individual variation stimulus (b) and on the combined stimulus (d). Amplitude was shown to be almost perfectly identified by musicians in the individual stimulus (c) but harder to identify on the combined one. Non-musicians exhibited similar results, although with lower success in identifying variation.
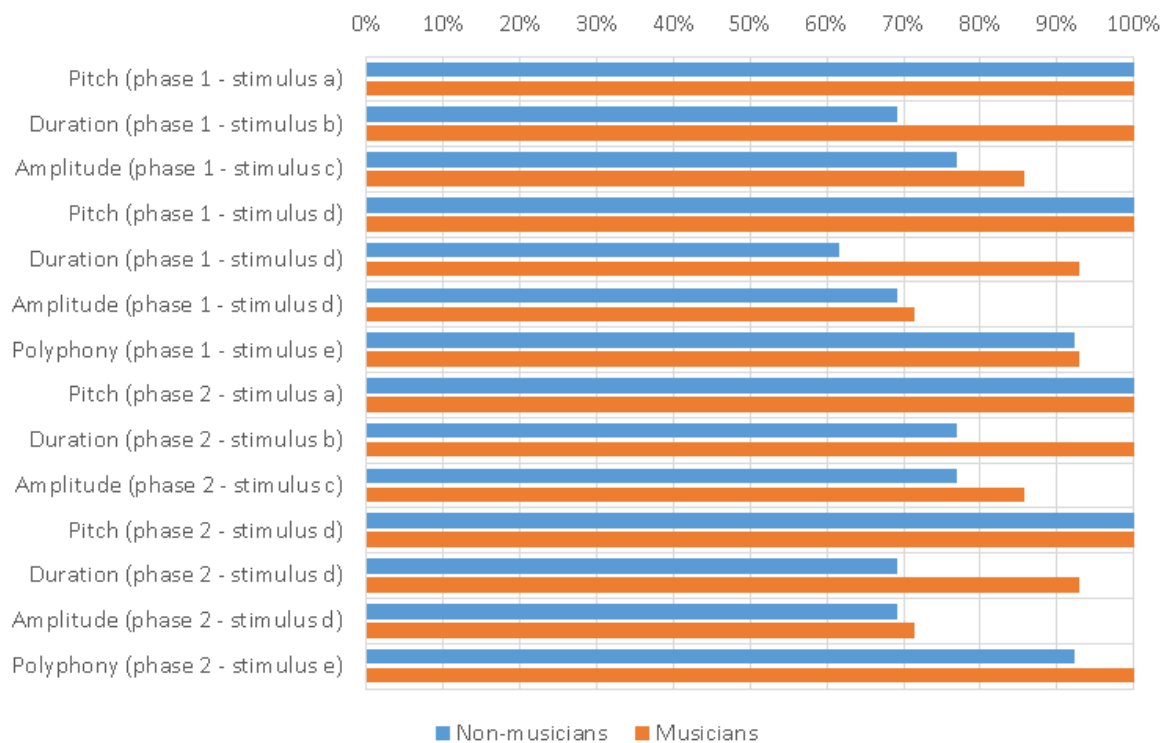
Figure 10. Relative frequencies of musical parameter variation perception on all stimuli across both experiment phases. *n* (non-musicians) = 13; *n* (musicians) = 14.

5.1. Result Interpretation and Trend Analysis

The gesture mapping choices participants provided in stimulus d, phase 2 should ideally represent their definitive mapping rule sets. The two-phase division of the experiment allowed the exploration and re-evaluation of this stimulus, combining all three analyzed musical parameters' variation at once, which would be the confirmation of their final mapping choices. Unfortunately, one of the problems we verified was that, even with the provided reviewing and discussion of their phase 1 performance and sound stimuli, several participants were unable to identify either duration or amplitude (or both) variation in this stimulus, resulting in incomplete gesture mapping sets (this is further discussed in Section 6.1). Instead of analyzing specific results for just that stimulus, we mist look at the mapping choices and approach for all stimuli from phase 2 to properly analyze emerging trends in mappings and propose the aforementioned guidelines.

In reviewing results for stimulus d, shown in Figure 11, we can observe that both note pitch and duration mapping have noticeably steady gestures. Amplitude, on the other hand, showed no immediate clear trending choice.

If we consider the categorized gesture results shown in Figure 12, there were clear trends in note Pitch and Duration mappings and more evenly distributed results concerning amplitude (categorization process described in Section 3.5, complete categorization listing in Appendix A.1).
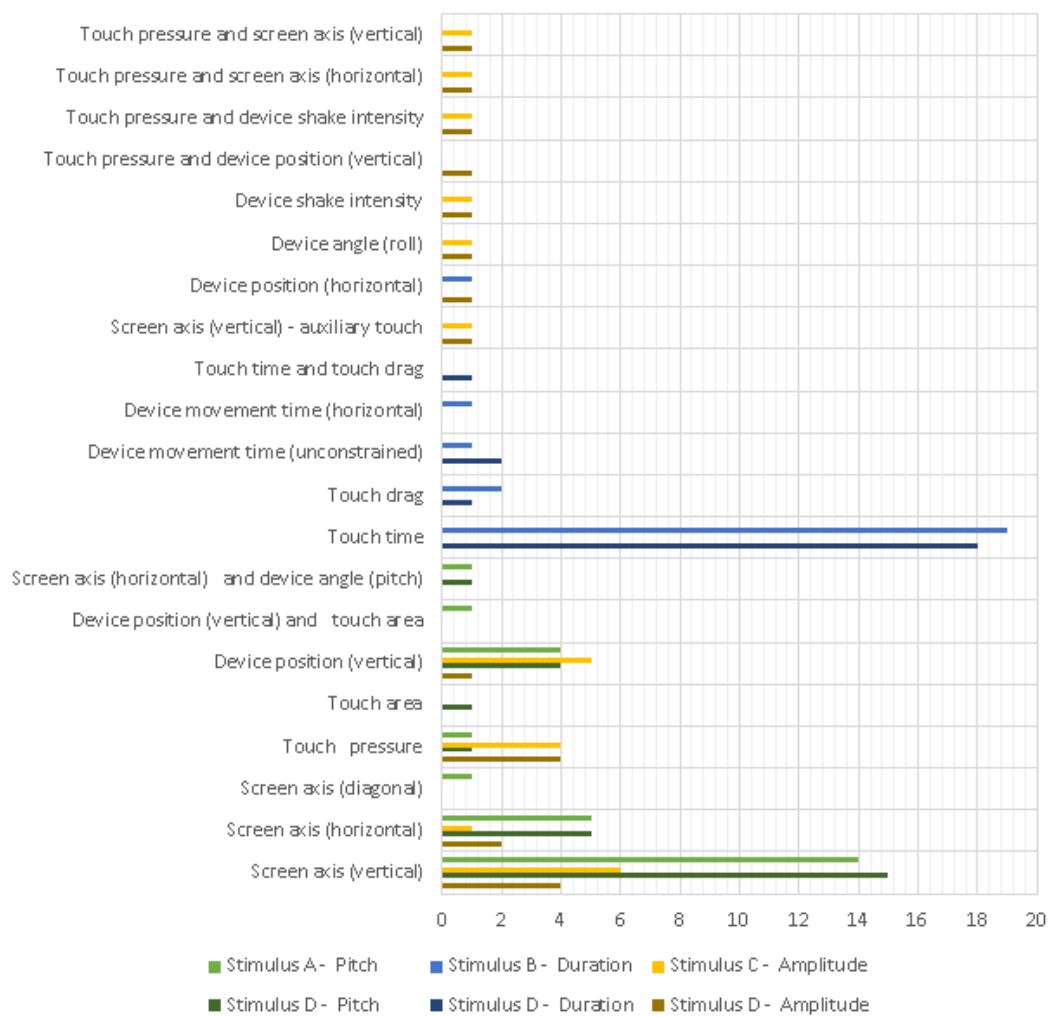
Figure 11. Frequency of participant (musician and non-musician) gesture choice in phase 2—uncategorized gestures. Full results reference: Tables A10–A15.
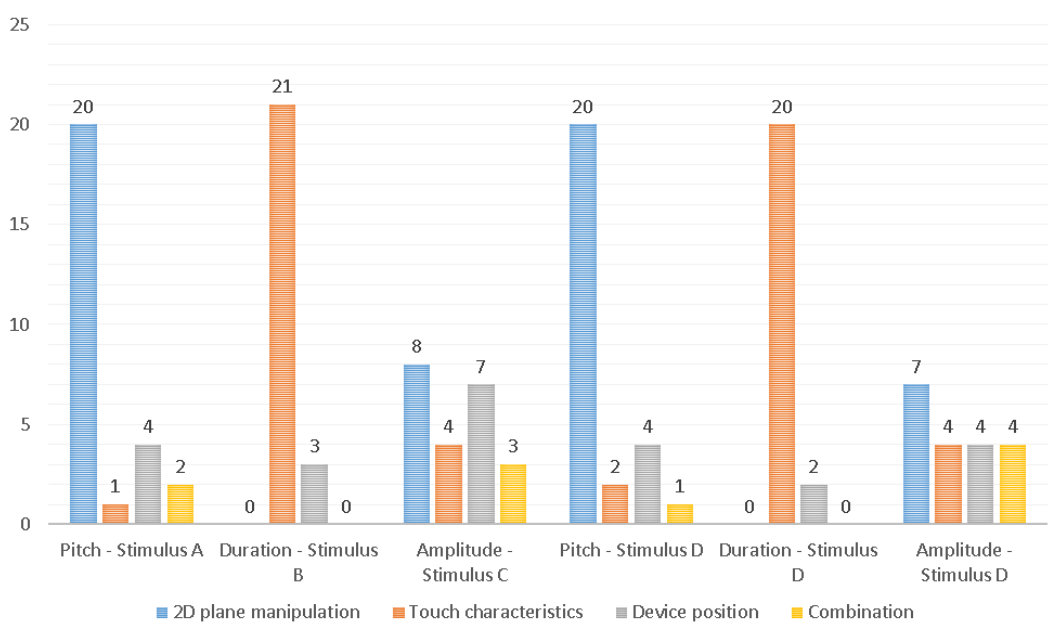
Figure 12. Frequency of participant (musician and non-musician) gesture choice in phase 2—categorized gestures.

5.1.1. Pitch Mapping

We identified a very pronounced trend towards using the device's screen y-axis (vertical) to map note pitch, with all twenty-seven participants able to perceive pitch variation. Fourteen participants chose this option in stimulus a and fifteen in stimulus d. The second most selected option (five participants in both stimulus a and d) used the device's screen x-axis (horizontal), followed by the physical manipulation of the device's vertical position, with four participants choosing it.

Referring back to Section 4.4, we saw that these choices show no association to participant profile, with non-significant results for both stimuli a and d pitch mapping for either participant profile. We can consider, thus, that pitch variation is most commonly associated with touch position mapping over an axis on the touchscreen.

5.1.2. Duration Mapping

Duration also exhibited a pronounced trend. Referring back to Figure 11, we see that twenty-two participants were able to identify duration change for stimulus d, while twenty-four did so successfully for stimulus b. Nineteen participants chose Touch Time to map note duration for stimulus b and eighteen for stimulus d. Considering the decrease in the number of successful variation perception, these can be considered equivalent. It is also interesting to note that only two choices (Device movement time—unconstrained and Device movement time—horizontal) were unrelated to touch time in itself out of all the choices. Both Touch drag and Touch time and touch drag represent, in essence, the same variable as Touch time: mapping note duration to the duration of the touch itself, leading us to consider that note duration is overwhelmingly associated with touch duration, in a behavior similar to a Note-on/Note-off MIDI event [31].

### 5.1.3. Amplitude Mapping

Amplitude, on the other hand, was the parameter whose variation participants failed the most to perceive consistently. When it was combined with other parameters, only nineteen participants (out of twenty-seven) could perceive amplitude variations in stimulus d. In contrast, twenty-two participants were able to perceive its individual variation (stimulus c). Furthermore, amplitude was the parameter whose mappings showed a wider range of choices. It is interesting to look deeper into each of the stimuli pertaining to amplitude variation (C and D) and analyze the rationale per participant profile. In the absence of a clearly defined trend, different conclusions have to be taken from these results. Considering the variation between the number of participants able to perceive amplitude variation in each stimulus, it is best to compare results through percentages instead of choice count.

If we focus on mappings for stimulus c, we see that non-musicians ($n = 10$) were divided between device vertical position (40%) and 2D axis touch positioning (40% as well, if we add up all different gestures making use of this approach). Musicians ($n = 12$), on the other hand, had a much more distributed array of choices, with Screen axis (vertical) and Touch pressure barely showing as the most selected options (25% each), and the other six choices each having 8.3%.

If we now look at mappings for stimulus d, we see that non-musicians ($n = 9$) chose a higher number of mappings making use of 2D-coordinate vertical or horizontal positioning (44.4%), followed by Touch pressure being chosen by 22.2%, and other choices all having 11.1% of participants—the prevalent choice stimulus c (Device position (vertical)) went from 40% to 11.1%.

There seems to be an even wider distribution of choice in musicians' case, with vertical-axis 2D position and Touch Pressure each being chosen by 20% of participants ($n = 10$) and all other options each being chosen at 10%. If, however, we take a closer look at the secondary choices, we see that *Touch Pressure* was a part of four of them, while 2D-coordinate positioning was part of two of them. We could then consider Touch pressure to be at least part of the choice for 60% of participants, while 2D-coordinate positioning

was part of 40%. It should be noted that, during the post-experiment discussion between the researcher and participants, many of them referred to Touch Pressure specifically, explaining that their first choice would have been to use it, but refrained from doing so because they knew that the sensor is not currently widely available in MHDs.

If we look at Figure 12, we can see this parameter's mapping once again had no clear trend in terms of the overall interaction approach, with 2D plane manipulation barely coming in front of other categories. Going a bit in-depth, referring back to Figure 11, we can observe that four of the choices under the Combination category make use of Touch characteristics, while two make use of 2D Plane Manipulation, putting the two gesture categories at the forefront of the participant choices.

In sum, we can observe that 2D-axis coordinates seem to be the mapping to which non-musicians gravitate, while Touch pressure seems to be the option towards which musicians gravitate. This might be indicative of practitioner's bias. Participants with instrumental background or knowledge associate the dynamics of a sound with the intensity of its note triggering mechanism, whereas participants with no instrumental knowledge view parameter variation as a whole in a scale-based visual way. Nonetheless, remaining within the analysis of collected data, we cannot present a definitive trend concerning note amplitude mapping and will further discuss the implications of these results afterward in this section.

5.2. Mapping Rationale

In terms of mapping rationale, and as stated in Section 4.2, we focused on the answers provided in the first three stimuli. If we look at Figures 7 and 8, we can observe a strong predominance of Instrument mimicking as the reason for mapping choices. Interestingly, and even though that predominance is more pronounced in the case of musicians, a considerable percentage of non-musicians (*circa* 40%) gave the same justification for their choices. Looking at Graphical representation details in Table A2, we can observe that answers under this category mainly focused on representing graphical elements commonly used to control or represent sonic parameters (e.g., knobs, sliders, waveform timelines), suggesting a strong connection between this approach and the operation of familiar music and sonic tools (e.g., music players—with visual volume and speed controls, and waveform visualization of songs). However, Intuition is harder to analyze, as participants seemed to provide these answers whenever they could not explain their choices as conscious decisions.

Interestingly, comparing the change in rationale between phases 1 and 2, we can observe that while in the case of non-musicians, changes were not very pronounced (with percentages changing very lightly); in the case of musicians, there was some gravitation towards Instrument mimicking, after being given the possibility of reflecting and rethinking their mappings.

5.3. Mapping Changes

Pitch mapping was shown to be the most stable mapping across all stimuli and phases, with the least number of mapping changes taking place either intra- or inter-phase. Non-musicians were shown to change their mappings between individual parameter stimuli (a–c) and combined stimulus (d) more often than musicians. This was verified for both phases, which can be seen as somewhat surprising. The twofold design of the experiment encompassed a reflection moment between phases of the experiment, allowing the participants to further structure and cement their approach and rule-set, now knowledgeable as to what the expectations were in terms of mappings. Nonetheless, as shown in Figure 9, mapping changes took place between individual and combined stimuli for both profiles. Musician participants had the same number of changes, while non-musicians increased the number of mapping changes on two of the three parameters (i.e., Duration and Amplitude). This is likely related to a failure or difficulty perceiving that specific parameter's variation and is further discussed in Section 6.

Changes, unsurprisingly, are visible between phases on stimuli a–c, where a re-evaluation of mappings was expected. Nonetheless, the changes shown between phases on stimulus d did not correspond to the changes seen on the individual stimuli. This is likely a byproduct of the aforementioned problems with parameter variation perception.

5.4. Profile Influence Analysis

Most of the analyzed variables were shown not to have any association with participant profile. Gesture mapping choices and the rationale provided for those mappings showed no association with profile, which is somewhat surprising considering the potential bias of musical experience to be expected in the context of this experiment.

However, there were exceptions tied to the perception of musical parameter variation, more specifically note duration. During phase 1, we found that note duration perception was significantly tied to participant profile in the case of stimulus b—where duration was the only varying parameter, and marginally significant (at the limit of becoming significant) in the case of stimulus d—where all three parameters varied in combination. In phase 2, these results changed. One non-musician was able to perceive duration variation in stimulus b. Stimulus d no longer showed any significant association between profiles, but this is due to more musicians failing to perceive duration variation as opposed to the higher degree of perception from non-musicians.

Duration was the only parameter showing an association to profile when looking at categorized gesture mappings. Even though both profiles had a very high percentage of choosing gesture mapping tied to touch, only 60% of non-musicians did so, compared to the total number of participants.

Even if statistical analyses have shown no significance between rationale and participant profile, it is interesting to look at Figures 7 and 8 and delve a bit deeper into the data. Instrument mimicking arose as the prevalent choice across both profiles, especially among musician participants (which is to be expected). Musicians overwhelmingly made choices based on Instrument mimicking and Musical bias, with these rationale accounting for 65% of answers on phase 1 and 73% on phase 2. Seemingly, the knowledge of affordances and constraints of the experiment and particularities of the sound stimuli allowed the musician participants to structure control mappings that fell into familiar musical rule sets. Non-musicians also seemed to favor framing their choices onto other familiar rule sets (e.g., graphical representation, attempting to mimic familiar interface elements associated with similar parameters: mimicking a volume knob's rotation to represent amplitude change), or took an instinctive approach to the gesture representation (this would be an exciting avenue for future work—ascertaining if Intuition comes from socially learned inherent musical bias, device familiarity, or otherwise).

5.5. Interpretation

Pronounced trends emerged in two of the three musical parameter mappings, notably pitch change and note duration. Note onset, inferred from other mappings, also showed a pronounced trend. This already provides a solid base for defining guidelines in terms of manipulating these parameters. Considering these results, one could argue that note-related parameters seem to be associated more directly with a touchscreen-based operation (i.e., touch coordinates and touch duration). Considering that the verified trends for the first two parameters are common interaction methods available on MHDs, both perceived gravitation and participant satisfaction with those choices can be attributed to their familiarity. Additionally, it is interesting to note that these results, when viewed in conjunction with the results of the mapping rationale, point towards mimicking an instrument with touch-based duration control and scale-based pitch control. In looking at the global context of such an experiment, one cannot discard cultural influence. Considering this, we believe that this natural gravitation towards this approach is intimately tied to the pervasiveness of the piano in our musical culture (Western European) as a whole and, in particular, in the representation of digital musical instruments and controllers. This seems to be backed by

the participants' answers concerning mapping rationale, with the piano being the most frequently targeted instrument for mimicry (Table A2).

On the other hand, note amplitude does not show as clear a trend as the other parameters, with reported ambiguity in the personal mappings for this parameter. Interestingly, one of the most selected control methods, touch pressure, is an interaction not yet common in low-to-mid-range MHDs, only available in very high-end or niche devices. Other highly selected approaches were tied to the device's physical positioning, either in terms of rotation or vertical/horizontal translation and 2D-axis touch coordinates (mainly in the case of non-musicians).

If we consider touch pressure, there is the immediate issue of sensor unavailability. Some attempts have been made to develop alternate sensing touch pressure [32–34], and other approaches to touch analysis can be used for the same objective. Considering that these approaches are untested in this experiment's time frame, we shall stick to the strict universal sensor availability on MHDs, which is (as noted) scarce.

Taking movement-based operation into consideration, the most selected control gestures would be the device's movement velocity (i.e., Shake intensity), vertical device position, and device roll angle (as illustrated in Figure 13).
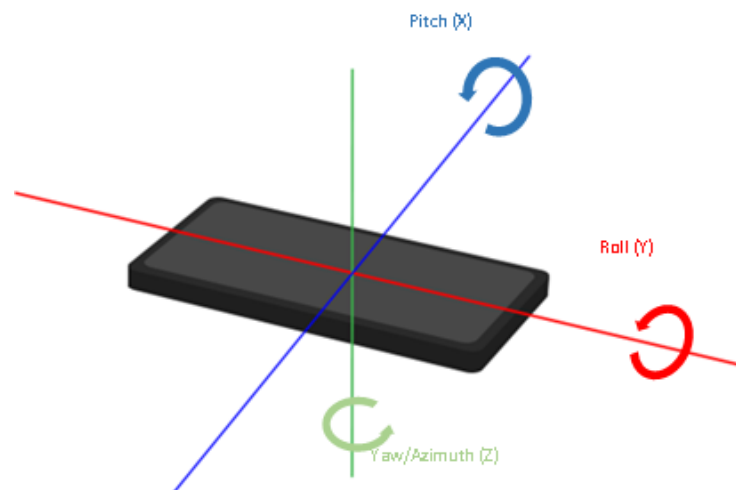
Figure 13. Device rotation angles/axes.

Using device movement velocity would make sense if the note onset approach was also movement-based or if the touch gesture was redundantly used with a device movement to take velocity from. Device vertical position (the height at which the device is held) or horizontal position (the position the device is held relative to the performer—much like a piano keyboard) is something challenging to measure, and one has to resort to either widely unavailable physical sensors or movement vector velocity calculations to determine device position (and, still, this would be relative to an arbitrary starting point). The third choice would be to map this parameter to the device Roll angle, which is easily measurable via the widely available force/acceleration sensors.

One could take, once again, the ambiguity of touch- and movement-based control and idealize a combination of both (anecdotally as proposed by one participant), using touch pressure to define the attack or starting amplitude of the note, and the device Roll angle to manipulate the amplitude envelope.

Summarizing our attempt to reach and define guidelines towards controlling these particular musical parameters in the context of a musical instrument, a specific approach can be taken from this experiment.

Leveraging on the detected trends, we can organize mappings as follows:

- Note onset and note duration would be mapped simultaneously to the touch on/off gesture on the touchscreen;

- Note pitch would be mapped to the wider axis (y-axis if the device is used in portrait mode, x-axis if the device is used in landscape mode) to accommodate a higher degree of detail within its bounds;
- As for note amplitude mapping, there is room for interpretation. Arguments could be made for taking a touchscreen-based approach or a movement-based approach. If we consider the movement-based approach and assume the difficulty in assessing the device's vertical and horizontal position, the remaining choice would be to map amplitude variation to device Roll angle. However, recent research [35] shows a notable lack of cross-device reliability in measuring this, which would impact this mapping's quality and detail. Touch pressure, which becoming available, would be the most immediate choice. In keeping with the premise of familiarity and availability, we would propose that note amplitude is associated with the touchscreen's secondary axis.

6. Conclusions

This article detailed an experiment to examine links across interaction control gestures and musical attributes, such as pitch, duration, and amplitude. This experiment was divided into two similar phases, aimed at having participants approach the experiment differently in each—firstly in a reactive approach, then in a reflected approach. Participants were asked to reproduce a series of sound stimuli and allowed to approach that task freely, with the sole constraint of sticking to touchscreen and motion-based control. This was an attempt to analyze how people approached musical gestures within the context of smartphone operation to define guidelines concerning control methods for mobile-based digital musical instruments.

The experimental results identify pronounced trends in terms of note onset, note pitch, and note duration control via touch-based interaction ,with pitch variation being associated with screen Y-Axis touch positioning, duration associated with the touch time (from touch start to touch release), and note onset consequently tied to the touch in itself. Note amplitude showed no major identifiable trend, with some approaches separating themselves nonetheless (i.e., touch pressure, device Roll angle, and screen x-axis touch positioning). Considering the unavailability (at present) of the first and the lack of reliability of the second, we proposed that the third approach (x-axis positioning) be adopted as the ideal representation for amplitude variation. We have identified and validated an informed approach for setting up the most basic actions on mobile-based musical instrument operation actions. Thus, we open up new avenues to build upon and further move towards comprehensive and more complete creative approaches using MHD.

This article also presents the protocol behind this experiment, which we propose as a systematized way of evaluating the mapping of musical parameter variation in the context of smartphone operation. All materials adopted in the protocol are made available in open access to the community at https://zenodo.org/record/4553522 (accessed on 20 February 2021).

Our main original contributions are these guidelines for mapping the four studied musical parameters in MHD musical tools, supported by concrete testing and evaluation of user operation, and the test protocol allowing for systematized evaluation and definition of the links between interaction control gestures and musical parameters in the context of MHDs.

6.1. Future Work

Although the results show some discernible trends, there is still room for additional testing to further cement these findings and explore the aspects where this experiment failed to provide definitive results. The global context in which this experiment was implemented (amid a global pandemic) resulted in considerable difficulties regarding user participation and limited numbers. It would, then, be interesting to re-implement this experiment with a greater sample size to corroborate the identified trends.

Considering the discussed cultural and societal significance of musical tradition in the collected results, it would be interesting to implement this experiment in a context where musical tradition is very different from that of Western European culture. Even amongst non-musician participants, one can identify, as discussed, an almost ubiquitous influence of the piano (as discussed in Section 5.5) and Western musical staving, introducing a transverse bias across both profiles. Although this does not present any limitation on the proposed interaction guidelines, which are inherently tied to a cultural context, it would likely present an additional route in the study of MHD musical usage and would perhaps even allow for the creation of broader, more universal interaction guidelines, or would allow for defining ways of bridging cultural differences in these devices' musical appropriation.

One interesting question that arises from this experiment concerning the participants' gesture mappings relates to the underlying motivation for said mappings—e.g., did the participants' gestures attempt to simply replicate the sound they heard or was there a conscious association between musical gesture and musical outcome? Some participants' answers seem to indicate different approaches in this regard. We believe we successfully established links between musical parameter variation and gesture mapping, but the experiment protocol does not encompass the analysis and study of the reasoning behind said mappings. It would be enriching to the study and development of these guidelines to understand this underlying reasoning better.

As is the case for most experimental studies, there is always room for improvement and correction. Specifically, it would be essential to integrate some corrections to the protocol to minimize some of the problems encountered during this experiment. One of the most prevalent issues lay in the difficulty of perception of some musical parameters: whereas this is not problematic during phase 1 of the experiment and allows for important profile comparative analysis, it is a hindrance in phase 2, where participants were expected to map the studied musical parameters fully. This could be addressed in task 2 of phase 1 while reviewing the participant's performance. After going through the data-collection questionnaire, the researcher would explain in detail which parameters were changing to guarantee the participant was indeed fully aware of expectations before going on to phase 2 of the experiment.

As for the additional functionalities to be added to the tools provided, one has its foundations laid down already (although not directly related to the objective of providing the desired interaction guidelines). As referenced in Section 3, this functionality relates to usage of raw data concerning participants' interaction. These values were collected from the device's sensors and consist of all physical manipulation data (i.e., accelerometer values) and touchscreen operation details (e.g., touch number, touch time, touch coordinate path). These data open up a new avenue of testing in the field of gestural analysis, allowing researchers to analyze the specific physical interaction with the device and compare it to the described intended gestures. For example, one could analyze how long a participant held a touch on-screen and compare that to the actual note duration of the sound stimulus or analyze how participants define the scale reach for pitch mapping. These are just some examples of interesting points to study further.

Another tool is currently in development alongside these potential analyses taken from the raw data, consisting of a visualizer for all the logged events. Through 3D representation, this tool will take the logged data of a participant's performance and replicate the device's movements and touch behavior. This would allow researchers to visualize each participant's performance.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as no ethical issues were involved (e.g., no vulnerable populations, no collection of sensitive issues, no distress situations, invasive activities, or collection of biological materials).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the subjects to publish this paper.

Data Availability Statement: The data presented in this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.4553522.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MHD	Mobile Handheld Device
HCI	Human–Computer Interaction
PD	Pure Data
DMI	Digital Music Instrument

Appendix A. Experiment Data

Appendix A.1. Data Categorisation

As explained in Section 3.5, both gestures and selection rationale were combined into broader categories to allow for a more comprehensive array of statistical analyses. Table A1 shows the categorization of all gestures selected by participants throughout the experiment, and Table A2 shows the same for the gesture selection rationale.

Table A1. Categorization of participants' gestures in the experiment.

Category	Gesture
2D Plane manipulation	Screen axis (vertical)
	Screen axis (horizontal)
	Screen axis (diagonal)
	Screen axis (vertical)—auxiliary touch
	Multitouch
	Sequential touches
	Multitouch and Screen axis (vertical)
	Multitouch and Screen axis (horizontal)
Touch characteristics	Touch time
	Touch drag
	Touch pressure
	Touch area
	Touch time and touch drag
	Touch pressure and touch drag
	Multitouch and touch pressure
	Multitouch and touch area
Device position	Device position (vertical)
	Device position (horizontal)
	Device roll
	Device pitch
	Device movement time (general)
	Device movement time (horizontal)
	Device shake intensity

Table A1. *Cont.*

Category	Gesture
Combination	Device position (vertical) and touch area Multitouch and device position (vertical) Multitouch and device roll Screen axis (horizontal) and device pitch Touch pressure and device position (vertical) Touch pressure and device shake intensity Touch pressure and screen axis (horizontal) Touch pressure and screen axis (vertical) Touch time and device movement time (horizontal)

Table A2. Categorization of gesture choice rationale.

Category	Reason Provided by Participant
Instrument mimicking	Reproduce piano *scale* (sic) Reproduce piano keys Reproduce piano playing Reproduce guitar scale
Graphical representation	Behave like a slider Behave like a knob Represent waveform visualization (horizontal time) Represent a numerical scale
Intuition	Just felt right Intuition It felt more natural
Physical mapping	The notes felt higher/lower The notes felt fuller/thiner The sound sounded heavier/lighter I felt this parameter needed some haptic representation
Musical bias	Represent musical staff Horizontal scale organization
Exploration	I was testing it out I was experimenting
Unsure	Do not know why I was not able to map this successfully
User Experience	Used to this interaction from gaming Used to this interaction from MHD use
Complementing other mappings	Had to do this considering previous mappings
Using previous mappings	I used the same mappings I did before
Combining previous mappings	I tried mixing the mappings I did before with these
Instrument mimicking, physical mapping	Combination of a reason from both categories

Appendix A.2. Gesture Selection Frequencies—Uncategorised

This section lists the complete frequency tables for mapping gesture choices and gesture choice rationale and their distribution between participant profiles. For reference, as explained in Section 3, total $N = 27$, with non-musicians $n = 13$ and Musicians $n = 14$. The bottom row of all tables presents total counts for each profile and overall count. Wherever totals are different from expected n, participants failed to identify the particular parameter

variation in that stimulus, and their mappings were, consequently, not considered; e.g., in Table A4, Non-musician total is 9, with expected n = 13, meaning four participants of that profile were unable to perceive note duration variation in stimulus b.

Appendix A.2.1. Phase 1 Frequencies

This section lists the complete frequencies for uncategorized gesture mappings in phase 1 of the experiment.

Table A3. Pitch variation mapping (uncategorized)—stimulus a.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Screen axis (vertical)	8	61.5%	6	42.9%	14	51.9%
Screen axis (horizontal)	2	15.4%	3	21.4%	5	18.5%
Touch pressure	0	0.0%	1	7.1%	1	3.7%
Device position (vertical)	2	15.4%	2	14.3%	4	14.8%
Device angle (roll)	0	0.0%	1	7.1%	1	3.7%
Device position (vertical) and touch area	1	7.7%	0	0.0%	1	3.7%
Screen axis (horizontal) and device angle (pitch)	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A4. Duration variation mapping (uncategorized)—stimulus b.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Touch time	6	66.7%	12	85.7%	18	78.3%
Touch drag	1	11.1%	1	7.1%	2	8.7%
Device movement time (unconstrained)	2	22.2%	0	0.0%	2	8.7%
Touch time and device movement time (horizontal)	0	0.0%	1	7.1%	1	4.3%
	9	100.0%	14	100.0%	23	100.0%

Table A5. Amplitude variation mapping (uncategorized)—stimulus c.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Screen axis (vertical)	3	30.0%	4	33.3%	7	31.8%
Screen axis (vertical)—auxiliary touch	1	10.0%	0	0.0%	1	4.5%
Touch pressure	1	10.0%	3	25.0%	4	18.2%
Device position (vertical)	3	30.0%	1	8.3%	4	18.2%
Device angle (roll)	1	10.0%	1	8.3%	2	9.1%
Device angle (pitch)	0	0.0%	1	8.3%	1	4.5%
Device shake intensity	0	0.0%	1	8.3%	1	4.5%
Touch pressure and touch drag	1	10.0%	0	0.0%	1	4.5%
Touch pressure and device shake intensity	0	0.0%	1	8.3%	1	4.5%
	10	100.0%	12	100.0%	22	100.0%

Table A6. Pitch variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Screen axis (vertical)	8	61.5%	6	42.9%	14	51.9%
Screen axis (horizontal)	2	15.4%	3	21.4%	5	18.5%
Touch pressure	0	0.0%	1	7.1%	1	3.7%
Touch area	1	7.7%	0	0.0%	1	3.7%
Device position (vertical)	2	15.4%	2	14.3%	4	14.8%
Device angle (roll)	0	0.0%	1	7.1%	1	3.7%
Screen axis (horizontal) and device angle (pitch)	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A7. Duration variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Touch time	5	62.5%	11	84.6%	16	76.2%
Touch drag	0	0.0%	1	7.7%	1	4.8%
Device movement time (unconstrained)	2	25.0%	0	0.0%	2	9.5%
Touch time and touch drag	1	12.5%	0	0.0%	1	4.8%
Touch time and device movement time (horizontal)	0	0.0%	1	7.7%	1	4.8%
	8	100.0%	13	100.0%	21	100.0%

Table A8. Amplitude variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Screen axis (vertical)	3	33.3%	2	20.0%	5	26.3%
Screen axis (horizontal)	1	11.1%	0	0.0%	1	5.3%
Screen axis (vertical)—auxiliary touch	1	11.1%	0	0.0%	1	5.3%
Touch pressure	2	22.2%	3	30.0%	5	26.3%
Device position (vertical)	1	11.1%	0	0.0%	1	5.3%
Device position (horizontal)	1	11.1%	0	0.0%	1	5.3%
Device angle (roll)	0	0.0%	1	10.0%	1	5.3%
Device angle (pitch)	0	0.0%	1	10.0%	1	5.3%
Device shake intensity	0	0.0%	1	10.0%	1	5.3%
Touch pressure and device position (vertical)	0	0.0%	1	10.0%	1	5.3%
Touch pressure and device shake intensity	0	0.0%	1	10.0%	1	5.3%
	9	100.0%	10	100.0%	19	100.0%

Table A9. Polyphony mapping (uncategorized)—stimulus e.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Device position (vertical)	1	8.3%	0	0.0%	1	4.0%
Device angle (roll)	1	8.3%	0	0.0%	1	4.0%
Multitouch	1	8.3%	0	0.0%	1	4.0%
Multitouch and screen axis (vertical)	7	58.3%	6	46.2%	13	52.0%
Multitouch and screen axis (horizontal)	2	16.7%	4	30.8%	6	24.0%
Multitouch and device position (vertical)	0	0.0%	1	7.7%	1	4.0%
Multitouch and device angle (roll)	0	0.0%	1	7.7%	1	4.0%
Multitouch and touch pressure	0	0.0%	1	7.7%	1	4.0%
	12	100.0%	13	100.0%	25	100.0%

Appendix A.2.2. Phase 2 Frequencies

This section lists the complete frequencies for uncategorized gesture mappings in phase 2 of the experiment

Table A10. Pitch variation mapping (uncategorized)—stimulus a.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Screen axis (vertical)	7	53.8%	7	50.0%	14	51.9%
Screen axis (horizontal)	2	15.4%	3	21.4%	5	18.5%
Screen axis (diagonal)	1	7.7%	0	0.0%	1	3.7%
Touch pressure	0	0.0%	1	7.1%	1	3.7%
Device position (vertical)	2	15.4%	2	14.3%	4	14.8%
Device position (vertical) and touch area	1	7.7%	0	0.0%	1	3.7%
Screen axis (horizontal) and device angle (pitch)	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A11. Duration variation mapping (uncategorized)—stimulus b.

	Profile					
	Non-Musician		**Musician**		**Total**	
	N	**%**	**N**	**%**	**N**	**%**
Touch time	6	60.0%	13	92.9%	19	79.2%
Touch drag	1	10.0%	1	7.1%	2	8.3%
Device position (horizontal)	1	10.0%	0	0.0%	1	4.2%
Device movement time (unconstrained)	1	10.0%	0	0.0%	1	4.2%
Device movement time (horizontal)	1	10.0%	0	0.0%	1	4.2%
	10	100.0%	14	100.0%	24	100.0%

Table A12. Amplitude variation mapping (uncategorized)—stimulus c.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Screen axis (vertical)	3	30.0%	3	25.0%	6	27.3%
Screen axis (horizontal)	1	10.0%	0	0.0%	1	4.5%
Screen axis (vertical)—auxiliary touch	1	10.0%	0	0.0%	1	4.5%
Touch pressure	1	10.0%	3	25.0%	4	18.2%
Device position (vertical)	4	40.0%	1	8.3%	5	22.7%
Device angle (roll)	0	0.0%	1	8.3%	1	4.5%
Device shake intensity	0	0.0%	1	8.3%	1	4.5%
Touch pressure and device shake intensity	0	0.0%	1	8.3%	1	4.5%
Touch pressure and screen axis (horizontal)	0	0.0%	1	8.3%	1	4.5%
Touch pressure and screen axis (vertical)	0	0.0%	1	8.3%	1	4.5%
	10	100.0%	12	100.0%	22	100.0%

Table A13. Pitch variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Screen axis (vertical)	8	61.5%	7	50.0%	15	55.6%
Screen axis (horizontal)	2	15.4%	3	21.4%	5	18.5%
Touch pressure	0	0.0%	1	7.1%	1	3.7%
Touch area	1	7.7%	0	0.0%	1	3.7%
Device position (vertical)	2	15.4%	2	14.3%	4	14.8%
Screen axis (horizontal) and device angle (pitch)	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A14. Duration variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Touch time	6	66.7%	12	92.3%	18	81.8%
Touch drag	0	0.0%	1	7.7%	1	4.5%
Device movement time (unconstrained)	2	22.2%	0	0.0%	2	9.1%
Touch time and touch drag	1	11.1%	0	0.0%	1	4.5%
	9	100.0%	13	100.0%	22	100.0%

Table A15. Amplitude variation mapping (uncategorized)—stimulus d.

	Profile					
	Non-Musician		Musician		Total	
	N	%	N	%	N	%
Screen axis (vertical)	2	22.2%	2	20.0%	4	21.1%
Screen axis (horizontal)	2	22.2%	0	0.0%	2	10.5%
Screen axis (vertical)—auxiliary touch	1	11.1%	0	0.0%	1	5.3%

Table A15. *Cont.*

	Profile					
	Non-Musician		Musician		Total	
	N	**%**	**N**	**%**	**N**	**%**
Touch pressure	2	22.2%	2	20.0%	4	21.1%
Device position (vertical)	1	11.1%	0	0.0%	1	5.3%
Device position (horizontal)	1	11.1%	0	0.0%	1	5.3%
Device angle (roll)	0	0.0%	1	10.0%	1	5.3%
Device shake intensity	0	0.0%	1	10.0%	1	5.3%
Touch pressure and device position (vertical)	0	0.0%	1	10.0%	1	5.3%
Touch pressure and device shake intensity	0	0.0%	1	10.0%	1	5.3%
Touch pressure and screen axis (horizontal)	0	0.0%	1	10.0%	1	5.3%
Touch pressure and screen axis (vertical)	0	0.0%	1	10.0%	1	5.3%
	9	100.0%	10	100.0%	19	100.0%

Table A16. Polyphony mapping (uncategorized)—stimulus e.

	Profile					
	Non-Musician		Musician		Total	
	N	**%**	**N**	**%**	**N**	**%**
Device position (vertical)	2	16.7%	0	0.0%	2	7.7%
Sequential touches	0	0.0%	1	7.1%	1	3.8%
Multitouch and screen axis (vertical)	7	58.3%	6	42.9%	13	50.0%
Multitouch and screen axis (horizontal)	2	16.7%	3	21.4%	5	19.2%
Multitouch and touch area	1	8.3%	0	0.0%	1	3.8%
Multitouch and device position (vertical)	0	0.0%	2	14.3%	2	7.7%
Multitouch and device angle (roll)	0	0.0%	1	7.1%	1	3.8%
Multitouch and touch pressure	0	0.0%	1	7.1%	1	3.8%
	12	100.0%	14	100.0%	26	100.0%

Appendix A.3. Mapping Rationale

Appendix A.3.1. Phase 1

This section lists the complete frequencies for gesture mapping rationale in phase 1 of the experiment.

Table A17. Mapping rationale—stimulus a.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	4	30.8%	7	50.0%	11	40.7%
Graphical representation	2	15.4%	1	7.1%	3	11.1%
Intuition	3	23.1%	1	7.1%	4	14.8%
Physical mapping	1	7.7%	0	0.0%	1	3.7%
Musical bias	2	15.4%	4	28.6%	6	22.2%
Complementing other mappings	1	7.7%	0	0.0%	1	3.7%
Instrument mimicking and physical mapping	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A18. Mapping rationale—stimulus b.

	Profile					
	Non-Musician		**Musician**		**Total**	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	6	66.7%	11	78.6%	17	73.9%
Graphical representation	1	11.1%	1	7.1%	2	8.7%
Intuition	2	22.2%	1	7.1%	3	13.0%
User Experience	0	0.0%	1	7.1%	1	4.3%
	9	100.0%	14	100.0%	23	100.0%

Table A19. Mapping rationale—stimulus c.

	Profile					
	Non-Musician		**Musician**		**Total**	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	3	30.0%	4	33.3%	7	31.8%
Graphical representation	4	40.0%	4	33.3%	8	36.4%
Intuition	3	30.0%	2	16.7%	5	22.7%
Physical mapping	0	0.0%	1	8.3%	1	4.5%
Exploration	0	0.0%	1	8.3%	1	4.5%
	10	100.0%	12	100.0%	22	100.0%

Table A20. Mapping rationale—stimulus d.

	Profile					
	Non-Musician		**Musician**		**Total**	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	1	7.7%	0	0.0%	1	3.7%
Intuition	1	7.7%	0	0.0%	1	3.7%
Unsure	1	7.7%	0	0.0%	1	3.7%
User Experience	1	7.7%	0	0.0%	1	3.7%
Using previous mappings	0	0.0%	1	7.1%	1	3.7%
Combining previous mappings	9	69.2%	13	92.9%	22	81.5%
	13	100.0%	14	100.0%	27	100.0%

Table A21. Mapping rationale—stimulus e.

	Profile					
	Non-Musician		**Musician**		**Total**	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	2	16.7%	1	7.1%	3	11.5%
Intuition	3	25.0%	0	0.0%	3	11.5%
Exploration	0	0.0%	1	7.1%	1	3.8%
Unsure	1	8.3%	1	7.1%	2	7.7%
Using previous mappings	6	50.0%	11	78.6%	17	65.4%
	12	100.0%	14	100.0%	26	100.0%

Appendix A.3.2. Phase 2

This section lists the complete frequencies for gesture mapping rationale in phase 2 of the experiment.

Table A22. Mapping rationale—stimulus a.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	4	30.8%	8	57.1%	12	44.4%
Graphical representation	2	15.4%	1	7.1%	3	11.1%
Intuition	3	23.1%	1	7.1%	4	14.8%
Physical mapping	1	7.7%	0	0.0%	1	3.7%
Musical bias	2	15.4%	3	21.4%	5	18.5%
Complementing other mappings	1	7.7%	0	0.0%	1	3.7%
Instrument mimicking and physical mapping	0	0.0%	1	7.1%	1	3.7%
	13	100.0%	14	100.0%	27	100.0%

Table A23. Mapping rationale—stimulus b.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	6	60.0%	12	85.7%	18	75.0%
Graphical representation	2	20.0%	0	0.0%	2	8.3%
Intuition	2	20.0%	1	7.1%	3	12.5%
User Experience	0	0.0%	1	7.1%	1	4.2%
	10	100.0%	14	100.0%	24	100.0%

Table A24. Mapping rationale—stimulus c.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	3	30.0%	6	50.0%	9	40.9%
Graphical representation	2	20.0%	3	25.0%	5	22.7%
Intuition	3	30.0%	1	8.3%	4	18.2%
Physical mapping	1	10.0%	1	8.3%	2	9.1%
Exploration	0	0.0%	1	8.3%	1	4.5%
Combining previous mappings	1	10.0%	0	0.0%	1	4.5%
	10	100.0%	12	100.0%	22	100.0%

Table A25. Mapping rationale—stimulus d.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	1	7.7%	0	0.0%	1	3.7%
Intuition	1	7.7%	0	0.0%	1	3.7%
Unsure	1	7.7%	0	0.0%	1	3.7%
User Experience	1	7.7%	0	0.0%	1	3.7%
Using previous mappings	0	0.0%	1	7.1%	1	3.7%
Combining previous mappings	9	69.2%	13	92.9%	22	81.5%
	13	100.0%	14	100.0%	27	100.0%

Table A26. Mapping rationale—stimulus e.

	Profile					
	Non-Musician		Musician		Total	
Mapping Reason	**N**	**%**	**N**	**%**	**N**	**%**
Instrument mimicking	2	16.7%	1	7.1%	3	11.5%
Intuition	2	16.7%	0	0.0%	2	7.7%
Exploration	0	0.0%	1	7.1%	1	3.8%
Unsure	1	8.3%	0	0.0%	1	3.8%
Using previous mappings	7	58.3%	11	78.6%	18	69.2%
Combining previous mappings	0	0.0%	1	7.1%	1	3.8%
	12	100.0%	14	100.0%	26	100.0%

Appendix A.3.3. Profile Association Results

This section lists the complete participant profile association χ^2 test results for gesture mapping (categorized and uncategorized).

Table A27. Phase 1 profile associations. NS: Value is constant.

Stimulus	Parameter	Variation Perception	Gesture Mapping	Mapping Reason	Mapping Change
a	Pitch	NS	$\chi^2(6) = 4.46, p = 0.62$	$\chi^2(6) = 5.79, p = 0.45$	–
b	Duration	$\chi^2(1) = 5.06, p = 0.04$	$\chi^2(3) = 4.11, p = 0.25$	$\chi^2(2) = 3.90, p = 0.14$	–
c	Amplitude	$\chi^2(1) = 0.35, p = 0.65$	$\chi^2(8) = 7.02, p = 0.54$	$\chi^2(4) = 2.18, p = 0.70$	–
d	Pitch	NS	$\chi^2(6) = 4.46, p = 0.62$	$\chi^2(10) = 9.37, p = 0.50$	$\chi^2(1) = 1.12, p = 0.29$
	Duration	$\chi^2(1) = 3.83, p = 0.08$	$\chi^2(4) = 6.42, p = 0.17$		$\chi^2(1) = 1.36, p = 0.33$
	Amplitude	$\chi^2(1) = 0.02, p = 0.90$	$\chi^2(10) = 9.37, p = 0.50$		$\chi^2(1) = 0.31, p = 0.68$
e	Polyphony	$\chi^2(7) = 1.12, p = 0.48$	$\chi^2(7) = 6.71, p = 0.46$	$\chi^2(4) = 5.68, p = 0.22$	-

Table A28. Phase 2 profile associations. NS: Value is constant.

Stimulus	Parameter	Variation Perception	Gesture Mapping	Mapping Reason	Mapping Change	Phase Change
a	Pitch	NS	$\chi^2(6) = 4.17, p = 0.65$	$\chi^2(6) = 5.83, p = 0.44$	–	$\chi^2(1) = 0.30, p = 1.00$
b	Duration	$\chi^2(1) = 3.64, p = 0.10$	$\chi^2(4) = 5.05, p = 0.282$	$\chi^2(3) = 4.80, p = 0.19$	–	$\chi^2(1) = 0.46, p = 0.60$
c	Amplitude	$\chi^2(1) = 0.35, p = 0.56$	$\chi^2(9) = 9.70, p = 0.38$	$\chi^2(5) = 4.05, p = 0.54$	–	$\chi^2(1) = 0.02, p = 1.00$
d	Pitch	NS	$\chi^2(5) = 3.23, p = 0.66$	$\chi^2(5) = 5.70, p = 0.34$	$\chi^2(1) = 2.33, p = 0.22$	$\chi^2(1) = 2.01, p = 0.48$
	Duration	$\chi^2(1) = 2.49, p = 0.17$	$\chi^2(3) = 5.45, p = 0.14$		$\chi^2(1) = 2.50, p = 0.17$	$\chi^2(1) = 0.003, p = 1.00$
	Amplitude	$\chi^2(1) = 0.02, p = 1.00$	$\chi^2(11) = 10.98, p = 0.45$		$\chi^2(1) = 0.31, p = 0.68$	$\chi^2(1) = 1.01, p = 0.60$
e	Polyphony	$\chi^2(1) = 1.12, p = 0.48$	$\chi^2(7) = 8.17, p = 0.32$	$\chi^2(5) = 6.10, p = 0.30$	–	$\chi^2(1) = 0.16, p = 1.00$

Table A29. Categorized gesture mappings profile associations.

Stimulus	Parameter	Phase 1	Phase 2
a	Pitch	$\chi^2(3) = 1.22, p = 0.75$	$\chi^2(3) = 0.96, p = 0.81$
b	Duration	$\chi^2(3) = 3.90, p = 0.14$	$\chi^2(3) = 4.80, p = 0.06$
c	Amplitude	$\chi^2(3) = 1.03, p = 0.80$	$\chi^2(3) = 4.50, p = 0.21$
d	Pitch	$\chi^2(3) = 1.22, p = 0.75$	$\chi^2(3) = 0.97, p = 0.81$
	Duration	$\chi^2(3) = 4.04, p = 0.13$	$\chi^2(3) = 3.18, p = 0.16$
	Amplitude	$\chi^2(3) = 3.64, p = 0.303$	$\chi^2(3) = 5.25, p = 0.15$
e	Polyphony	$\chi^2(3) = 4.97, p = 0.17$	$\chi^2(3) = 4.93, p = 0.18$

Appendix B. Experiment questionnaires

Appendix B.1. Pre-Experiment Questionnaire

These were the questions asked to participants to establish their participation eligibility and define their musical proficiency profile.

1. **What is your age?**
   20–30 years old: proceed to question 2
   Other: Not eligible for the experiment
2. **Do you use a smartphone or tablet every day?**
   Yes: Proceed to question 3
   No: Not eligible for the experiment
3. **Do you regularly do any of the following tasks on your device (at least once per day)?**
   - **E-book or office document reading or creation;**
   - **Photo editing;**
   - **Gaming.**

   Yes: Proceed to question 4
   No: Not eligible for the experiment.
4. **Do you play or have in the past played any instrument (including singing or digital instruments)?**
   Yes: Proceed to question 5
   No: Proceed to question 7
5. **Considering your most active time, how many days a week do/did you play or practice per day?**
   Two *or more times a week*: Musician
   Under two times a week: Proceed to question 6
6. **How long has it been since you last played/practiced your instrument?**
   Over five years: Non-musician
   Under five years: Musician
7. **What is the maximum level of formal musical training you have?**
   High school music classes or lower: Proceed do question 8
   Conservatory or college-level classes: Musician
8. **Roughly how many hours would you say you listen to music during one week?**
   Over three hours/week: Non-musician
   Under three hours/week: Not eligible for the experiment

Appendix B.2. Post-Task Questionnaire

These were the questions asked to participants while reviewing the video recording of their test performance. The questions were asked for each of the sound stimuli the participant had to attempt to reproduce.

1. "In this sound example, what variation did you perceive between the notes, if any?"
2. "How did you try to represent that variation as a gesture?"
3. "Why did you feel this was the most adequate choice?"

References

1. Levin, G. DIALTONES (A TELESYMPHONY). 2001. Available online: http://www.flong.com/storage/experience/telesymphony/index.html (accessed on 2 February 2021).
2. Han, Q.; Cho, D. Characterizing the technological evolution of smartphones: Insights from performance benchmarks. In Proceedings of the ACM International Conference Proceeding Series, Suwon, Korea, 17–19 August 2016; [CrossRef]
3. Holst, A. Number of Smartphone Users Worldwide from 2016 to 2021 (in Billions). 2019. Available online: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide (accessed on 2 February 2021).
4. Brinkmann, P.; Mccormick, C.; Kirn, P.; Roth, M.; Lawler, R. Embedding Pure Data with libpd. In Proceedings of the Fourth International Pure Data Convention, Weimar, Germany, 8–12 August 2011; pp. 291–301.
5. Clément, A.R.; Ribeiro, F.; Rodrigues, R.; Penha, R. Bridging the gap between performers and the audience using networked smartphones: The a. bel system. In Proceedings of the of ICLI 16, Brighton, United Kingdom, 29 June–3 July 2016.
6. Clément, A.; Rodrigues, R.; Penha, R. Tools and Template Development for Live Networked Musical Performance System. In Proceedings of the 1st Doctoral Congress in Engineering, Porto, Portugal, 11–12 June 2015.
7. Tanaka, A.; Parkinson, A.; Settel, Z.; Tahiroglu, K. A Survey and Thematic Analysis Approach as Input to the Design of Mobile Music GUIs. In Proceedings of the International Conference on New Interfaces for Musical Expression, Ann Arbor, Michigan, USA, 21–23 May 2012; pp. 200–203.
8. Essl, G.; Lee, S.W. Mobile Devices as Musical Instruments -State of the Art and Future Prospects. In Proceedings of the 13th International Symposium on CMMR, Matosinhos, Portugal, 25–28 September 2017; pp. 364–375.
9. Turchet, L. Smart Musical Instruments: Vision, Design Principles, and Future Directions. *IEEE Access* **2019**, *7*, 8944–8963. [CrossRef]
10. Magnusson, T. Affordances and constraints in screen-based musical instruments. In Proceedings of the 4th Nordic Conference on Human-Computer Interaction, Oslo, Norway, 14–18 October 2006; Volume 189, pp. 441–444. [CrossRef]
11. Papetti, S.; Fröhlich, M.; Schiesser, S. The TouchBox: An open-source audio-haptic device for finger-based interaction. In Proceedings of the IEEE World Haptics Conference, Tokyo, Japan, 9–12 July 2019; pp. 491–496. [CrossRef]
12. Tanaka, A. Mapping Out Instruments, Affordances, and Mobiles. In Proceedings of the International Conference on New Interfaces for Musical Expression, Sydney, Australia, 15–18 June 2010; pp. 15–18.
13. Maes, P.J.; Leman, M.; Lesaffre, M.; Demey, M.; Moelants, D. From expressive gesture to sound: The development of an embodied mapping trajectory inside a musical interface. *J. Multimodal User Interfaces* **2010**, *3*, 67–78. [CrossRef]
14. Magnusson, T. *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions*; Bloomsbury Academic: New York, NY, USA, 2019. [CrossRef]
15. Magnusson, T. Designing Constraints: Composing and Performing with Digital Musical Systems. *Comput. Music. J.* **2010**, *34*, 62–73. [CrossRef]
16. Hunt, A.; Wanderley, M.M.; Paradis, M. The Importance of Parameter Mapping in Electronic Instrument Design. In Proceedings of the 2002 Conference on New Interfaces for Musical Expression, Dublin, Ireland, 24–26 May 2002; Volume 32, pp. 149–154. [CrossRef]
17. Hunt, A.; Wanderley, M.M. Mapping performer parameters to synthesis engines. *Organised Sound* **2003**, *7*, 97–108. [CrossRef]
18. Gillian, N.E. Gesture Recognition for Musician Computer Interaction Doctoral dissertation, Faculty of Arts, Humanities and Social Sciences, Queen's University, Belfast, Ireland. Available online http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.725.1727&rep=rep1&type=pdf (accessed on 2 February 2021).
19. Cont, A.; Coduys, T.; Henry, C. Real-time gesture mapping in PD environment using neural networks. In Proceedings of the 2004 Conference on New Interfaces for Musical Expression, Hamamatsu, Japan, 3–5 June 2004; pp. 39–42.
20. Michon, R.; Smith, J.O.; Wright, M.; Chafe, C.; Granzow, J.; Wang, G. Mobile music, sensors, physical modeling, and digital fabrication: Articulating the augmented mobile instrument. *Appl. Sci.* **2017**, *7*, 1311. [CrossRef]
21. INFOGRAPHIC: How Computing Power Has Changed over Time. Available online: https://www.businessinsider.com/infographic-how-computing-power-has-changed-over-time-2017-11 (accessed on 2 February 2021).
22. Mobile CPUs Are Now as Fast as Most Desktop PCs. Available online: https://www.howtogeek.com/393139/mobile-cpus-are-now-as-fast-as-your-desktop-pc/ (accessed on 19 February 2021).
23. Jordà, S. Digital Instruments and Players : Part I—Efficiency and Apprenticeship. In Proceedings of the 2004 Conference on New Interfaces for Musical Expression, Hamamatsu, Japan, 3–5 June 2004; pp. 59–63.
24. This App Blows! (Ocarina 2 Launches Today)—Smule. Available online: https://blog.smule.com/this-app-blows-ocarina-2-launches-today (accessed on 19 February 2021).
25. The Motion Synth: Turn Movement into Music by AUUG. Available online: https://www.kickstarter.com/projects/1892750571/the-motion-synth-turn-movement-into-music (accessed on 19 February 2021).

26. 9 Artists Pioneering the Art of the Musical Mobile App. Available online: https://flypaper.soundfly.com/discovery/artists-pioneering-musical-mobile-apps/ (accessed on 19 February 2021).

27. Software by Miller Puckette. Available online: http://msp.ucsd.edu/software.html (accessed on 2 February 2021).

28. Cifter, a.S.; Dong, H. User characteristics: Professional vs. lay users. In Proceedings of the fifth international conference on inclusive design, London, United Kingdom, 8–10 April 2009.

29. Pritschet, L.; Powell, D.; Horne, Z. Marginally Significant Effects as Evidence for Hypotheses: Changing Attitudes Over Four Decades. *Psychol. Sci.* **2016**, *27*, 1036–1042. [CrossRef] [PubMed]

30. Olsson-Collentine, A.; van Assen, M.A.; Hartgerink, C.H. The Prevalence of Marginally Significant Results in Psychology over Time. *Psychol. Sci.* **2019**, *30*, 576–586. [CrossRef]

31. Summary of MIDI 1.0 Messages. Available online: https://www.midi.org/specifications-old/item/table-1-summary-of-midi-message (accessed on 2 February 2021).

32. Hwang, S.; Bianchi, A.; Wohn, K.Y. VibPress: Estimating pressure input using vibration absorption on mobile devices. In Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, Munich, Germany, 27–30 August 2013; pp. 31–34. [CrossRef]

33. Winter, A.E.; Cox, B.R.; Ginn, L.K.E.; Whitt, D.O.; Fitz-Coy, A.A.; Picciotto, C.E.; Yun, G.G.; Nelson, J.J. Input Device Haptics and Pressure Sensing, United States of America, Patent No.: US 9448631 B2. Available online: https://patentcenter.uspto.gov/#!/applications/14698318 (accessed on 2 February 2021).

34. Tung, Y.C.; Shin, K.G. ForcePhone: Software Lets Smartphones Sense Touch Force. *IEEE Pervasive Comput.* **2016**, *15*, 20–25. [CrossRef]

35. Kuhlmann, T.; Garaizar, P.; Reips, U.D. Smartphone sensor accuracy varies from device to device in mobile research: The case of spatial orientation. *Behav. Res. Methods* **2020**, *53*, 22–33. [CrossRef]