# End-User Action–Sound Mapping Design for Mid-Air Music Performance

Dominic Brown

A Thesis submitted in partial fulfilment of the requirements of the University of the West of England for the degree of Doctor of Philosophy

Department of Computer Science and Creative Technologies
Faculty of Environment and Technology
University of the West of England, Bristol
June 2020

# Abstract

The mapping between a performer's actions and an instrument's sound response is a consistent theme in Digital Musical Instrument (DMI) research. Previously, mapping was an activity for creators, but more recent work has exposed mapping design to musicians through dedicated software, democratising mapping design and providing a new avenue for creative expression.

DMIs often suffer from a lack of practitioners beyond their designer, and few are used by musicians in professional performance contexts. The Mi.Mu Gloves are one example that is used by several musicians in their professional practice. The research in this dissertation investigates end-user mapping practice with the Gloves, and the influences on Glove musicians' design decisions as their practice develops, examining the question: *How do end-users of a glove-based mid-air DMI design action–sound mapping strategies for musical performance?*

A study of the mapping practice of existing Glove musicians revealed differences in the mapping of experienced and novice musicians, with novices evoking conceptual metaphors of movement and music, while experienced musicians designed ergonomic mappings that minimised performer error.

An examination of the initial development period of mapping practice also found that novices used conceptual metaphors, with transparency and the audience's perception being important factors. Creative mapping was hindered by system reliability and poorly trained posture recognition.

An investigation into the practice of expert glove musicians supported observations that they focus on error minimisation and ergonomic controls, but also found that they embellished these simple controls with ancillary gestures to communicate aesthetic meaning. They also suffered from system reliability, and had developed techniques to mitigate them.

A study into the effects of this system-related error on skill acquisition with the gloves found that a relatively small rate of system error had a significant effect.

Finally, design heuristics applicable to DMI design, mid-air interaction design and end-user mapping design are presented.

# Acknowledgements

Like all PhD candidates, I owe a debt of gratitude to many people who have encouraged, critiqued, supported and inspired me throughout my time studying for this degree.

I would like to thank my supervisors Dr Tom Mitchell and Dr Chris Nash and my director of studies Professor Larry Bull for their guidance, encouragement and knowledge; Dr Andrew McPherson and Professor Jim Smith for their insights, analysis and critique during the examination process, and Professor Aniko Varadi for chairing; Sam Hunt, for his camaraderie and companionship along our PhD journeys; Dr Adam Stark at Mi.Mu Ltd and Mehdi Sobhani at the Bristol Robotics Laboratory for their technical expertise and assistance; as well as the staff and students of the Music Technology department at UWE.

I would also like to thank my family: Mum, Dad, Michael, James, Ella, Poppy and Hattie for their endless love and emotional support, as well as my good friends: Ash Mukherjee, Tom Sanger, James White, Tom Nicholas, Dan Mills, Josie Goodinson, Mikey Burton, Charlotte Barnes, Henry Roper, Will Gould, Tommie James, Harry Willmott, Jed Milton and many others who have provided love and support over the years. Also, I would like to thank my fellow masochists at UWE's Doctoral Society: Dr Richard Mayne, Dr Ola Michalec, Dr Hannah Stott, Dr Abdullah Khatib, Dr Vevila Dornelles, Dr Beth Jones, David Richter, Dr Leah Morris and many others for their fellowship throughout the PhD process.

A special thanks is due for The South West Creative Technology Network and those involved in the Immersion Fellowship for their provoking discussions and sharing of knowledge, stories, ideas and experiences.

I am also indebted to the doctors, nurses and staff at Southmead Hospital's Dermatology and Plastic Surgery departments, the Teenage Cancer Trust and Macmillan Cancer Support for their expert and continuing care.

Lastly I would like to thank Dr David Creasey, who encouraged me to pursue a PhD in the first place, and for his guidance and occasional grilling along the way.

# Declarations

This thesis contains content included in the following publications.

In Chapter 4:

Brown, D., Nash, C. and Mitchell, T. (2017) A User Experience Review of Music Interaction Evaluations. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 370–375.

In Chapters 5 and 6:

Brown, D., Nash, C. and Mitchell, T. (2018) Understanding User-Defined Mapping Design for Mid-Air Musical Performance. In: *Proceedings of the 5th International Conference on Movement and Computing (MOCO)*. Genoa, Italy, June 2018. ACM. 27:1–27:8.

In Chapters 5 and 8:

Brown, D., Nash, C. and Mitchell, T. (2018) Simple Mappings, Expressive Movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity*. 29 (2–3), pp. 129–148.

Other publications that have informed and guided the research.

Brown, D., Renney, N., Stark, A., Nash, C. and Mitchell, T. (2016) Leimu: Gloveless Music Interaction using a Wrist Mounted Leap Motion. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2016*. Brisbane, Australia, July 2016. pp. 300–304.

Brown, D., Nash, C. and Mitchell, T. (2016) GestureChords: Transparency in Gesturally Controlled Digital Musical Instruments through Iconicity and Conceptual Metaphor. In: *Proceedings of Sound and Music Computing (SMC) 2016*. Hamburg, Germany, Sept. 2016. pp. 85–92.

Brown, D. (2016) The Application of Established Gestural Languages in the Control Mappings of Free-hand Gestural Musical Instruments. In: *Proceedings of the International Conference on Live Interfaces (ICLI) 2016*. Brighton, UK, July 2016. pp. 181–186.

# Contents

# List of Figures

# List of Tables

# Acronyms

**DMI** Digital Musical Instrument.

**ML** Machine Learning.

**HMM** Hidden Markov Model.

**DTW** Dynamic Time Warping.

**HCI** Human-Computer Interaction.

**IMU** Inertial Measurement Unit.

**TUI** Tangible User Interface.

**UX** User Experience.

**NIME** New Interface for Musical Expression.

**NIME** The International Conference on New Interfaces for Musical Expression.

**SMC** The Sound and Music Computing Conference.

**ICMC** The International Computer Music Conference.

**MMI** Mimetic Motor Imagery.

**MMA** Mimetic Motor Action.

**DODE** Domain-Oriented Design Environment.

**DAW** Digital Audio Workstation.

**MSI** Musical Sophistication Index.

**GUI** Graphical User Interface.

**CHI** The ACM CHI Conference on Human Factors in Computing Systems.

**XML** Extensible Mark-up Language.

**OSC** Open Sound Control.

# 1. Introduction

Technology has historically shaped the aesthetics, production and dissemination of music. Breakthroughs in craftsmanship, manufacturing and materials gave rise to the earliest acoustic instruments and, more recently, the most sophisticated digital synthesisers. The microprocessor sparked an unprecedented period of innovation, which has impacted every dimension of human existence, and developments in Human-Computer Interaction (HCI) and gestural interaction technology is redefining the ways in which humans interact with music.

As Digital Musical Instruments (DMIs) become ever more technologically sophisticated, their creation more democratic, and their adoption more widespread, an increasingly relevant line of inquiry in DMI research is how a DMI can effectively facilitate a musician's musical expression, to the extent that the musician integrates the instrument into their long-term, professional creative practice.

One such DMI that is succeeding at forming a community of dedicated, professional practitioners is the Mi.Mu Gloves: a pair of wearable gestural controllers that enable musicians to control music through mid-air movements. A significant aspect of Mi.Mu Glove musicianship is end-user mapping design, where a dedicated software application allows glove musicians to define their own connections, or *mappings*, between their actions and the resulting sound.

Established thought around mapping in DMI design is that it has a significant impact on an instrument's ability to facilitate musical expression (see Section 2.5.1).

If, as with the Mi.Mu Gloves, end-user musicians of DMIs are able to design mappings to facilitate their musical expression, an important question must be considered: How do glove users design action–sound mapping strategies for musical performance?

The research presented in this dissertation seeks to address this, through examining the end-user mapping design practice of existing glove users, tracking mapping practice over the initial development of glove musicianship, exploring expert-level glove musicianship in professional musical practice, and analysing the effects of system reliability on skill acquisition.

## 1.1. Mid-Air Digital Musical Instruments

DMIs differ from their acoustic counterparts in that they do not rely on the physical excitation of resonant materials by a player to produce sound, rather they use computer systems that are able to sense the player's interactions and synthesise an auditory response. Consequently, a wide variety of interaction methods have been proposed, broadly fitting into four categories (Miranda and Wanderley, 2006):

- *Augmented instruments.* Acoustic instruments that have additional electronic sensors and controls, such as the Hyper-Flute (Palacio-Quintin, 2003);
- *Instrument-like gestural controllers.* Designed to mimic existing acoustic instruments as closely as possible, such as typical MIDI keyboards and the Akai EWI (Rovan *et al.*, 1997);
- *Instrument-inspired gestural controllers.* Influenced by existing instruments but do not replicate them, such as the Sequential Drum (Mathews and Abbott, 1980);
- *Alternate gestural controllers.* Controllers not modelled on any particular instrument, instead seeking to explore new methods of musical interaction.

For augmented, instrument-like and instrument-inspired DMIs, existing acoustic instruments are used to inform their design. For alternate gestural controller DMIs that seek to move away from existing instruments, it can often be the case that no acoustic musical instrument counterpart exists, and as such their design can be more challenging, but also more diverse, due to the plethora of interaction interfaces and techniques that can be used in these instruments.

One method often used in alternate gestural controllers is mid-air interaction, where actions are expressed through the articulation of body movements rather than the manipulation of physical, tangible objects. This is most commonly achieved either through cameras and computer-vision based algorithms (Aggarwal and Cai, 1999) or through wearable solutions such as data gloves (Sturman and Zeltzer, 1994). The ability to manipulate music through the unencumbered movements of one's body has inspired creators of musical instruments for nearly a century, with mid-air interactions being used in the Theremin, and by many DMI creators since: Michel Waisvisz's *The Hands* (Torre, Andersen and Baldé, 2016; Waisvisz, 1985), Laetitia Sonami's *Lady's Glove* (Sonami, n.d.) and Imogen Heap's *Mi.Mu Gloves* (Mi.Mu Ltd., 2018), for example.

Within DMI research, discussed in detail in Chapter 2, a *mapping* refers to the relationship between the player's actions (or gestures) and the instrument's auditory response (Arfib *et al.*, 2002; Rovan *et al.*, 1997; Winkler, 1995), and plays an important role in defining the characteristics of a DMI (Hunt, Wanderley and Paradis, 2003).

Mapping has been the subject of much research (Visi *et al.*, 2017; Tanaka, 2010; Paine, Stevenson and Pearce, 2007; Arfib *et al.*, 2002), and the challenge of designing mapping strategies has previously been framed as *the mapping problem* (Hunt, Wanderley and Paradis, 2003), as problems can arise if poor mapping choices are made: overly simple mappings can lead to boring instruments that musicians soon grow tired of (Hunt, Wanderley and Paradis, 2003), while complex mappings that fail to provide a perceivable, “correct” response to a musician’s actions can result in perceptual disconnection (Nakra, 2002).

This perceived “correctness” is rooted in the embodied link between gestures and music; that is, the musical meaning that humans naturally ascribe to particular physical movements through their knowledge and experience of the physical world (Cox, 2016; Johnson and Larson, 2003), discussed in more detail in Section 3.2.

Several projects are exposing mapping design to the end-users of DMIs, by either making physically hackable instruments like the D-Box (Zappi and McPherson, 2015), or by providing Graphical User Interfaces (GUIs) that expose interface and synthesis parameters and allow users to design their own action–sound connections (Di Donato, 2017; Fiebrink, Trueman and Cook, 2009).

Allowing the end-users of a musical instrument to create their own mappings provides researchers with an opportunity to gain insights into how DMI mappings are approached from the perspective of the performer, and how the context of a musician’s practice influences their mapping decisions.

## 1.2. Research Questions

This research seeks to investigate end-user mapping design with a mid-air DMI within the context of designing mappings for music performance, to provide new insights into DMI mapping, end-user mapping design tools and mid-air interaction.

The research question this dissertation explores is: How do end-users of a glove-based mid-air DMI design action–sound mapping strategies for musical performance?

The research focuses on musicians using a mid-air data glove interface: the Mi.Mu Gloves (Mitchell, Madgwick and Heap, 2012; Mitchell and Heap, 2011), and its dedicated mapping design software: Glover, both described in Chapter 5.

To address this research question, the following questions are considered.

1. What factors influence mapping design decisions for glove-based music performance?
2. How does mapping practice develop as glove musicians gain experience?
3. How do these factors impact a musician's ability to acquire skill with glove instruments?

## 1.3. Methodological Approach

To explore these research questions, a mixed-methods approach was taken. Quantitative observations obtained through task-based interaction studies into mapping practice are supported by qualitative observations of a musician's creative practice in-the-world. This approach is influenced by User Experience (UX) research methods (Section 4.2) and previous work in the study of creative practice (Nash, 2011; Collins, 2005).

## 1.4. Contributions

This thesis makes the following research contributions.

1. *Empirical findings.* A series of empirical findings into end-user mapping design, focusing on the factors that influence and affect mapping choices in the context of musical performance.
2. *Design Guidelines.* A set of recommendations, drawn from the empirical findings, for designing mid-air DMIs as well as end-user mapping tools.

3. *Methodological Review.* Insights into the methodological approaches used within the DMI field and recommendations for future research.

These contributions provide insights relevant to DMI design, mid-air interaction within the HCI field, and the use of software in performance practice.

## 1.5. Structure of Dissertation

This dissertation (summarised in Table 1.1) first presents background on DMIs and mapping (Chapter 2), embodied cognition and the relation between gesture and sound (Chapter 3), and a systematic review of UX research methods and their use in the DMI research field (Chapter 4). The Mi.Mu Gloves and their dedicated mapping application Glover are then presented in detail (Chapter 5), followed by empirical studies into end-user mapping practice with the system, looking at how existing users design mappings for a given task (Chapter 6), how mapping design develops over the initial learning period of glove musicianship (Chapter 7), and how expert-level glove musicians design mappings in their personal creative practice (Chapter 8). Following from insights gained in these chapters, an empirical study into the effects of system-related error on skill acquisition with the gloves is presented (Chapter 9). Finally, a series of design heuristics and recommendations, applicable to the design of future DMIs and mid-air interaction tools, are presented (Chapter 10).

**Introduction**

1: *Introduction.*
Frames and describes the research question investigated in the empirical work of this dissertation.

**Background**

2: *Digital Musical Instruments.*
Presents literature from the research field of DMI design, focusing on mapping.

3: *Movement, Music and Meaning.*
Presents literature around embodiment and its role in our perceptual link between movement and music.

4: *A Review of DMI Research from a UX Perspective.*
Presents literature around UX methodologies and their use within the DMI field through a systematic literature review.

**Description of Technology**

5: *Glover and the Mi.Mu Gloves.*
Describes the DMI and dedicated end-user mapping application examined in the empirical work.

**Empirical Work**

6: *Mapping Practice of Existing Glove Users.*
Investigation into existing mapping practice, observing a marked difference between novice and expert mapping practice.

7: *Development of Mapping Practice in Novice Glove Musicians.*
Detailed examination into novice mapping practice, tracking the development of mapping practice longitudinally.

8: *Experienced Musical Practice with the Gloves.*
Investigation into expert mapping within professional musical performance practice.

9: *Effect of System Error on Performer Skill Acquisition.*
Examines the effects of system reliability, an influencing factor on end-user mapping design, on performers' ability to acquire skill with the gloves.

**Discussion**

10: *Discussion.*
Summarises the empirical findings and presents recommendations for future research.

Table 1.1.: Dissertation Structure.

# 2. Digital Musical Instruments

The research presented in this thesis centres around end-user mapping design practice with a mid-air Digital Musical Instrument (DMI), the Mi.Mu Gloves. In this chapter, the literature around DMIs is presented, discussing musical expression, mid-air interaction in DMIs, mapping and end-user mapping. Finally, different perspectives in DMI interaction, error in DMIs and the importance of expert practice is discussed. The following chapter (Chapter 3) presents literature around embodiment and the perceptual relationship between gesture and sound.

## 2.1. Defining Digital Musical Instruments

Attempting to define and classify DMIs is prone to difficulties, as the landscape of instruments continues to evolve and develop (Magnusson, 2017). The following definition has been used in the work presented in this dissertation.

A DMI, or New Interface for Musical Expression (NIME), is a musical instrument that uses sensors, computation and audio synthesis rather than acoustic means to create sonic, and additionally haptic and visual, responses to a musician's actions. They consist of a control surface or gestural controller, which drives the musical parameters of a sound synthesis engine in real time, and a sound generation unit (Miranda and Wanderley, 2006; Wanderley, 2001).

An important term to define in the discussion of DMIs is gesture. Gesture is discussed in depth in Chapter 3, but within the discussion of this chapter, gesture refers to the bodily movement that interacts with musical instruments and computing systems (Section 3.1).

Gestural controllers (Wanderley, 2001) have a broad spectrum of possible interaction methods, which have been classified as follows (Mulder, 2000):

- *Touch controllers* that require the performer to touch a control surface;
- *Expanded-range controllers* that require physical contact in a limited form or that do not require physical contact but have a limited range of gestures;
- *Immersive controllers* with few or no restrictions on performer movements.

Multiple methods exist to capture the gestures applied to a gestural controller (Wanderley, 2001):

- *Direct*, where one or more sensors are used to directly monitor performer's actions;
- *Indirect*, where gestures are isolated from the structural properties of the sound produced, usually by analysing an instrumental signal and inferring the gestures used to create it;
- *Physiological*, the analysis of physiological signals, such as EMG. This technique is complicated by the difficulty in isolating meaningful parts of the signal.

The Mi.Mu Gloves can be considered to be *immersive controllers* that use a *direct* gesture capture method.

## 2.2. Musical Expression

Much of the literature emphasises the importance of designing DMIs that allow performers to perform with musical expression (Dobrian and Koppelman, 2006;

Wanderley and Orio, 2002). To understand how instruments provide scope for expression, it is important to consider its contributing physical and musical aspects.

Musical expression can be regarded as the aesthetic choices made by the performer to "deform" symbolic data (such as notes from a score) into the music that is heard (Müller and Mazzola, 2003; Taruskin, 1995). In other words, expression can be described as the intentional manipulation of musical features in a performance, often to provoke a certain emotional response from an audience.

Musical expression is also linked to a performer's personal style, where the deformations made are unique to each performer, facilitated by the specificity of structure in a musical interaction (Gurevich, Stapleton and Bennet, 2009). For example, an interaction structure of "playing a piano" provides more room for individual expression than "playing a C major triad on the right hand".

It has been argued that an instrument must provide control for several sound features to achieve more expressive and varied performances (Jordà, 2004b). However, simply providing controls for multiple sound features does not directly lead to expression, the role of the instrument is to convey the performer's expression by enabling control of these expressive cues in real time (Dobrian and Koppelman, 2006).

An alternative interpretation of musical expression, the dominant model (Gurevich and Treviño, 2007), holds that musical expression is a multi-directional communication between performers, composers and audiences. In this model, the listener determines whether a performance is expressive, and the focus is on the listener's experience of either the emotions performers and composers intended to communicate, or at least their ability to understand them. The idea of expression being about a connection between several stakeholders reflects performance ecologies (Green, 2011), which considers the wider contexts surrounding performance and musical practice and the linkages between audiences, performers and composers to be integral, particularly its effects on each stakeholder's ability to understand the expression being communicated.

An important discussion within musical expression with digital instruments is that of the balance of agency between a performer and their instrument. A communications-oriented perspective focusing on the agency of the performer, while a materials-oriented perspective highlights the agency of the digital instrument (Mudd, 2019), acknowledging that the digital tools used by performers and composers shape and influence the music created with them. From this approach, DMIs can be viewed more as collaborators and co-creators rather than just conduits through which expressive music is created.

## 2.3. Mid-Air Interaction

Mid-air interaction is an interaction method that does not rely on a user physically manipulating an interface, but instead relies on the detection of a user's bodily movements directly; thus a user interacts with the space, or in the air, around them. Mid-air interaction has been used extensively for controlling music, and is a popular interaction method in DMIs (Müller *et al.*, 2014; Silva *et al.*, 2013; Bailly *et al.*, 2012).

The gestural sensing capabilities of an instrument has an influence on the types of interactions a performer can make (Miranda and Wanderley, 2006). Instruments that wish to employ mid-air gesture in their control mappings rely on interfaces that can accurately and comprehensively sense the movements of a user's hands and arms, and as such must use *direct* gestural acquisition methods. The most commonly used interfaces for freehand gestural control broadly lie within one of three categories: vision-based, wearable sensor-based and combined methods.

### 2.3.1. Vision-based Interfaces

Vision-based interfaces typically use cameras or other optical sensors, applying machine vision algorithms to estimate a subject's pose and motion (Aggarwal and Cai, 1999; Gavrila, 1999). The advantage of this is the lack of physical restrictions on users, as well as the relative ease of sensing large scale body movements.

Limitations in vision-based interfaces arise due to the camera's typically static position, which limits the interaction space to within the camera's field of view (Kim *et al.*, 2012). Tracking errors are also introduced when subjects become occluded (Wang and Popović, 2009), and a lack of detail in captured images can prevent vision-based systems from accurately detecting subtle movements and small scale hand gestures. Historically, the low frame rates common in cameras coupled with computationally expensive machine vision algorithms could extend a system's reaction time to beyond that which is acceptable for real time manipulation of audio (Wessel and Wright, 2002). However, it will not be long before the continuing advancements in computer vision algorithms and processing power (Kiani Galoogahi *et al.*, 2017) circumvent this issue.

### 2.3.2. Wearable Interfaces

Wearable solutions typically employ a series of malleable sensors (such as pressure or flex sensors) to detect subtle hand and arm movements, usually embedded within gloves (Jiang *et al.*, 2014; Mitchell and Heap, 2011; Jessop, 2009). Other methods include attaching sensors to the wrist and other body parts, such as an array of sensors around a user's wrist to detect muscle movements (Vogt and Wood, 2014). These methods can provide detailed and accurate representations of the hand, and are well suited to detect precise finger motions with little noise. However, they can often be highly costly, invasive (Rehg and Kanade, 1994), cumbersome

(Pavlovic, Sharma, Huang *et al.*, 1997), and can suffer from maintenance and reliability problems (Gniotek and Krucinska, 2004).

### 2.3.3. Combined Interfaces

Combined approaches combine vision-based and wearable solutions, typically by attaching an optical sensor to the body in some way; solutions such as Digits (Kim *et al.*, 2012) and Leimu (Brown *et al.*, 2016) couple infra-red cameras with an Inertial Measurement Unit (IMU), while the Lightglove (Howard and Howard, 2001) enables virtual typing and pointing using LED scanner/receiver arrays and a two dimensional accelerometer. These approaches can minimise occlusion and field of view problems associated with using static cameras, but can be cumbersome to wear, and while they can have an improved tracking accuracy over static cameras, they remain less precise than glove-like wearable interfaces (Brown *et al.*, 2016).

## 2.4. Mid-air interaction in DMIs

There is a rich history of mid-air interfaces being used for gestural control of music. Notably *The Hands* by Michel Waisvisz, first exhibited in 1984, which incorporated a diverse set of sensors including potentiometers, push keys and ultrasonic transmitters, mapped to MIDI messages (Torre, Andersen and Baldé, 2016; Waisvisz, 1985). The VAMP system (Jessop, 2009) detects hand and arm movements through flex sensors on the elbow and wrist and an accelerometer on the forearm, while more recent work (Klipfel, 2017) uses force sensors for detecting finger tapping as well as finger flex sensors and accelerometers. The Mi.Mu Gloves (Figure 2.1), the DMI used to produce the empirical research presented in this dissertation, detect hand posture and orientation in space using flex sensors and

Figure 2.1.: Chagall Van Den Berg performing with the Mi.Mu Gloves (Image by Ben Houdijk, courtesy of Chagall Van Den Berg).

IMUs, transmitting sensor data over WiFi via Open Sound Control (OSC) (Serafin *et al.*, 2014; Mitchell, Madgwick and Heap, 2012; Mitchell and Heap, 2011).

With all of these solutions, what is notable is the wide diversity of sensor layouts and control mappings chosen by the designers. This diversity is due to the fact that each of these solutions was designed with specific interaction styles and goals in mind, for example, *The Hands* were the personal project of Waisvisz, and were designed in accordance with his own personal aesthetics, while the VAMP system uses choral conducting as its interaction metaphor. Klipfel (2017) decides to afford tapping gestures in his gloves' interaction, while the Mi.Mu Gloves' are designed with open fingertips, allowing musician's to also interact with other instruments or devices.

This highlights the fact that mid-air interfaces must make certain decisions and trade-offs. Additional sensors can capture more movement information, but the technology becomes heavier, requires more power, and constrains natural hand movements and interactions with the world. Meanwhile, vision-based interfaces consume more computational resources, and require more elaborate set-up and calibration procedures to capture more detailed information.

## 2.5. Mapping

In the context of DMIs, mapping refers to the connection between control parameters (derived from performer's actions) and sound synthesis parameters (Hunt, Wanderley and Kirk, 2000). Due to the lack of a physically defined link between these parameters in DMIs, a specified mapping strategy is required to connect a user's actions to a musical response. As such, mapping is an integral part of DMI design, and can be instrumental in establishing the nature of an instrument (Hunt, Wanderley and Paradis, 2003).

Individual mappings between a performer's actions (control parameters) and sound response (audio parameters) are categorised into one of four relationships (Hunt and Kirk, 2000; Rovan *et al.*, 1997):

- *One-to-one* One control parameter affects one auditory parameter;
- *One-to-many* One control parameter affects multiple auditory parameters;
- *Many-to-one* Multiple control parameters affect a single auditory parameter;
- *Many-to-many* Multiple control parameters affect multiple auditory parameters.

Many-to-one mappings have also been described as "convergent", and one-to-many "divergent" (Hunt and Kirk, 2000). In this work, the "many-to-one" and "one-to-many" terms will be used.

A one-to-one mapping has been described as the least expressive, with one-to-many, many-to-one and many-to-many mappings being more expressive as they reflect the complex nature of control in acoustic instruments (Rovan *et al.*, 1997). A violin, for example, uses a many-to-one mapping of bow-speed, bow pressure, choice of string and finger position to control auditory loudness (Hunt and Kirk, 2000). However, one-to-one mappings are often more satisfying to play in DMIs, due to the clear relationship between the control and sound response (Tanaka, 1993).

Traditionally in computer science, a many-to-one mapping is seen as a several inputs combined through a logical OR relationship, where the inputs can each independently be used to control a single output. Alternatively, the combination in a many-to-one can use an AND relationship, where several inputs are combined to produce an output. In DMIs research, mapping literature tends towards conceptualising many-to-ones as using an AND relationship, where several inputs are required to control a single output (Paine, Stevenson and Pearce, 2007; Hunt and Wanderley, 2002; Rovan *et al.*, 1997). For example, Hunt and Kirk (1999) discusses how several inputs, bow speed, pressure and finger position are all necessary to control volume on a violin. These logical AND many-to-one mappings could be referred to as compound mappings (Tanaka, 2010), however, many-to-one will be used to refer to these types of mappings in this dissertation.

### 2.5.1. The Mapping Problem

A common issue when designing DMIs is that the mapping strategy must provide an engaging interaction for performers and audiences alike (Hunt, Wanderley and Paradis, 2003; Rovan *et al.*, 1997). If the mapping is too simple, performers quickly master the instrument and it is rendered little more than a musical toy, while a complex mapping can result in a perceptual disconnect (Nakra, 2002), where a mapping is too difficult to be understood by both audiences and musicians. This is often referred to as *the mapping problem* (Hunt, Wanderley and Paradis, 2003;

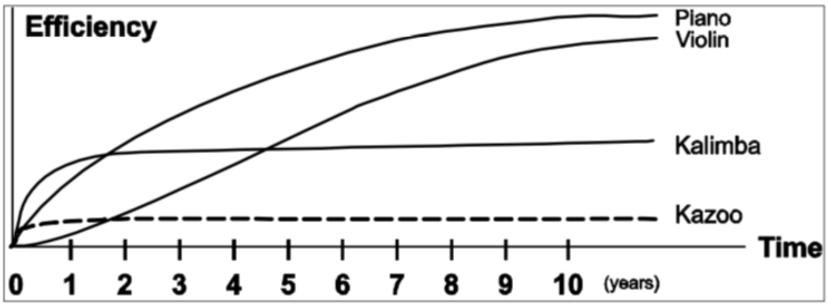


Figure 2.2.: Examples of instruments' learning curves (Jordà, 2004a).

Rovan *et al.*, 1997). Successful mappings are those that allow for novice players to achieve some musicality with little investment, while allowing experienced players to achieve highly expressive and virtuosic performances, or a "low entry fee and no ceiling on virtuosity" (Wessel and Wright, 2002). This is an idea adopted from general Human-Computer Interaction (HCI) research, which advocates that user interfaces for creative tools should have "low floors" easily accessible for new users, "high ceilings" that allow experienced users to perform complex tasks, and "wide walls" that support a diverse range of interactions and styles (Nash and Blackwell, 2011; Resnick *et al.*, 2005; Myers *et al.*, 2000).

Building from this, successful mappings are those that support the psychological state of Flow (Csikszentmihalyi, 1996), a mental state of focus and enjoyment that is reached through a balance between challenge and ability. "Good" DMI mappings are those that find that balance: mappings that are too simple do not to provide rich experiences, and mappings that are too complex alienate the user before their richness can be extracted from them (Levitin, McAdams and Adams, 2002).

A mapping can also provide an engaging interaction for both novices and experts if it promotes an *embodied interaction* (Dourish, 2004), which makes use of our understanding of and relationship with the world around us to influence interaction design. Embodied Interaction is explored in depth in Chapter 3.

Jordà (2004a) explores the idea of supporting both novice and expert interaction in instruments using the "learning curve" (Figure 2.2). An instrument's learning curve is a graphical representation of the measure of the cost of performing an action against the number of times it has been performed. Simple instruments, such as the kazoo, are mastered quickly but do not allow for complex performance, and thus have a steep curve that plateaus quickly. The violin's curve rises slowly and reaches a high peak, reflecting the difficulty of mastering the instrument and its potential for virtuosic performance. Instruments like the piano have a relatively steep learning curve that reaches a high peak, and could be considered an instrument with a "low entry fee and no ceiling on virtuosity" (Wessel and Wright, 2002).

The learning curve and virtuosic potential of an instrument can be greatly affected by its mapping alone. Hunt, Wanderley and Paradis (2003) explore the effects of an "accidental theremin", in which the mapping of two parameters, traditionally mapped to pitch and amplitude in a one-to-one strategy, are instead mapped in a more complex many-to-many strategy. The results of their analysis showed that users of the one-to-one mapping quickly mastered the instrument and grew tired of using it, while users of the many-to-many mapping took longer to master the instrument but found it much more rewarding to play.

In contrast, it can be argued that virtuosity can be found in even the simplest musical instruments, as performers explore all of the possibilities provided by the constraints of the instrument, as well as exploiting "non-obvious" affordances, that may not have been consciously designed (Zappi and McPherson, 2014b; Gurevich, Stapleton and Marquez-Borbon, 2010; Magnusson, 2010).

### 2.5.2. Explicit Mapping

The design of DMI mappings follows one of two approaches: implicit mapping, where the action-sound relationship is inferred using machine learning techniques (Section 2.5.3); or explicit mapping, where the connection is exactly defined (Hunt, Wanderley and Kirk, 2000).

Explicit mapping is when the action–sound relationship in a DMI has been specifically defined by the instrument's designer. This gives designers the benefit of being able to fully perceive and understand the relationship, but it has been argued that designers could create simple one-to-one relationships between raw input parameters and synthesiser parameters due to the simplicity of implementation (Hunt, Wanderley and Paradis, 2003).

Explicit mappings are often implemented in programming frameworks (Malloch, Sinclair and Wanderley, 2013) or visual programming tools like Max (Bellona, 2006; Van Nort and Wanderley, 2006; Bevilacqua, Müller and Schnell, 2005). Dedicated domain-specific Graphical User Interface (GUI) tools are also being developed to facilitate explicit mapping design for end-users (Section 2.7).

### 2.5.3. Implicit Mapping with Machine Learning

Machine Learning (ML) is a body of statistical analysis methods that employs machines to complete tasks by learning from explicit examples, or by finding patterns within a large corpus of data (Caramiaux and Tanaka, 2013). Advancements in computational technologies has led to the application of these techniques in DMI mapping (Fiebrink, Trueman and Cook, 2009; Modler, Myatt and Saup, 2003). Using ML allows designers to move away from explicitly defining the relationships between parameters and instead develop connections implicitly that are "taught by example" (Merrill and Paradiso, 2005) to the instrument.

When using ML in DMI design, care must be taken when deciding which algorithm to use. Classification algorithms, such as Hidden Markov Model (HMM) or Dynamic Time Warping (DTW), provide discrete results from input data, which can be inappropriate for musical instruments, where continuous control over the sound output is desired.

One solution is to find the variance of the incoming gesture, which can be achieved by calculating parameters that measure differences between the target and performed gestures (Caramiaux *et al.*, 2014b). This allows for fine details in performance variation, such as the size and speed of the gesture, to be used for expressive control.

However, there remain some issues with using these methods in designing DMIs. As multiple training examples are used to define a class, the amalgamation of many different performances of the same gesture results in a fairly arbitrary "correct" movement representing an intermediate gesture rather than anything explicitly designed. Additionally, machine learning can be considered a "black box", with users often unaware of the underlying processes. While this allows musicians to abstract their thinking away from individual parameters, it can make precise control in performance scenarios difficult (Fiebrink *et al.*, 2010). Additionally, there is an inherent amount of unreliability with ML techniques, which will nearly always fall short of 100% successful classification rates (Shepperd *et al.*, 2019). This has major implications in music, as system-related errors may not be acceptable for musicians.

### 2.5.4. Features of Mapping

When defining a mapping strategy, it is necessary to discuss the features of the control and sound parameters. In other words, what inputs are being mapped to what outputs?

One possible conceptualisation is to simply have raw controller sensor readings as mapping inputs and low-level synthesis parameters as mapping outputs. For example, values from sensors (e.g. button states and slider values) are given direct relationships to audio parameters (in FM synthesis for example: modulation index, carrier frequency and modulator frequency). To facilitate higher-level conceptualisations of auditory control, it is important to present the raw controller values and direct synthesiser parameters in abstracted, semantic layers describing gesture and sound and their relation to perception.

This work will refer to the gesture and sound elements of this perceptual mapping layer as gestural *features* and sound *parameters*. The same conceptualisation has been referred to through several terms, such as: gestural and sound semantics (Malloch, Sinclair and Wanderley, 2007); related-to-gesture perception and related-to-sound perception parameters (Arfib *et al.*, 2002); and gestural and sound parameters (Kvifte and Jensenius, 2006).

To facilitate for higher-level abstraction, Arfib *et al.* (2002) present a "chain" (Figure 2.3) of the mapping process. In the mapping chain, raw sensor data is processed into gestural features (which Arfib *et al.* refer to as gesture-related perception parameters), mapped to sound parameters (sound-related perception parameters) that relate to synthesis parameters. It is between these semantic layers where the creative element of mapping the process takes place. Raw data from a gestural controller may be passed through a gesture recogniser algorithm to become a gestural feature like "downward stroke". On the other end, we may have the sound parameter "brightness", which then maps to a synthesis process that controls high frequency content (e.g. a low pass filter).

The gestural features that are available to an instrument is highly dependent on the affordances of its interface. For example, using a single camera positioned in front of the performer with their entire body in frame will allow for large arm gestures to be used for control, while subtler finger movements may be unintelligible. Similarly, if the camera is positioned much closer to the performer, so that

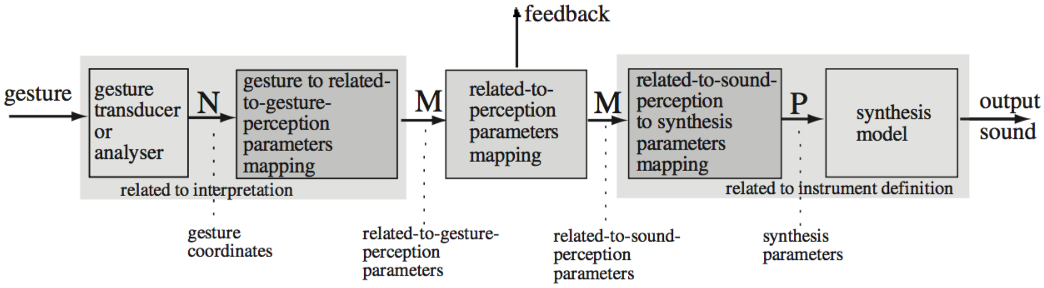


Figure 2.3.: The mapping chain (Arfib *et al.*, 2002).

only their hand is within frame, small finger movements will be easily detected, but the performer will be restricted from using larger arm motions.

Paine, Stevenson and Pearce (2007) argue that there are a series of underlying features of physical instrument control: Pressure, Speed, Angle and Position. Features within these categories are mapped to sound parameters. On a violin for example, changes in *bow* pressure, bowing speed, the angle of the bow against the string (more or less bow hairs in contact with the string) and the position of the bow all contribute, in differing significance, to the control of dynamic level; while on a flute, changes in *breath* speed, *finger* positions and *mouth* angle contribute to dynamics.

Sound parameters can be drawn from definitions of expressive features of musical performance, of which there are five categories (Paine, Stevenson and Pearce, 2007; Poepel, 2005; Juslin, 2003):

- *Timing*: tempo, timing, pauses;
- *Pitch*: intonation, vibrato;
- *Dynamics*: sound level;
- *Articulation*: tone attacks, tone decays;
- *Timbre*: tonality.

These can be either be discrete, such as a note being triggered; or continuous, such as an amplitude envelope (Kvifte and Jensenius, 2006). Tanaka (2010) presents a similar model that distinguishes three types of musical features: *Binary*, sound being activated; *Basic*, fixed parameters changed such as pitch selection or audio effect; and *Expressive*, continuously varying parameters that follow a gesture, such as vibrato or tremolo.

## 2.6. End-User Mapping

In HCI, eliciting the end-user in the design of mappings is often used as a method to try and find intuitive and ergonomic solutions for an interaction, the logic being that a general "gesture vocabulary" will exist between users, and can be utilised to provide natural interactions for all (Nielsen *et al.*, 2004). Previous HCI work in end-user designed gestures has found that while some gestures occurred frequently between users, a great deal were novel and diverse (Kray *et al.*, 2010), and that the context of the interaction plays an important role on what users believe a gesture should control (Wobbrock, Morris and Wilson, 2009). For example, a touch-screen gesture of splaying five fingers outward on an object on screen should enlarge it, but doing so on the screen's background should zoom in.

End-user designed interaction has also been explored in computer music for similar purposes, exploring the relationship between mid-air gestures and musical material. Previous work has looked at mapping by demonstration (Françoise, 2015; Caramiaux *et al.*, 2014a), in which participants are asked to give gestural responses to musical stimuli. Similar work (Godøy, Haga and Jensenius, 2006) asked participants to "trace" a gesture in response to sounds they heard. This work reveals insights into the embodied link between gesture and music (discussed in Chapter 3), however, it focuses on participants' gestural responses to sound instead of active gestural control of sound.

While user-defined gestures provide a mapping that is meaningful for the performer who created it, mappings created using this technique often cannot be generalisable between performers or audiences. This technique also requires a lengthy set-up process on the part of the performer, and mappings may also need to be set in a prescribed order, leading to premature commitment (Nash, 2015) as they may be difficult to alter later.

## 2.7. End-User Mapping Software

The Glover software (described in Chapter 5) provides an environment in which users define their own connections between gestural features and sound parameters. Glover is a tool for creativity support (Shneiderman, 2007), and can be considered a Domain-Oriented Design Environment (DODE). DODEs support the creation and development of applications and systems within a specific context domain, and empower individual creativity by hiding the computational complexity of the computer and focusing the user's attention to the task at hand (Fischer, 1999), and aim to reduce the large conceptual distance between the semantics and conceptual perceptions of the problem-domain (Fischer, 1994). Fischer and Girgensohn (1990) provide a taxonomy for what features DODEs require to facilitate end-user modifiability: setting parameters, adding functionality to existing objects, creating new objects by modifying existing objects, and defining new objects from scratch.

Fischer and Girgensohn (1990) argue that end-user modifiability is not a luxury, but a necessity, particularly in cases of personal aesthetics and creativity, which is fostered when the end-user is given agency over the system's behaviour. This is evident in most music production software, where the user is given the tools and responsibility to create any form of digital music they desire from a starting point of silence. While allowing for as much creative freedom as possible, this "blank canvas" approach can be problematic, with many researchers observing that it can be the most challenging and daunting part of the creative process (Collins, 2007; Alty, 1995).

Several pieces of end-user mapping software exist in the literature. Prominent examples include WebMapper[1], GECO[2], junXion[3] and MyoMapper[4].

### 2.7.1. WebMapper

WebMapper (Figure 2.4) is a browser based application designed for devices built using the libmapper mapping framework (Malloch, Sinclair and Wanderley, 2013). The tool automatically detects software devices built using the framework and presents the available inputs and outputs, both in the OSC message format, for the user to map together.

Mapping in WebMapper follows a patch cord design that reflects the traditional visualisation of instrument mapping. The inputs and outputs themselves are uneditable by the user, as they are defined by the libmapper devices that exist independently from WebMapper. A user instead edits the patch cord connections, which, once selected, can be modified through the toolbar at the top of the application. WebMapper supports users defining the mapping function between the input and output by editing a $y = mx + c$ style mathematical expression, as well as the scaling and offset of the input and output value ranges.

WebMapper supports one-to-one and one-to-many mappings. Many-to-ones are not possible, as the software has no feature for combining inputs. Instead, the output simply alternates between the two input signals.

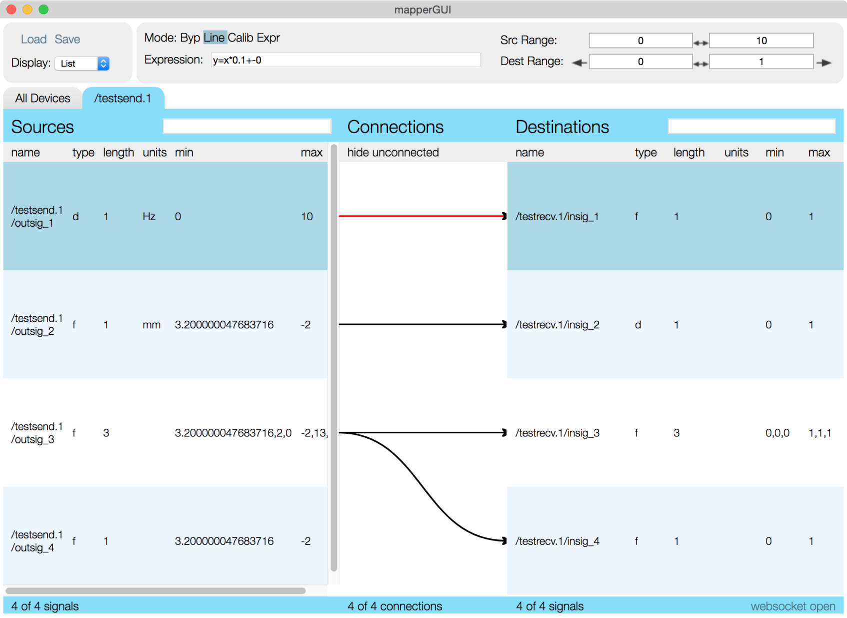
---

[1] libmapper.github.io/ecosystem/user_interfaces.html

[2] gallery.leapmotion.com/geco-midi

[3] www.steim.org/product/junxion

[4] www.balandinodidonato.com/myomapper/



Display: List

Expression: y=x*0.1+-0

Src Range: 0 10

Dest Range: 0 1

| name | type | length | units | min | max |
|---|---|---|---|---|---|
| /testsend.1 /outsig_1 | d | 1 | Hz | 0 | 10 |
| /testsend.1 /outsig_2 | f | 1 | mm | 3.200000047683716 | -2 |
| /testsend.1 /outsig_3 | f | 3 | | 3.200000047683716,2,0 | -2,13, |
| /testsend.1 /outsig_4 | f | 1 | | 3.200000047683716 | -2 |

| name | type | length | units | min | max |
|---|---|---|---|---|---|
| /testrecv.1/insig_1 | f | 1 | | 0 | 1 |
| /testrecv.1/insig_2 | d | 1 | | 0 | 1 |
| /testrecv.1/insig_3 | f | 3 | | 0,0,0 | 1,1,1 |
| /testrecv.1/insig_4 | f | 1 | | 0 | 1 |

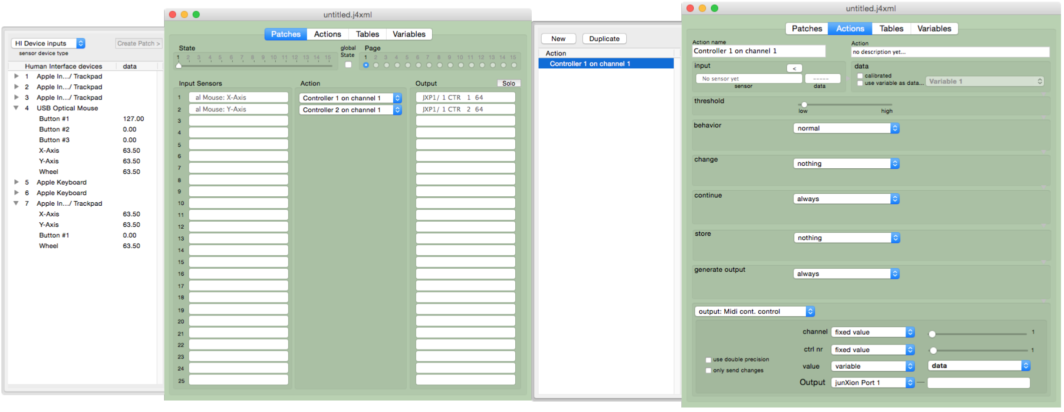
Figure 2.4.: WebMapper

### 2.7.2. GECO

GECO is a gesture-to-MIDI mapping tool designed for the Leap Motion gestural controller. GECO takes data from the Leap Motion and presents the user with gesture features: up-down, left-right and back-forth axes, tilt, open hand (fingers abducted) and closed hand (fingers adducted); which can be mapped to MIDI messages. GECO supports one-to-one mappings and many-to-one mappings, where specific gesture features are combined. Scaling the input signal from the Leap Motion to the 0–127 of MIDI is done by the program. The only signal attenuation afforded to users is the ability to offset the output value.

### 2.7.3. junXion

junXion (Figure 2.5) is a mapping tool built by STEIM that natively recognises joysticks and gestural devices such as Nintendo Wii remotes, and allows users to map sensor data to MIDI and OSC.



| Human Interface devices | data |
|---|---|
| ▶ 1 Apple In.../ Trackpad | |
| ▶ 2 Apple In.../ Trackpad | |
| ▶ 3 Apple In.../ Trackpad | |
| ▼ 4 USB Optical Mouse | |
| Button #1 | 127.00 |
| Button #2 | 0.00 |
| Button #3 | 0.00 |
| X-Axis | 63.50 |
| Y-Axis | 63.50 |
| Wheel | 63.50 |
| ▶ 5 Apple Keyboard | |
| ▶ 6 Apple Keyboard | |
| ▼ 7 Apple In.../ Trackpad | |
| X-Axis | 63.50 |
| Y-Axis | 63.50 |
| Button #1 | 0.00 |
| Wheel | 63.50 |

| | Input Sensors | Action | Output |
|---|---|---|---|
| 1 | al Mouse: X-Axis | Controller 1 on channel 1 | JXP1/ 1 CTR 1 64 |
| 2 | al Mouse: Y-Axis | Controller 2 on channel 1 | JXP1/ 1 CTR 2 64 |

(a) Patches tab.

Action: Controller 1 on channel 1

Action name: Controller 1 on channel 1
Action: no description yet...
input: No sensor yet (sensor) ----- (data)
data: ☐ calibrated ☐ use variable as data... Variable 1
behavior: normal
change: nothing
continue: always
store: nothing
generate output: always
output: Midi cont. control
channel: fixed value 1
ctrl nr: fixed value 1
value: variable data
Output: junXion Port 1

(b) Actions tab.

Figure 2.5.: junXion

junXion uses tabs for different elements of the mapping process. The Patches tab (Figure 2.5a) provides an overview of all mapping connections; the Actions tab (Figure 2.5b) allows the user to modify a single mapping path; the Tables tab allows the user to create custom transform function tables; and the Variables tab allows the user to create and monitor variables, which are used to store values to be used across junXion.

Mappings are organised by input parameter. Once an input is added to the Patches list, it can then be edited in the Actions tab. The user can modify the input data through a series of combo boxes, which junXion uses for nearly all user-modifiable features, before selecting a corresponding output parameter. The software also allows for input values to be stored in variables, which is the only way of using the same input value across multiple mappings.

### 2.7.4. MyoMapper

MyoMapper allows end-users to map gestural data from a Myo armband to OSC data, and has been applied in Virtual Reality, robotics, and dance performances. The software is organised into four sections: calibration and scaling, feature extraction, feature scaling, and message streaming (Di Donato, Bullock and Tanaka, 2018).

The software's calibration and scaling functionality allows users to set minima and maxima from the raw rotation value (pitch, yaw and roll) provided by the Myo armband ($-2\pi$–$2\pi$), scaling this to $0$–$1$. There is also a "flip" function, that applies a $y = 1 - x$ function to the scaled sensor data. The software's feature scaling functionality derives a series of features from the Myo sensor data such as the moving average, minimum, maximum, and absolute values, zero crossing rate and first order and second order derivations. These are intended to be used in third-party ML tools such as the Wekinator (Fiebrink, Trueman and Cook, 2009), but can also be used as gesture features in explicit mapping. The feature selection window allows users to route input features to one of two OSC ports, one for explicit mapping and one for ML mapping, while the OSC streaming window allows users to send the data from these inputs as user-defined OSC outputs.

## 2.8. Different Perspectives in DMIs

It is important to consider that a performer's perception of an instrument in terms of playability and expressivity will be different than the perceptions of audiences, who may only be concerned with resulting auditory and visual experience (O'Modhrain, 2011).

Transparency (Fels, Gadd and Mulder, 2002) is a concept that considers the perspectives of audiences and performers in DMI design, and is interested in the ability of audiences and performers to perceive the connections between a performers movements and the instrument's sound response. The Mapping Transparency Scale (Figure 2.6) consists of two axes, one for performers and one for audiences, each ranging from opaque to transparent. The framework does not

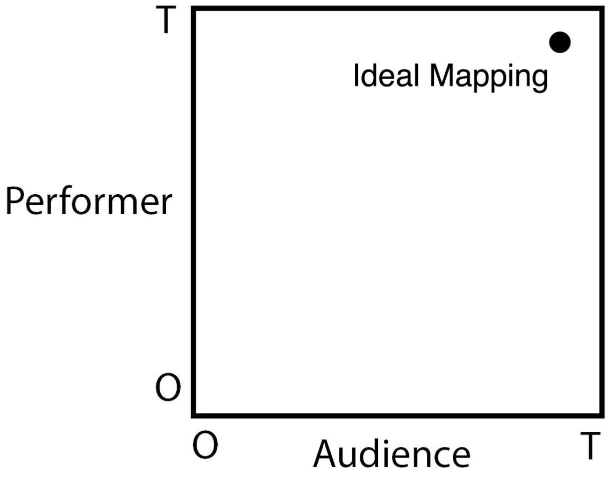


Figure 2.6.: The Mapping Transparency Scale (Fels, Gadd and Mulder, 2002).

indicate how transparency can be measured, but does provide a helpful design concept that views perceptions of spectators as equally important as the perceptions of performers. Similar concepts have been presented in the broader HCI literature (Reeves *et al.*, 2005), where a spectator's ability to perceive a user's *manipulations* and the resulting *effects* are measured on a scale from hidden to amplified.

O'Modhrain (2011) describes a series of stakeholders in a DMI: Audience, Performer/Composer, Designer and Manufacturer, and puts forward a series of evaluation goals related to them: Enjoyment, Playability, Robustness and Achievement of Design Specifications. Under each evaluation goal, each stakeholder has a different means of evaluating the instrument's effectiveness. For example, under *Enjoyment*, a performer might evaluate an instrument through reflective practice and long-term engagement, while a manufacturer may use market surveys and sales figures.

## 2.9. Error in DMIs

A spectator's understanding, and potentially enjoyment, of a musical performance relies on their ability to perceive the skill of a performer (Gurevich and Fyans, 2011). A metric for an audience's perceptual understanding of a DMI is their ability to perceive performer error (Fyans, Gurevich and Stapleton, 2009a; Fyans, Gurevich and Stapleton, 2009b), which has been used in empirical studies of DMIs (Bin, 2018). These ideas of a spectator's understanding of a performance can be linked to the idea of Liveness (Auslander, 2008), which highlights an audience's desire to witness "authentic" music making. A sense of Liveness is achieved when a performance "feels real", with the visual aspect of a live performance being intrinsically linked to, and reinforcing, the authenticity of the auditory aspect in the minds of the audience.

As well as performer-related error, system-related error is a factor that must be considered in DMIs, particularly as ML techniques become ever more popular in DMI and music interaction design (Macionis and Kapur, 2018; Vogl and Knees, 2017; Scurto, Bevilacqua and Françoise, 2017). Mid-air and gestural interaction systems are particularly susceptible to error (Arif *et al.*, 2014; Norman and Nielsen, 2010), due to a lack of feedback (Gustafson, Bierwirth and Baudisch, 2010) and the difficulties recognition systems have distinguishing between similar gestures (Morris, Wobbrock and Wilson, 2010).

DMI research into system-related issues has focused on understanding and mitigating the latency inherent in digital signal processing (Jack, Stockman and McPherson, 2016; McPherson, Jack and Moro, 2016; Pardue *et al.*, 2014), with Wessel and Wright (2002)'s infamous 10ms latency limit for effective control intimacy remaining influential. Recent work has found that unpredictability in latency has more of an impact than the length of latency itself, with a 10ms delay with 3ms of jitter being perceived to be as bad as a 20ms delay (Jack, Stockman and McPherson, 2016).

## 2.10. Expert Performers for DMIs

Often in HCI and DMI research, the focus of an evaluation is on the experience of novice users (Siegel, 2012). While examining a novice's experience of technology can be useful, especially for initial impressions, the potential scope of the evaluation becomes limited. Nielsen (1994) argues that a small number of expert users can find usability issues that would elude a large number of novices, as their intimate knowledge of the problem domain and current software solutions gives them a knowledge and appreciation of niche program features that novices would not reach in the short amount of time allotted to most usability studies.

In a similar manner to heuristic evaluation, the skill and expertise of a small number of experienced musicians can be elicited to provide meaningful insights into new DMIs (Johnston, 2009). However, very few DMI studies make use of experienced users in their evaluations. This is mostly due to a lack of experienced performers for new DMIs beyond the initial designer (McPherson and Kim, 2012), owing to the newness of the instrument and a lack of existing practitioner communities (McPherson and Kim, 2012), the instrument being developed for a specific piece (Paine, 2009), or that the instrument is a fragile, expensive prototype with no mass-market presence.

This lack of experience can be mitigated by making use of the existing expertise in traditional instruments that is abundant in society, and is useful in cases of augmented instruments, which offer extended performance techniques on top of existing playing methods (Eldridge and Kiefer, 2017; McPherson, Gierakowski and Stark, 2013). However, while their interactions are based on traditional instruments, augmented instruments become new instruments in their own right, and while knowledge of the relevant traditional instrument can give musicians a head start towards expertise, there can be a whole new set of affordances and constraints that musicians must learn. This practice is also limited in the case of mid-air instruments that have no grounding in traditional instrument interaction styles. What is needed for a thorough investigation into the nature of a musical instrument's capacity to facilitate expression is for it to have expert practitioners of their own.

The development of expert-level users for a given DMI is often dependent on the instrument having a community of dedicated practitioners, who train new musicians through formalised pedagogy, build a body of standard works and pieces, (Marquez-Borbon and Stapleton, 2015) and discover new instrumental techniques (McPherson and Kim, 2012). The Mi.Mu Gloves, the DMI examined in this research, is one such instrument that has a community of practitioners, a number of which are experts who use the instrument in their professional performance practice.

## 2.11. Summary

This chapter has presented literature around DMIs. Section 2.1 set out DMIs classifications, while Section 2.2 discussed the importance of facilitating musical expression within DMIs. Section 2.3 discussed mid-air interaction and how it is achieved through various interface solutions, and Section 2.4 then discussed how mid-air interaction is used within the design of DMIs. Section 2.5 discusses the importance of mapping in the DMI design process, how mappings are implemented and the features that make up mapping strategies. Section 2.6 and 2.7 discussed end-user mapping in music interaction and existing software solutions. Section 2.8 discussed the different perspectives that exist within DMI interactions and the mapping transparency concept. Section 2.10 discussed the importance of soliciting expert performer perspectives in DMI design and evaluation, and the current lack of established practitioners.

# 3. Movement, Music and Meaning

The previous chapter (Chapter 2) presented literature around the Digital Musical Instrument (DMI) field, with a focus on mapping design. The next chapter (Chapter 4) presents literature around User Experience (UX) concepts and their application in studying and evaluating DMIs.

This chapter discusses embodiment and the perceptual link between movements and sound, which informs the theoretical position from which the research presented in later empirical chapters is conducted (Chapters 6, 7, 8 and 9). The chapter begins by discussing gesture and the distinction between bodily gesture and gesture in music, before discussing embodiment and its implications in cognition, Human-Computer Interaction (HCI) and our understanding of music. Conceptual Metaphor Theory is then discussed, which is used in the later empirical chapters as a way of revealing the embodied meaning behind the mapping decisions of glove musicians. This chapter also touches on alternative theoretical positions of understanding gesture and music: ecological approaches and semiotics.

## 3.1. Music and Gesture

Our experience with music is intrinsically linked to bodily gesture (Cox, 2016; Leman and Godøy, 2010; Jensenius, 2007). Arguing that musical experience is inseparable from movement, Leman and Godøy (2010) conceive that gestures are way of expressing of a “fundamental connection” between music and movement. Discussing how listeners mimic musical performance through performing air guitar, miming conducting or copying the body language of singers, they propose the following questions:

> Why is it that so many listeners are able to spontaneously make gestures that seem to fit the music? Why do they make these gestures? How are these gestures related to the music, and how are these gestures related to the gestures of performers? (Leman and Godøy, 2010)

The idea of gesture is central to DMI research. There are many definitions of gesture with relative applications in the DMI field, with 62% of The International Conference on New Interfaces for Musical Expression (NIME) papers using the term (Jensenius, 2014). From a body-related perspective, the term is used to refer to a bodily movement that interacts with musical instruments and computing systems and in relation to DMI mapping, bodily gesture features are mapped to sound parameters. From a sound-related perspective, musical gesture describes the motion in sound (Leman and Godøy, 2010). In music interaction design, both approaches are important to consider, as our cognitive understanding of sound motion is underpinned by our bodily motions and movements (Cox, 2016; Leman and Godøy, 2010), and it is this link which can be harnessed to recouple a musician’s actions and the sound response in a DMI. In these cases, the physical gesture must elicit one’s understanding of musical motion in order for the action to seem genuine.

### 3.1.1. Body-Related Gesture

Body-related gestures comprise postures and movements that operate within biomechanical constraints of the human body (Leman and Godøy, 2010). Jensenius *et al.* (2009) present a three-point classifications of bodily gesture:

- *Communication*, when gestures transmit meaning in social interactions;
- *Control*, where gestures are used to manipulate a system, often computational and interactive;
- *Metaphor*, where gestures project concepts of physical movement, sound or other perceptions of cultural topics.

Cadoz and Wanderley (2000) present a similar three-point classification, where communication gestures are referred to as *semiotic* gestures, control as *ergotic*, and metaphor as *epistemic*. Jensenius *et al.* (2009) also propose a classification of bodily gestures relevant to musical performance. In these classifications, a single gesture may fit within multiple categories at once (Dahl *et al.*, 2010).

- *Sound-producing* gestures, movements used to create sound (Control);
- *Communicative* gestures, intended to communicate meaning, either between performers or performers and their audiences (Communication);
- *Sound-facilitating* gestures, support the sound-producing gestures (Control);
- *Sound-accompanying* gestures, not involved in sound production, but follow the music (Metaphor).

Sound-producing gestures, being related to direct control of auditory processes in DMIs, has received more attention from the DMI field, and the category has been broken down further (Godøy, Haga and Jensenius, 2006; Cadoz, 1988).

- *Excitation* gestures, which provide the energy that will be present in the perceived phenomena. Either instantaneous (e.g. percussive or plucking) or continuous (e.g. bowing);

- *Modification* gestures, which affect the relation between the excitation and the resulting sound. Either parametric, a continuous or discrete variation (e.g. vibrato), or structural, a categorical modification (e.g. a mute inserted in a trumpet).
- *Selection* gestures, which consist of choosing between multiple similar elements in an instrument (e.g. fretting a chord on a guitar).

### 3.1.2. Sound-Related Gesture

Sound gestures are created when we experience expressive qualities in sound material that form a coherent gestalt of pitch, temporal and timbral information.

Many arguments have been presented to explain the “gestural” nature of sound. Schneider (2010) discusses the concept of sympathetic movement, which suggests that musical “forces” affect a listener, and that specific motor behaviours are induced by music. This approach alludes to a certain passivity in the listener, and that they have little control over the way in which they react to musical stimuli, and are simply resonating sympathetically with the sound stimulus. Similar to this is motor theory, which holds that the gestures we perceive in music is our way of processing sensory information through mentally simulating the associated movement (Godøy *et al.*, 2016).

An embodied approach, which states that all abstract ideas and concepts are grounded in our bodily experience, ascribes more agency to the listener, suggesting that the perception of gestures in musical material is caused through our mimetic comprehension of the music, imagining what it is like to *be* or *do* the sound (Cox, 2016). This is discussed further in Section 3.2.2.

## 3.2. Embodiment

Embodiment, or embodied cognition, is an extension of the philosophy of phenomenology (Jensenius, 2007), and stands at odds with Cartesian duality, arguing that the mind and body are not two separate things, and that our cognitive processes are dependent on our understanding and relation with the world around us (Leman, 2008; Dourish, 2004; Winograd and Flores, 1986). Embodiment argues that the ways in which we think about and give meaning to the world are all based on our bodily interactions with it. In embodied cognition, meaning is more than just words and concepts, and exists in relation to ourselves and our environments.

> An embodied view of meaning looks for the origins and structures of meaning in the organic activities of embodied creatures in interaction with their changing environments. It sees all our higher functioning as growing out of and shaped by our abilities to perceive things, manipulate things, move our bodies in space, and evaluate our situation (Johnson, 2008).

Embodied cognition draws on the ideas of Heidegger, particularly his distinction between *vorhanden* (or *present-at-hand*) and *zuhanden* (or *ready-to-hand*). The classical example of the distinction is the use of hammer. The person doing the hammering is not actively thinking about the hammer, and it becomes an extension of their body (*ready-to-hand*). If there is a break down in the interaction, if the hammer breaks or slips, the hammerer becomes aware of the hammer (*present-at-hand*). This phenomenon has been noted in DMI research: Hunt, Wanderley and Paradis (2003) discuss a moment when users stopped thinking about their interactions and “just played” an instrument, “as if their conscious mind were not in control”.

Leman and Godøy (2010) build on ideas of embodiment using *body schema* and *body image*. They describe body schemas as gestures and movements that can be carried out with no conscious effort, such as picking up a glass of water. These movements take place without the need to be aware of the role of each individual muscle. Leman and Godøy discuss how this manifests in the performances of trained musicians, who do not have to think about the movement of their fingers on their instrument; they simply play the intended melody, which appears to "come out of the body." This idea of body schema clearly reflects Heidegger's original concept of *Zuhanden* (ready-to-hand), where the tool becomes an extension of the body; while body image, which refers to one being aware of the body in relation to the environment, such as consciously learning a new melody and the movements required to perform it, is a reworking of *Vorhanden* (present-at-hand), where the tool is the focus of a user's attention.

Regarding movement, Johnson (2008) argues that there are four recurring qualitative dimensions of all bodily movement: *Tension*, bodily exertion felt through muscular tension, *Linearity*, the spatial, directional qualities of the movement, *Amplitude*, the amount of expansion or contraction in the range of motion, and *Projection*, the vectoral quality of forceful bodily actions. These four dimensions also reflect the four dimensions of movement described by Paine, Stevenson and Pearce (2007) of Pressure, Speed, Angle and Position.

### 3.2.1. Embodied Interaction

In HCI, embodied interaction is an approach to the design and analysis of interaction that takes embodiment to be central to the whole phenomenon, and attempts to move computation and interaction out of the world of abstract cognitive processes and into the same phenomenal world as our other sorts of interactions (Dourish, 2004; Winograd and Flores, 1986).

Dourish (2004) describes this concept in HCI, and how in a typical computer interaction one performs tasks on a computer *through* the mouse, as an extension of ones hand (zuhanden). When a user reaches the limitations of the mouse's capabilities (for example, reaching the edge of the screen), the user becomes concious of the mouse and it becomes the object of their attention (vorhanden).

Interactions designed on embodied principles can lead to intuitive interactions, where a user is able to immediately understand an interface, and the interface behaves as expected (Antle, Corness and Droumeva, 2009; Spool, 2005). Intuitive interaction occurs when, in a movement-based system, users enact appropriate input actions unconsciously or automatically, rather than consciously learning, step-by-step, how to interact with the system (Antle, Corness and Droumeva, 2009). However, a user finding these interactions to be intuitive depends on the user's previous experiences in a given domain.

Thus an interaction device is successful in supporting an embodied interaction if the user begins to act *through* the device, instead of acting *on* the device. Embodied interactions also occur in musical performance; once a musician becomes an expert they act through the instrument in order to create an intended, specific sound, instead of being consciously aware of their actions upon the instrument (Hunt, Wanderley and Paradis, 2003).

Much research in the DMI field considers embodiment and embodied interaction to be important for understanding and providing instruments that strike the right balance between boredom and challenge (Nash and Blackwell, 2011; Wilkie, Holland and Mulholland, 2010; Jensenius, 2007).

### 3.2.2. Embodiment and Music

Music is meaningful as it presents the flow of human experience, feeling and thinking in an embodied form, with musical structure, temporal flow, pitch contours and intensity being linked to patterns within bodily experience.

Cox (2016) argues that we comprehend music mimetically, and that we consider two fundamental questions in our musical experiences: "what is it like to *do* that?" and "what is it like to *be* that?" The *doing* question relates to our imagining of the overt movements involved in creating a sound, for example, the physical exertions of the performer, which Cox calls Mimetic Motor Action (MMA). The *being* question relates to our imagining of the movement of the music that does not manifest in overt actions, or Mimetic Motor Imagery (MMI). The Mimetic Hypothesis argues that our experience of sound is that it is produced by physical events, and so sounds signify the physicality of their source. He argues that while some music "resists" mimetic engagement, all acoustical features; such as pitch, duration, timbre, strength, location; can be mimetically represented.

## 3.3. Metaphor

The concept of metaphor is important in HCI and subsequently DMI research. The term metaphor has a large scope of potential meanings in HCI and requires careful contextualisation (Blackwell, 2006). Within this research, the two forms of metaphor most applicable are the ideas of embodied and conceptual metaphor.

### 3.3.1. Embodied Metaphor

To understand and make sense of abstract concepts, they must be rooted in our embodied experience, and expressed through metaphors (Lakoff and Johnson, 1980). An embodied metaphor is when concepts relating to one domain, grounded in our bodily experience of the world (source domain), are used to express ideas relating to an abstract concept (target domain). An example, from Lakoff and Johnson (1980), is the metaphor ARGUMENT IS WAR, in which the abstract concept of arguing (target) is explained in the conceptual system of war (source): "he *attacked the weak points* in my argument," "your claims are *indefensible*."

Antle, Corness and Droumeva (2009) draw from embodied metaphors (Lakoff and Johnson, 1980) to support the design of intuitive interactions in an interactive audio environment. A series of metaphors for music and movement are ascertained by interviewing dancers and choreographers. It is important to note that these metaphors are dependent on the participants individual experiences, which, due to all of the participants experience being in dance, may differ from metaphors drawn from a more varied pool of participants:

- Tempo: fast movement is fast tempo, slow is slow.
- Volume: more movement is loud, less is quiet.
- Pitch: nearer to the sensor is high, farther is low.
- Rhythm: smooth movement is rhythmic, choppy is chaotic.

Two versions of the interactive environment were devised: one that followed these embodied metaphors, and another that subverted them. It was found that the system that followed the embodied metaphors was "easier to use" than the non-embodied system.

### 3.3.2. Conceptual Metaphor

A conceptual metaphor also expresses an idea in one domain through the terms of another, but while a conceptual metaphor can be grounded in embodiment, it is often the case that the source and target domains are linguistically or semantically linked, and as such conceptual metaphors are highly influenced by the language used to express them (Casasanto, 2009). For example, English uses metaphors that map time to spatial distances while Greek uses metaphors that map time to an amount of substance (Casasanto, 2009).

Conceptual metaphor is generally used in HCI to provide a cognitive basis for facilitating intuitive interactions (Antle, Corness and Droumeva, 2009), by explaining the behaviour of computer software via an association with a more familiar domain (Marx, 1994).

### 3.3.3. Metaphor in Music

Embodied cognition, and by extension metaphor, underpins our experience of music (Cox, 2016; Johnson, 2008; Brower, 2000). It is argued that our understanding of music is entirely metaphoric, grounded in three of our basic bodily experiences of physical motion: MOVING MUSIC, MUSICAL LANDSCAPE and MUSIC AS MOVING FORCE (Johnson and Larson, 2003). The MOVING MUSIC perspective is that the music moves relative to a stationary observer, "*here comes* the recapitulation". The MUSICAL LANDSCAPE perspective is that the observer moves through a stationary musical environment "we are *approaching* the Coda". The MUSIC AS MOVING FORCE perspective is that the music is the force that moves us through the world, "The chorus *blew me away*".

These perspectives reflect, and are built upon, the similar perspectives of the conceptual metaphoric understanding of our perception of time (Lakoff and Johnson, 1980). Moving observer: "we're *coming up on* Christmas", moving time: "Christmas *is coming*". The phenomenon of language reflecting our embodied and metaphoric understanding of the world can be exploited to investigate how musicians, and non-musicians, conceptualise music (Wilkie, Holland and Mulholland, 2010; Brower, 2000).

| Source (Physical Motion) | | Target (Music) |
|---|---|---|
| Physical object | –> | Musical event |
| Physical motion | –> | Musical motion |
| Speed of motion | –> | Tempo |
| Location of observer | –> | Present musical event |
| Objects behind observer | –> | Past musical events |
| Path of motion | –> | Musical passage |
| Starting/ending point of motion | –> | Beginning/end of passage |
| Temporary cessation of motion | –> | Rest |
| Motion over same path again | –> | Recapitulation, repeat |

Table 3.1.: The MOVING MUSIC metaphor (Johnson and Larson, 2003).

Cox (2016) builds on the discussion of the metaphoric logic of musical and temporal motion, and that as anticipation is correlated with LOCATION-AHEAD and APPROACH, presence is correlated with LOCATION-HERE and ARRIVAL, and memory is correlated with LOCATION-BEHIND and DEPARTURE; when we experience these phenomena in non-spatial domains like music, these conceptual correlations motivate a sense of musical locations.

### 3.3.4. Metaphor in Mapping Design

Wessel and Wright (2002) argues for the application of conceptual metaphors in DMI mapping design, using metaphors relating to spatial representations of pitch, as well as more specific metaphors intended for specific musical interactions, such as "scrubbing" for temporal control, "dipping" for volume control and "catch and throw" for selecting and deselecting musical material are explored.

Fels, Gadd and Mulder (2002) also explore metaphor in mapping design, examining how its use leads to "transparent" mapping strategies (see Section 2.8), examining instruments that use metaphors in their mapping strategies they find that metaphors aided the understanding of certain aspects of an instrument, but that their explanatory power broke down if the instruments behaviour did not match the framework of the conceptual metaphor. This reflects an effect described by

Lakoff and Johnson (1980), where in using the structure of one concept to describe the behaviours of another, certain features of the explained concept are highlighted by the metaphor while others are minimised or lost.

Similar to Antle, Corness and Droumeva (2009), Wilkie, Holland and Mulholland (2010) draw from dialogue between musicians to establish conceptual metaphors that exist within music, finding examples such as HARMONIC PROGRESSION IS MOVEMENT ALONG A PATH, A KEY/CHORD IS A CONTAINER FOR NOTES and MUSICAL SILENCE IS A BLOCKAGE TO MOVEMENT, which are then used to analyse musical software by establishing whether a certain metaphor is apparent in the design and use of the software. They conclude that using conceptual metaphor in interaction design can create intuitive interaction models for expert users while providing affordances for novice users.

### 3.3.5. Image Schemas

An image schema is a "dynamic, recurring pattern of organism-environment interactions" (Johnson, 2008). Image schemas are ways in which we conceptualise our understanding of the world via spatial relationships, and they form the backbone of conceptual metaphor, playing an important role in language in general. The term was coined by George Lakoff and Mark Johnson, and are important primarily because "they help to explain how our intrinsically embodied mind can at the same time be capable of abstract thought" (Johnson, 2005). Hampe (2005) provides a condensed categorisation:

- Image schemas are directly meaningful ("experiential" / "embodied"), preconceptual structures, which arise from, or are grounded in, human recurrent bodily movements through space, perceptual interactions, and ways of manipulating objects.

- Image schemas are highly schematic gestalts which capture the structural contours of sensory-motor experience, integrating information from multiple modalities.
- Image schemas exist as continuous and analogue patterns beneath conscious awareness, prior to and independently of other concepts.

Image schemas are direct building blocks that linguistic constructs are built on top of, particularly conceptual metaphor. Examples include SOURCE-PATH-GOAL, CONTAINER, OBJECT, UP-DOWN and FRONT-BACK (Hampe, 2005), which are also present in our cognitive understanding of music, for example: *up*beat, voice *leading*, harmonic *progression* and *goals* (Brower, 2000).

### 3.3.6. Image Schema in Music

Brower (2000) provides a distinction between image schema, patterns abstracted from bodily experience, and *music schema*, or patterns abstracted from musical conventions. He describes how melody fits the SOURCE-PATH-GOAL schema, as it can be described through terms of moving, with the goal of returning to the home chord: “Melody moves primarily by diatonic step, secondarily by chromatic step or arpeggiation. An unstable melodic pitch normally resolves downward and/or to its nearest stable neighbour. Melody normally comes to a point of final rest on the tonic.”

Godøy, Haga and Jensenius (2006) describes a similar concept in embodied sound cognition, which holds that listeners develop generalised schemata based on previous experiences of sound generation, which is then used in both the perception of familiar and unfamiliar sounds.

Wong (2011) explores image schema in DMI mapping design using Tangible User Interfaces (TUIs), particularly applying different image schema in mapping pitch, drawn from different languages and cultural perceptions. Wong explores the UP-DOWN image schema in pitch mapping, expressed in the conceptual metaphor MORE-IS-UP, as well as the SIZE image schema, used in the Kpelle language, and argues that both mappings provide a coherent interaction, although no empirical evaluations of the mappings are conducted.

## 3.4. Semiotics

As well as an embodied approach to meaning, the meaning of a gesture can also be analysed using semiotic principles. In fact, it has been argued that a mid-air gesture's only function is its *semiotic* property, as it does not provide action against another object (Miranda and Wanderley, 2006; Cadoz and Wanderley, 2000).

Semiotics was introduced by the linguist Ferdinand de Saussure and the philosopher Charles Sanders Peirce, and is concerned with the study of signs (Chandler, 2007). Saussure presents a two-part model of the sign:

- The *signified*, the concept being represented;
- The *signifier*, the representation of the concept.

For example, the word "tree" is a signifier, being a representation of a physical concept, a physical tree, which is the signified. Saussure stressed that the signified and signifier are "intimately linked" (Saussure, 1983), although stresses there is no intrinsic, direct or inevitable relationship between them, and that the connection is entirely arbitrary (Saussure's *rule of arbitrariness*).

Peirce's model consists of three parts:

- The *representamen* - the form which the sign takes;
- The *interpretant* - the sense made of the sign;
- The *object* - to which the sign refers.

An example is the traffic light sign for “stop”, which consists of: a red light facing traffic at an intersection (the representamen); vehicles stopping (the object); and the idea that a red light indicates that vehicles must stop (the interpretant) (Chandler, 2007).

Peircian signs are categorised as either a *symbol*, *index* or *icon* (Peirce, 1986). A symbolic sign is that in which the relation between the representamen and object is made by arbitrary rules, such as the noun “cat” in the English language; while an indexical sign is a sign that is indicative of something else, such as a tally kept by a farmer that indicates the number of his sheep; and the iconic sign is rendered significant in its resemblance to its referent in some form, such as a photograph of a face is an iconic sign of that person.

An icon is defined by Peirce (1986) as “a sign which refers to the object that it denotes merely by virtue of characters of its own, and which it possesses, whether any such object actually exists or not.” In other words, the iconic mode is that in which the signifier is perceived as *resembling* or imitating the signified (recognisably looking, sounding, feeling, tasting or smelling like it) - being similar in possessing some of its qualities (Chandler, 2007).

Semiotics has a history of use in musicology. What differs in semiotic musical meaning is that music carries immanent meaning, with what matters being the identification of sonic events internal to the musical system, without a reference to anything outside of it (Reybrouck, 2017). Semiotics allows for the study of “musical sense-making”, where music users construct and organise their knowledge of music using their own subjective observational tools (Reybrouck, 2012), thinking about sound at a symbolic level of representation, recognition and identification over a real-time experience (Reybrouck, 2017).

## 3.5. Ecological Approach

An ecological approach to musical meaning looks at the idea of subject and environment interactions through action/perception processes (Leman, 2008).

> When the subject perceives the sound as being produced by a physical mechanism, it may form the impression that the sound is real; and hence, that this environment has a high degree of presence, which in turn may facilitate the experience of being immersed (Leman, 2008, p. 166).

Leman (2008) argues that these action/perception processes are highly similar to embodied cognition, with the difference being that embodied cognition relates to the human cause of a sound, while the ecological approach explains our understanding of the cause of sound independent from human action. However, while a sound source may not be caused by human action, an embodied understanding of a sound is able to conceptualise the human action that *could* have caused it, as well as how the characteristics of that sound relate to the body. In other words, what it is like to *do* something and what it is like to *be* something (Cox, 2016).

## 3.6. Summary

In this chapter, literature around Embodiment has been presented. Section 3.1 discussed gesture in terms of bodily movement and perceived movement in music. Section 3.2 presented Embodiment and its application in interaction design and our understanding of music. Section 3.3 discussed Conceptual Metaphor Theory, which is underpinned by Embodiment, and its ability to reveal embodied understanding of music through language, as well as its use in DMI mapping design. Image Schemas are also discussed, which form the backbone of conceptual metaphors through spatial relationships. Section 3.4 presents a brief description of Semiotics, another approach to our understanding of music, while Section 3.5 discusses musical ecologies.

# 4. A Review of DMI Research from a UX Perspective

Chapter 2 presented literature around Digital Musical Instrument (DMI) design, with a focus on mid-air interaction, end-user mapping and the perspectives of different stakeholders in DMI interactions. Chapter 3 discussed embodiment and the perceptual link between movement, gesture and music. This chapter discusses how to effectively study DMI interaction, and the use of User Experience (UX) research methodologies in the DMI field.

UX methodologies are used to influence the empirical research undertaken in Chapters 6–9. This chapter presents a meta-review of the use of UX methods in the evaluation and study of DMIs. A series of UX dimensions (Bargas-Avila and Hornbæk, 2011) is used to examine a corpus of 132 papers from the proceedings of The International Conference on New Interfaces for Musical Expression (NIME), The Sound and Music Computing Conference (SMC) and The International Computer Music Conference (ICMC), to determine how UX is being used in the DMI research field.

A common theme in DMI research is how to effectively evaluate DMIs (Barbosa *et al.*, 2015; Jordà and Mealla, 2014; Hattwick and Wanderley, 2012; Kiefer, Collins and Fitzpatrick, 2008). The New Interface for Musical Expression (NIME) community, due to its historic connection to The ACM CHI Conference on Human Factors in Computing Systems (CHI), has often drawn from Human-

Computer Interaction (HCI) for its evaluation methodologies and frameworks (Hsu and Sosnick, 2009; Kiefer, Collins and Fitzpatrick, 2008; Wanderley and Orio, 2002). Traditionally, HCI evaluation methods focused on the concept of Usability (Nielsen, 1993), discussed in Section 4.1, but HCI research has moved on towards a focus on UX, discussed in Section 4.2.

The contents of this chapter have previously appeared in the following publication.

Brown, D., Nash, C. and Mitchell, T. (2017) A User Experience Review of Music Interaction Evaluations. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 370–375.

## 4.1. Usability

Usability is often simplified to refer to how easy something is to use. However, usability is a more nuanced idea with multiple components. When defining usability, research tends towards the ISO 9241 standard of "the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments". These three qualities are defined as:

- *Effectiveness*, the accuracy and completeness that a task can be achieved;
- *Efficiency*, the resources expended to achieve a specific goal;
- *Satisfaction*, the comfort and acceptability of a system.

Nielsen (1994) describes usability as five attributes: Learnability, Efficiency, Memorability, Errors and Satisfaction. These usability attributes are reflected in NIME research, with learnability and the learning curve (Jordà, 2004a; Wanderley and Orio, 2002) having been discussed in Section 2.5.1. While learnability in music and HCI shares similarities, in both cases it is seen as continuous measurement of improvement and not a binary “learned or not learned” state. In traditional usability, the user is considered to have learned the interaction once they can complete a task successfully (Nielsen, 1994), while the same is often the case in musical interaction, other more intrinsic measurements can be more useful, such as when a musician has developed enough skill so that the experience of playing is rewarding (Vertegaal and Eaglestone, 1996). Success in the former can be easily quantitatively measured, through metrics such as completion time or number of errors, while the nature of musical performance makes learning “success” difficult to determine, and is, individual to each performer. Performances can be graded by an examiner or adjudicator, such as in music examinations (ABRSM, 2019), but this is a measurement from an audience perspective, and is most often measured through qualitative, subjective metrics, with quantitative measurements rarely used.

Jordà (2004a) discusses efficiency in musical interaction, drawing from the definition from engineering of the ratio of useful energy output to energy input. His equation for musical instrument efficiency is as follows:

$$\text{Music Instrument Efficiency} = \frac{\text{Musical Output Complexity}}{\text{Control Input Complexity}}$$

This mathematical approach is potentially problematic, as it is trying to quantitatively express non-numerical and semantically subjective qualities such as “sonic richness”. Also, while an instrument with the simplest interaction and the most complex musical response may be the most musically “efficient”, it is questionable whether this is the metric DMI designers and researchers should prioritise. For example, an arpeggiator could be considered as more efficient than a piano as it has more musical output complexity. Efficiency is also a term associated with the domains of work, notably the early 20th century idea of Taylorism, a system for minimising "waste" movements performed by factory workers through the observation and analysis of their movements (Gertler, 2009), and the subsequent development of a "one best way" of performing factory work (Blythe *et al.*, 2004). While this led to a significant increase in production, the mindless, rote nature of Taylorised work lead to both repetitive strain injuries and the reduction in the quality of the worker's experience (Blythe *et al.*, 2004).

Similarly, error, previously discussed in Section 2.9, is an area of contention in music interaction. Fyans, Gurevich and Stapleton (2009b) argue that error is an important metric in DMI evaluations from the audience perspective, and that a spectator's ability to perceive error is necessary for them to be able to assess a performer's skill, and is linked to their mental model and understanding of a new musical instrument, as well as their perception of the instrument's *transparency* (Fels, Gadd and Mulder, 2002). From the performers perspective, the case for using error as an evaluation metric is debatable. While it can be argued that better performances are those that minimise errors, building instruments with the intention of minimising the potential for performance errors can lead to "musical toys" (Wessel and Wright, 2002), instruments that, being easily mastered, fail to provide engaging interactions for musicians.

Usability can sometimes be useful in musical contexts, but as discussed in this section the paradigm often does not fit the context of a musical interaction. When performing music, the objective is rarely to complete a task as efficiently and effectively as possible, but to communicate musical expression, recreationally enjoy music and provide entertainment to audiences. As such, the NIME field has seen a move away from usability as an evaluation metric, towards more subjective-based, experiential methods (El-Shimy and Cooperstock, 2016; Morreale, De Angeli and O'Modhrain, 2014; Makelberge *et al.*, 2012). This move away from usability is reflected in the wider HCI community, who have moved on to use UX theories and principles.

Despite this, usability should not be dismissed, as it sometimes is in UX research (El-Shimy and Cooperstock, 2016; Dillon, 2002, for example), and it remains a continually discussed concept in music interaction evaluations: Wanderley and Orio's seminal paper on usability for musical controllers (Wanderley and Orio, 2002) still receives multiple citations every year.

## 4.2. User Experience

UX represents a growing trend within HCI that focuses on a user's subjective experience of an interaction, and grew out of a counter-movement to usability (Hassenzahl and Tractinsky, 2006) that saw the paradigm as insufficient, placing too much emphasis on speed and accuracy in interaction (Dillon, 2002). It is often referred to as the "Third Wave" of HCI (Kiefer, Collins and Fitzpatrick, 2008), and while the exact definition of UX is one of great debate, researchers agree that it represents a dynamic, subjective and context-dependent approach to studying interaction (Law *et al.*, 2009).

Differences between UX and usability are often boiled down into "hedonic" qualities and "pragmatic" qualities respectively, something reflected in the popular AttrakDiff questionnaire tool (Hassenzahl, Burmester and Koller, 2003). While this represents an attempt to clearly distinguish between UX and usability, it gives an overly simplistic impression, and obstructs many facets that contribute to an experience, such as eudaimonic qualities of self-worth and meaning (Mekler and Hornbæk, 2016), that are evident in musical interactions.

More nuanced paradigms have been put forward for UX. Dillon (2002) defines a three point paradigm, mirroring the three aspects of usability, of *process*, what the user does; *outcome*, what the user attains; and *affect*, what the user feels. Another three-pronged approach is presented by Hassenzahl and Tractinsky (2006):

- *Beyond the Instrumental*, going beyond the task-based model of interaction;
- *Emotion and Affect*, how a user feels and is affected by the interaction;
- *The Experiential*, the context and temporality in which the technology is used.

This conceptualisation is useful for DMI research, especially for musical performance contexts. While some aspects of musical interaction can be considered task-based, such as learning a scale or practising a specific piece, considering a musical performance to be a task that a musician completes can leave out important aspects such as musical expression. In Hassenzahl and Tractinsky's UX framework, the emotion and feelings felt during an interaction, which could include musical expression, are explicitly considered. It is also important to consider the context in which a musical performance takes place, as this has a significant impact on the relationship between and experiences of performers and their audiences.

A user experience framework has been explicitly established for musical interface design: MINUET (Morreale, De Angeli and O'Modhrain, 2014). MINUET is based on the PACT framework (Benyon, Turner and Turner, 2005), a user-centred design technique that focuses on *People*, the end-user, *Activities*, the interactions taking place, *Contexts*, when and where interactions are happening, and *Technologies*, the software and hardware involved. MINUET applies this to a musical context (Table 4.1), and considers technological specifications over actual hardware or software implementations.

In these UX paradigms, what is clear is the importance of context in an evaluation: there is no "one size fits all" solution to evaluation in music interaction (El-

| PACT | MINUET |
|---|---|
| People | Performers<br>Audiences |
| Activities | Motivation<br>Collaboration<br>Learning Curve<br>Ownership |
| Contexts | Musical Style<br>Physical Environments<br>Social Environments |
| Technologies | Control<br>Mapping<br>Operational Freedom<br>Embodied Facilitation<br>Input<br>Feedback |

Table 4.1.: The MINUET framework (Morreale, De Angeli and O'Modhrain, 2014) in relation to PACT (Benyon, Turner and Turner, 2005).

Shimy and Cooperstock, 2016), or in UX. In each case, the study must take experiential, contextual properties into account. Another similarity is the distinction between instrumental and non-instrumental studies. A study of a new DMI using a usability-based approach may ask a participant to perform a piece of music (a task), while quantifiable measurements are taken of the accuracy of the performance (Wanderley and Orio, 2002). A UX study may ask a participant to use the instrument with little or no guidance or direction, and elicit more qualitative data about their experience (Stowell, Plumbley and Bryan-Kinns, 2008). This could be over a short space of time or over several weeks or months (Nash and Blackwell, 2011; Gelineck and Serafin, 2012). Previous work (Nash and Blackwell, 2011; Collins, 2007; Shneiderman, 2007) advocates for longitudinal, observational and qualitative approaches in studying creative tools, moving past the "old strategies" of laboratory-controlled, short-term usability studies.

## 4.3. Meta-Review Objectives

The rest of this chapter details a meta-review of DMI papers from the perspective of UX methods. The meta-review seeks to provide an alternative perspective of music interaction evaluations than previous work (Barbosa *et al.*, 2015), and highlight areas of potential for new research in the DMI field.

Bargas-Avila and Hornbæk (2011) conducted a meta-review of empirical studies in the UX field to understand how its ideas are being used in primary research, finding a series of "dimensions" of UX that are evaluated regularly: generic UX, emotion, enjoyment, aesthetics, hedonic qualities, engagement, motivation, enchantment and frustration. To observe recent trends of usability and UX in music interaction, a meta-review of recent papers from the computer music field was conducted from the perspective of UX, using these dimensions to classify the evaluations.

Analysing recent literature from NIME, SMC and ICMC involving empirical user-focused evaluations, the review focuses on:

1. The stakeholders considered in the evaluations.
2. The dimensions of UX that are evaluated.
3. What participant tasks are used.
4. How data is collected.

## 4.4. Method

The method was drawn from an adaptation of the QUOROM method (Moher *et al.*, 1999) used by Bargas-Avila and Hornbæk (2011). The corpus as been filtered as follows (Figure 4.1).

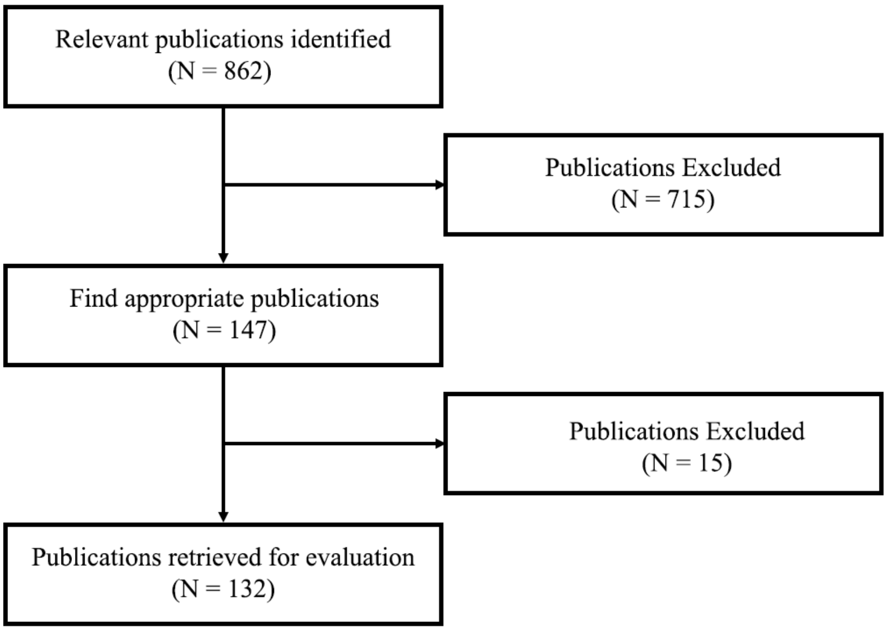


Figure 4.1.: The corpus filtering procedure.

1. **Identify sources.**

   Source Selection: Conference proceedings of NIME, SMC, and ICMC for 2014 – 2016 (N = 862).

2. **Find appropriate publications.**

   Screening criteria: Papers that mention an empirical user study in the title or abstract, using the keywords *Evaluat[e,ion,ed,ing], User, Study*. (N = 147).

3. **Publications retrieved for detailed evaluation.**

   Screening Criteria: Papers of which the evaluation focuses on the user (N = 132).

4. **Final Corpus.**

   The final corpus for the meta-analysis consisted of 132 papers.

For the UX dimensions, similar dimensions were chosen to those found by Bargas-Avila and Hornbæk (2011) in their meta-analysis to be prominent aspects evaluated in the UX field. It was also noted when papers focused on usability, allowing for its comparison against the dimensions of UX. The following definitions were used in categorising the corpus.

1. **Usability** Evaluations cover concepts such as ease of use, effectiveness and ergonomics, efficiency and learnability (Nielsen, 1993).
2. **Generic UX** Evaluations take a holistic approach and seek to explore the participants' experiences as a whole, without focusing on any specific dimensions.
3. **Aesthetics** Evaluations focus on the aesthetic, artistic properties of the experience (Luhtala *et al.*, 2012), such as appeal, taste, style, and expression (Danto, 1981).
4. **Emotion** Evaluations measure the emotional response and feelings of participants.
5. **Enchantment** Evaluations focus on the affective attachment of people to technology (McCarthy *et al.*, 2006).
6. **Engagement** Evaluations study flow (Csikszentmihalyi, 1996), intrinsic interest and curiosity (Chapman, Selvarajah and Webster, 1999).
7. **Enjoyment** Evaluations focus on the hedonic qualities of interaction (Blythe *et al.*, 2004).
8. **Motivation** Evaluations focus on what drives a participant's decisions and behaviour (Evans, 1975).
9. **Frustration** Evaluations focus on the participant's dislikes and hindrances during an interaction (Mendoza and Novick, 2005).

In a similar method to Barbosa *et al.* (2015), The stakeholders in each evaluation have been identified, using the following categories:

1. **Performers** Participants with agency, actively affecting their experience of real-time auditory interaction.

| Year | Conference | No. of papers |
|---|---|---|
| 2014 | NIME | 28 |
| | ICMC/SMC | 32 |
| 2015 | NIME | 19 |
| | ICMC | 6 |
| | SMC | 16 |
| 2016 | NIME | 15 |
| | ICMC | N/A |
| | SMC | 16 |

Table 4.2.: Breakdown of papers.

2. **Audiences** Participants without agency in the evaluation, passively involved in the experience.
3. **Designers** Participants with agency in evaluations that involve creating or designing hardware or software.
4. **Composers** Participants with agency in evaluations that involve composing or creating artistic material, but not performing.

It is important to give these definitions as some of the evaluations do not follow a traditional performance framework (Grani *et al.*, 2016, for example), where each participant is asked to play an auditory game. As the participant is actively engaging in a task, they have been categorised as a performer.

## 4.5. Results

Due to the analysis taking place before the ICMC 2016 proceedings were available, the small number of relevant ICMC 2015 papers (N = 6) and the joint ICMC/SMC conference of 2014, the decision was made to group the ICMC and SMC papers together in the analysis. The breakdown of papers used in the analysis can be found in Table 4.2.

The analysis was non-exclusive, with some evaluations covering more than one UX dimension, data collection method, stakeholder or participant task. If more than one evaluation was included in a single paper they were recorded as separate results. After the analysis, the following categories were identified for the participant tasks:

1. **Specific Task** Participants are asked to perform a pre-determined exercise, such as listen to auditory stimulus, or perform certain tasks with an instrument.
2. **Open Exploration** Participants are free to do as they please during an interaction.
3. **Guided Exploration** Participants have some freedom, but are guided by certain constraints.
4. **Watch Performance** Participants watch a performance given by a musician, in either a concert or laboratory setting (e.g. watching a video).
5. **Prepare and/or Give Performance** Participants are asked to prepare a piece and give a performance as part of the evaluation.
6. **Workshop** Participants' interactions take place in a workshop setting.
7. **In The World Use** Participants use the technology in their own personal environments.
8. **Other** Any other task that does not fit in the above categories.

The following data collection methods were also identified:

1. **Questionnaires** Specific questions used to gather responses.
2. **Likert Scales** Questionnaires use the Likert format.
3. **Comparisons** Participants are asked to compare stimulus, and give ratings; perform pair-wise comparisons and the like.
4. **Interviews** Either structured or unstructured.
5. **Field Notes** Observations are taken by researchers during the evaluation.
6. **Audio/Video Recording** Recordings of experiment are used in the analysis.

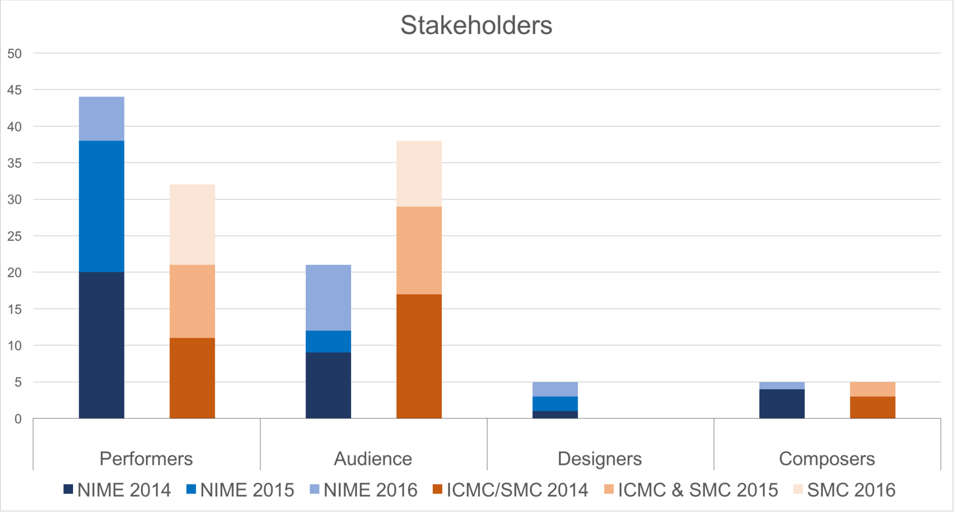


Figure 4.2.: Stakeholders

7. **Interaction Log** The user's interaction with an interface is logged.
8. **Open/Informal Comments** Unstructured feedback is provided.
9. **Created Materials** Things made by participants during the evaluation are analysed, e.g. (McPherson and Zappi, 2015).
10. **Physiological Measurements** Methods such as EEG, ECG and the like are used to record a participant's body.
11. **Other** Any other method that does not fit in the above categories.
12. **NS** The data collection method is not specified.

### 4.5.1. Stakeholders

The most popular stakeholder used in evaluations was the performer (50.7%), followed by the audience (39.3%), while designers (3.3%) and composers (6.7%) perspectives were rarely evaluated (Figure 4.2). While it has been suggested that performers are the most important stakeholders in digital music (Birnbaum *et al.*, 2005), the results suggest that the perspectives of designers and composers could be better represented during evaluations, as these perspectives may reveal aspects of musical interactions that have previously been overlooked.

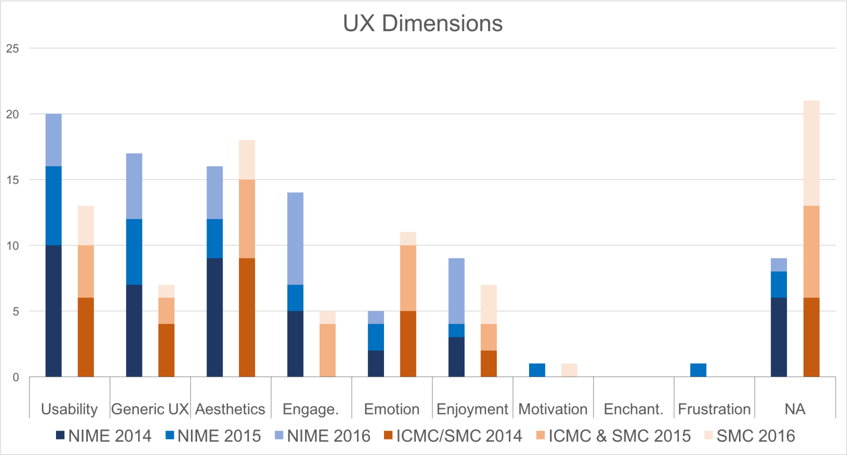
The stakeholders results are quite different to those of Barbosa *et al.* (2015), whose stakeholders results were Performers: 52, Designers: 28 and Audience: 20. This is most likely because of their inclusion of technical evaluations as evaluations from the designer's perspective. Since the focus of the research was on evaluations with participants the designers result is low, as a designer's subjective experience is not usually solicited during technical evaluations.

### 4.5.2. UX Dimensions

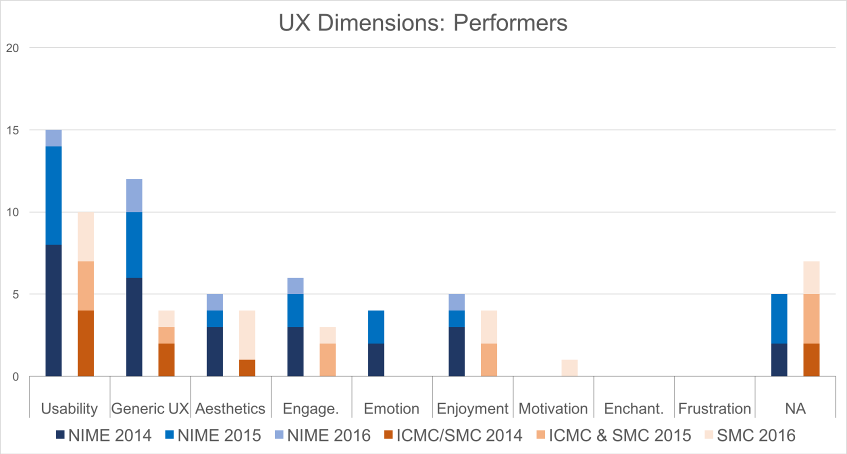
The results (Figure 4.3) indicate that although UX concepts are being applied in computer music research, usability remains a popular metric in NIME papers (21.7%), while within ICMC and SMC, the largest proportion were not applicable to dimensions of UX, such as when an audience's perception of vibro-tactile feedback is measured (Fontana *et al.*, 2016). A high amount of not applicable papers is to be expected, and is most likely due to the fact that empirical evaluations in computer music research do not always share the same targets as UX research, and so a large number of papers will not fit within the scope of this research.

Of the dimensions of UX, aesthetics is the most commonly used (19.4%), followed by generic UX (13.7%) and engagement (10.9%). This reflects the literature of the field, which highlights the importance of expression (Dobrian and Koppelman, 2006), style (Gurevich, Stapleton and Marquez-Borbon, 2010; Jordà, 2004b) and engagement (Wessel and Wright, 2002) in computer music research. Generic UX papers often included evaluations with less formal structures, such as when a group of children are used to evaluate a museum experience through open exploration and group interview (Jørgensen *et al.*, 2015). This reflects the ideas of Stowell, Plumbley and Bryan-Kinns (2008) in their proposed qualitative method.

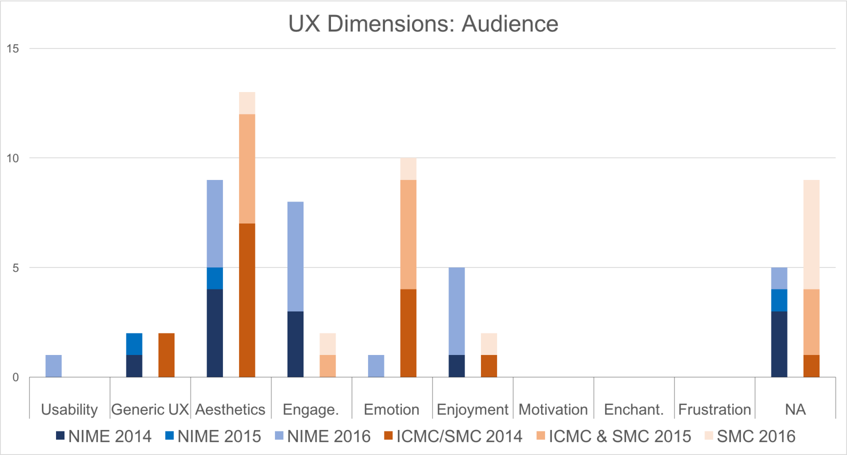
Emotion and enjoyment were evaluated in relatively equal measure (9.1%), but emotion evaluations in ICMC/SMC occurred only from the audience's perspect-



(a) Use of UX Dimensions for all DMI research.



(b) UX Dimension use in performer-perspective research.



(c) UX Dimensions use in audience-perspective research.

Figure 4.3.: Use of UX Dimensions in DMI research.

ive. Three dimensions: motivation, enchantment and frustration; were evaluated for either rarely or not at all. This suggests that these are areas of UX that are currently overlooked in music interaction, and represent an opportunity for new directions in research. For example, studying how musicians become affectionately attached to an instrument may help us understand long term uptake of DMIs, while studying motivation may allow us to explore their appeal over traditional instruments.

Although frustration is often linked to measurements of user error used in usability studies, in UX, frustration represents a qualitative exploration of negative aspects of a user's experience, for example the work by Blythe *et al.* (2006).Its study could help the computer music community identify areas for improvement in the design of DMIs and music interaction technology.

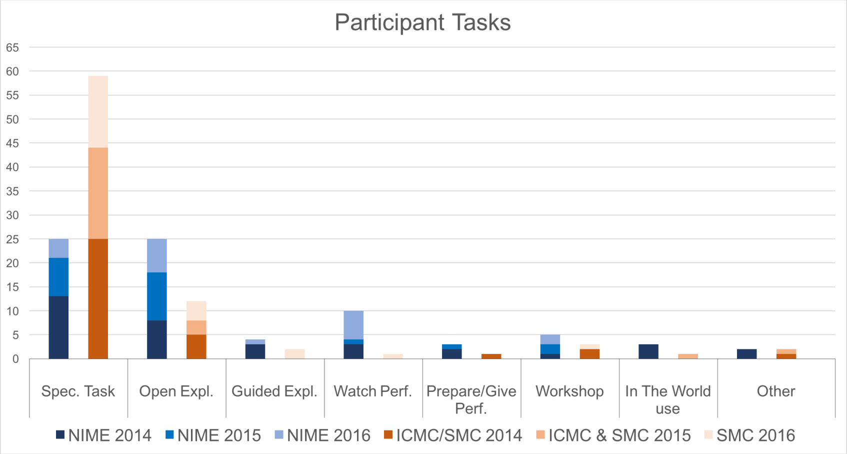
### Performers

From the performer's perspective, usability was found to be the most prominent dimension (29.4%), followed by generic UX (18.8%) while aesthetics, engagement and enjoyment share a similar proportion (10.5%). While NIMEs and ICMC/SMC have different quantities of performer evaluations, they have a similar spread of evaluation dimensions, with usability being the most popular.

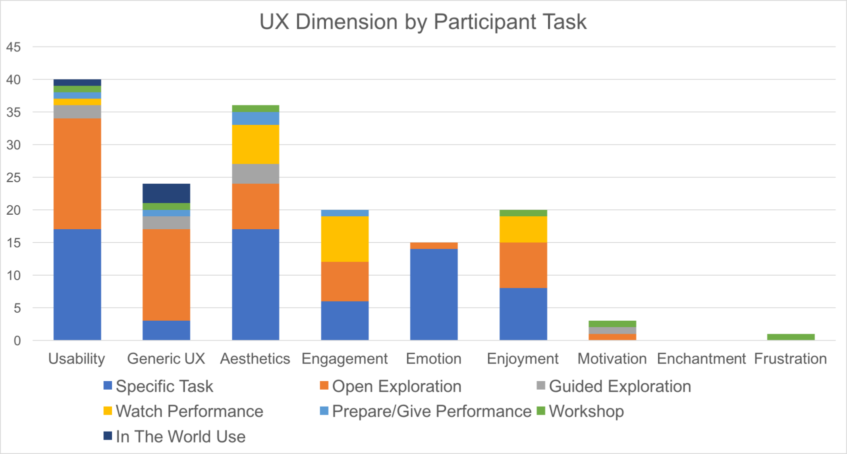
Usability remains prominent most probably because of its close relation to ideas of learnability and playability, which are important ideas in NIME and computer music research.

### Audience

Aesthetics was the most prominent dimension from the audience's perspective, in both ICMC/SMC and NIME. Interestingly, emotion was commonly studied within SMC and ICMC, while it was rare within NIME evaluations. NIME often



(a) Tasks given to participants.



(b) Tasks used for UX dimensions.

Figure 4.4.: Participant Tasks

focused on engagement and enjoyment while ICMC/SMC evaluations rarely did so.

### 4.5.3. Participant Tasks

Overwhelmingly, the most popular participant tasks (Figure 4.4) were specific tasks (53.1%), which make up the majority of ICMC and SMC evaluations. Meanwhile, NIME evaluations use specific tasks and open exploration in equal measure. The other tasks were used much less frequently.

Interestingly, NIME evaluations include watching performances much more than ICMC and SMC. This could be due to NIME's focus on instruments, which suit audience evaluation through performance.

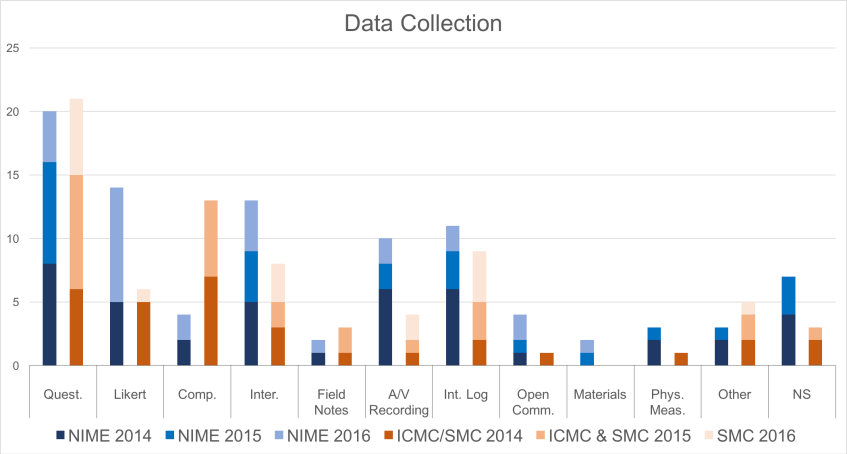
When filtered by UX dimension, it is interesting to observe that while questionnaires are the most popular technique for most dimensions, open exploration is the most popular for generic UX. This reflects the dimension's less focused approach, in that via open exploration, any aspect of the interaction may be explored by participants. Similarly, "in the world" use appears mostly in generic UX, as this technique also encourages an open response from participants.

Meanwhile, emotion is studied nearly exclusively using specific tasks, with evaluations often asking audience participants to report on their emotions after listening to musical stimuli. The dimensions of aesthetics, engagement and enjoyment are each studied using a wide range of tasks, but most prominently specific tasks, open exploration and watching performances.

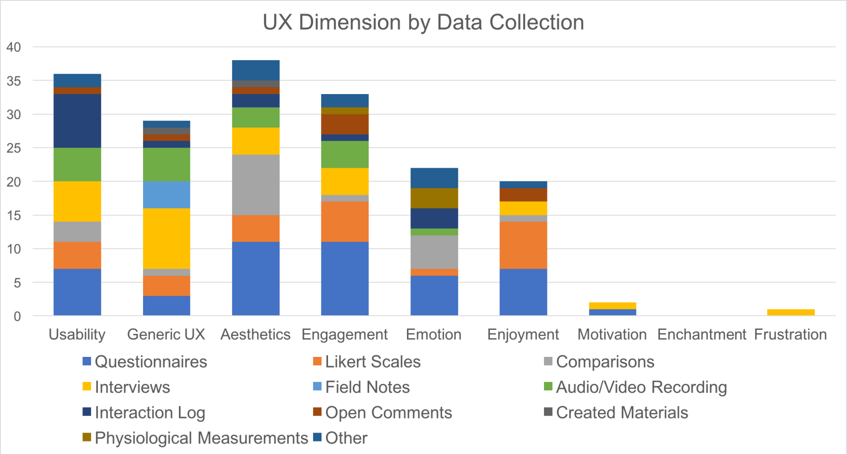
### 4.5.4. Data Collection

The most popular method of collecting data (Figure 4.5) was by questionnaire (24.6%), and the results reflect those of Barbosa *et al.* (2015). Due to their prominence, questionnaires formatted as Likert scales were included in their own category (12.0%). Questionnaires most likely remain a popular technique as they give evaluations an ability to focus on specific aspects, and quantitatively analyse otherwise qualitative elements of an interaction.

Interaction logs are used mainly to measure usability. This reflects the evaluation technique of Wanderley and Orio (2002), as well as Kiefer, Collins and Fitzpatrick (2008), which use interaction logs to provide quantitative data for usability measurements.



(a) Data collected by conference.



(b) Data collected by UX Dimensions.

Figure 4.5.: Data Collection.

Interviews and field notes were mostly used to measure generic UX, while questionnaires are rarely used. This also reflects the open nature of the dimension, as interviews and field notes do not limit a participant's response.

Comparisons, such as pair-wise comparisons and preference ranking, are most commonly used to measure aesthetic qualities.

While it was found that emotional responses are elicited using specific tasks, they are collected using a wide variety of methods, including specific emotion measurement tools like the Self-Assessment Manikin (SAM).

## 4.6. Discussion

The results indicate that there is a strong correlation between UX and the evaluation criteria used in computer music research. However, usability remains the most prominent idea from HCI used in the field, despite efforts to move the field towards UX theories and principles.

This research has found three common dimensions in UX research: motivation, enchantment and frustration; that are evaluated rarely or not at all in computer music interactions. These areas could help to address key questions regarding digital musical instruments, and help us to better understand the nature of the instruments and technologies we create. For example, looking at enchantment and the way in which musicians become emotionally attached to DMIs may help to us to understand how short-term experimenters become long-term practitioners; understanding what motivates and influences musicians to choose DMIs could enable us to design in ways that encourage new players; and studying frustration in DMIs could help us to design more enjoyable and engaging music interaction experiences.

While these dimensions are inherently very different from each other, they share a very qualitative nature. Examples of their previous use in HCI literature use descriptive case studies (Ross *et al.*, 2008) and highlight the need for *"rich personal accounts"* (Sengers *et al.*, 2008). This more qualitative perspective is also shared with much research in the DMI community, and highlights the growing trend in both UX and music interaction towards deeper explorations of a user's subjective experience (such as the work presented in Chapter 8), as well as the potential ease with which these dimensions could be adopted into music interaction research.

As well as the UX dimension findings, the research has found that specific tasks are the most popular participant task used in evaluations, and data is most commonly collected through questionnaires. While these are tried and tested methods, it indicates that there is room within computer music evaluations for the use of alternative methods, which may help us to evaluate technologies more thoroughly. For example, studying how musicians use instruments in their own personal environments ("in the wild") allows us to better examine their creative process, as it is difficult to capture this in laboratory environments (Gelineck and Serafin, 2012).

Similarly, the tasks of watching and preparing for a performance reflect real world use cases for musical technology, and we can learn much from studying the dynamics behind these processes. As every evaluation needs to be tailored to the specific goals and needs of the research in question (O'Modhrain, 2011), a full discussion of how these findings should affect future evaluations is beyond the scope of this study, and is an area that researchers must consider for their own evaluations.

The analysis may have benefited from delineating between individual and group stakeholders, which would have provided a deeper insight into the user experience of multi-user interactions, such as collaborative installations. Also, breaking specific tasks into subcategories (for example into listening exercises and performance tasks) would have allowed for more detailed analysis of participant tasks.

By reviewing which areas of UX are commonly evaluated in music interaction research and which are overlooked, alongside the participant tasks and data collection methods used, this research has provided a new perspective on the interaction evaluations taking place, and revealed alternative qualities to be considered in future NIME research.

The results suggest that while much work highlights NIMEs move away from usability and traditional HCI research, its ideas are still being remain significant in the music interaction field.

One aspect of a user's experience with a DMI that has not been covered by this analysis is self-efficacy, or the "the conviction that one can successfully execute the behaviour required to produce the outcome" (Bandura, 1997). Within music performance, self-efficacy can be used to determine a user's belief that they would be able to achieve certain performance goals with an instrument, and is linked to a performer's intrinsic motivation to learn and develop skill with an instrument (McPherson and McCormick, 2006).

In terms of evaluating a DMI, measuring self-efficacy before and after a user's interaction with an instrument would reveal a well designed system, with a user able to perceive more expressive possibilities once they have spent time exploring the instrument, and as their confidence with it increases. For example, through exploratory play, new users to the D-Box were able to perceive and develop creative and innovative performance techniques as their familiarity and confidence with the system increased (McPherson *et al.*, 2016).

This ability to perceive expressive capabilities could also be affected by existing reference material, such as online videos and forum discussions, which would allow new musicians to learn about the performance techniques of existing practitioners. This is particularly apparent for the Mi.Mu Gloves, whose high profile users have videos with hundreds of thousands of views (WIRED UK, 2013).

Understanding self-efficacy would also help DMI researchers to better contextualise and account for the nature of their participants, often self-selecting, who are more likely to be confident musicians with no qualms about experimenting with potentially unreliable or unpredictable instrument prototypes.

## 4.7. Summary

This chapter presents a meta-review of DMI research that examines how UX methods are being used to evaluate DMI interaction. Sections 4.1 and 4.2 presented background on the concepts of Usability and UX respectively, and Sections 4.3 to 4.6 present the meta-review.

The meta-review found that usability is still commonly used to evaluate DMIs, while aesthetics is most the commonly used dimension of UX used in evaluations. Three areas of UX: motivation, enchantment and frustration; are often overlooked in current interaction evaluations, and represent potential avenues for future evaluations. As well as this, it has been found that questionnaires are the most popular method of data collection, and specific tasks are the most common participant tasks.

This chapter concludes the review of existing literature that has informed the empirical work that makes up the rest of this dissertation. The next chapter (Chapter 5) will introduce the Mi.Mu Gloves and their dedicated end-user mapping software, *Glover*, that allows glove musicians to map mid-air movement to musical MIDI data (Mi.Mu Ltd., 2018). The Mi.Mu Gloves and Glover are used in the later empirical research (Chapters 6–9) to examine the research questions proposed by this dissertation.

# 5. Glover and the Mi.Mu Gloves

The previous chapters (Chapters 2, 3, and 4) have presented existing literature relating to Digital Musical Instruments (DMIs), embodiment, and a systematic review of the use of User Experience (UX) research techniques and methodologies within DMI literature.

This chapter presents the Mi.Mu Gloves, the DMI used in the empirical research (Chapters 6–9), and their dedicated end-user mapping application: Glover, which allows glove musicians to map the mid-air movements detected by the gloves to musical MIDI data. The chapter presents the mapping affordances and constraints of the system.

The next chapter (Chapter 6) begins the empirical work conducted with the Mi.Mu Gloves, presenting an investigation into the end-user mapping practice of existing glove musicians.

Elements of this chapter have previously been presented in the following publications:

Brown, D., Nash, C. and Mitchell, T. (2018) Understanding User-Defined Mapping Design for Mid-Air Musical Performance. In: *Proceedings of the 5th International Conference on Movement and Computing (MOCO)*. Genoa, Italy, June 2018. ACM. 27:1–27:8.

Brown, D., Nash, C. and Mitchell, T. (2018) Simple Mappings, Expressive Movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity*. 29 (2–3), pp. 129–148.

## 5.1. The Gloves

The Mi.Mu Gloves are data gloves that provide a wearable mid-air interface for controlling music (see Section 2.3). The Glove project began through a collaboration with Tom Mitchell and Imogen Heap, and Heap's motivation to take electronic music performance away from the laptop and into a visual, gestural performance domain (Mitchell and Heap, 2011). The gloves detect a wearer's finger positions with flex sensors and hand orientation with an Inertial Measurement Unit (IMU), providing an accurate representation of a hand's posture and orientation in space. They also include auxiliary sensors and feedback devices in the form of buttons, LEDs and vibration motors (Figure 5.1).

The Mi.Mu Gloves have been fortunate enough to have developed a community of musicians beyond their initial practitioner. The commercial and artistic success of Heap before her adoption of the Gloves into her performances may lend a level of authenticity to the gloves: Heap's audiences may be more open to accepting her new instrument as they already have an appreciation and admiration of her compositional and performing ability. The wealth of online content and videos from Heap, Grande and other Glove musicians (Grande, 2015; WIRED UK, 2013) also provides a cultural context that other instruments lack, with this content allowing new Glove musicians to discover performance techniques and ideas that have been developed by experienced users. The demand for the gloves has now resulted in the incorporation of a company (Mi.Mu Ltd., 2018), with 2019 seeing the gloves become more widely commercially available.

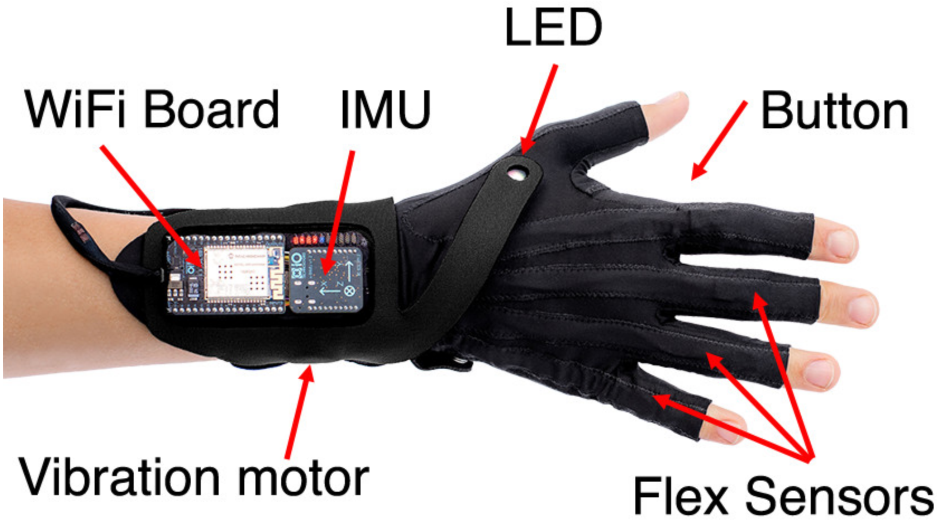
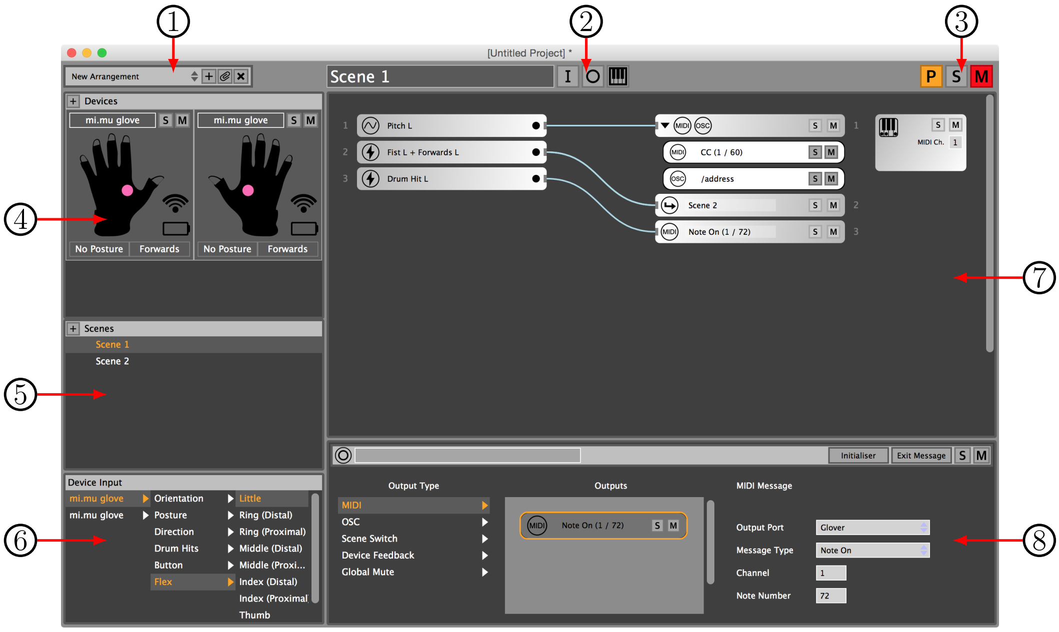


Figure 5.1.: A Mi.Mu Glove.

The glove's community of practitioners, although small, represents a rare resource in the DMI field. Few new DMIs have a community of dedicated users, let alone users who have been practitioners for several years.

## 5.2. Glover

The Mi.Mu Gloves are used alongside a dedicated mapping software application: *Glover* (Figure 5.2), which allows end-users to design mappings between gesture features and sound parameters (Section 2.6). The software converts the raw sensor data provided by the gloves to a variety of parameters that represent gesture features, such as postures, movements and hand direction (more detail in Table 5.1), and provides an interface for connecting these to musical features in the form of MIDI or OSC messages, as well as to glove-based feedback in the form of vibration pulses or LED settings (Figure 5.3).



Output Port: Glover
Message Type: Note On
Channel: 1
Note Number: 72

Figure 5.2.: An example Glover project.

1. Arrangement selection.
2. New mapping options.
3. Perform mode, solo and global mute.
4. Device overview panel.
5. Scene overview panel.
6. Device input panel.
7. Mapping panel.
8. Inspector panel.

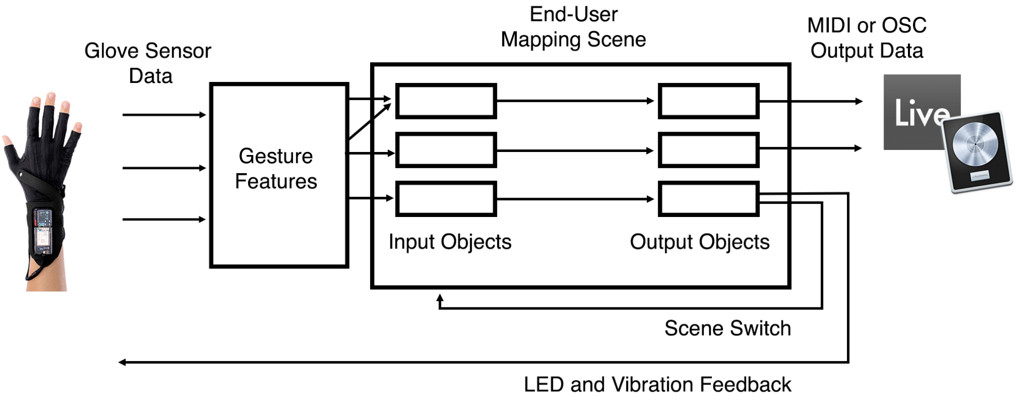


Figure 5.3.: A schematic representation of end-user mapping with Glover.

| *Type* | *Category* | *Parameter* |
|---|---|---|
| Movements | Orientation | Roll |
| | | Pitch |
| | | Yaw |
| | Flex | Average Finger Flex |
| | | Individual Finger Flex |
| Qualifiers | Postures | Fist |
| | | Open Hand |
| | | Puppet Hand |
| | | One Finger Point |
| | | Custom Postures |
| | Directions | Up |
| | | Down |
| | | Left |
| | | Right |
| | | Forwards |
| | | Backwards |
| | Button | Button Up |
| | | Button Down |
| Events | Hits | Slap |
| | | Drum Hit |
| | | Wrist Flick |

Table 5.1.: The mapping lexicon of Glover.

### 5.2.1. Gesture Features in Glover

Glover presents users with three types of gesture features:

- *Movements:* continuous controls derived from body movements, such as the pitch, yaw and roll of the wrist, and the amount of flex of each finger.
- *Events:* controls that notify that a specific action has occurred, often set gestures such as "drum hits".
- *Qualifiers:* state-based controls that can either be occurring or not, such as specific hand postures or directions.

Within these feature types, mapping design with Glover uses a specific set of gestural features, which make up a lexicon of control options (Table 5.1).

Figure 5.4.: The default postures (clockwise from top left: fist, puppet hand, open hand, one finger point).

### 5.2.2. Posture Recognition

A significant aspect of glover interaction is the ability for users to train the software to recognise postures using machine learning techniques through the Glover software's *Posture Recogniser*. Originating from Heap's early work with the Gloves, the postures "Fist", "Puppet Hand", "Open Hand" and "One Finger Point" (Figure 5.4) have become a standard part of practice in glove performance, and are included by default in Glover's Posture Recogniser.

## 5.3. Mapping in Glover

Mapping in Glover is achieved by explicitly defining the connections between gestural and musical features (Section 2.5.2). The main Input–Output Object mapping interface uses a patch cord metaphor (Figure 5.5) in a similar manner to visual

Figure 5.5.: Left hand pitch (movement) and Open Hand posture (qualifier) control the value of a MIDI CC message.

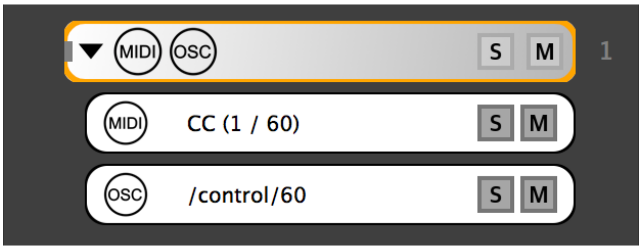
programming tools such as PureData and Max/MSP. Users create new mapping input and output objects using tool bar controls, and connections between them are made by clicking-and-dragging from an input object to an output object. This interface provides the most mapping affordances, with any gestural feature being able to be mapped to any output type. The Instruments (see Section 5.3.3) represent mapping *macros* that provide higher level abstractions of common mappings used by glove musicians, and were added to the software after feedback from glove musicians. They have less diverse utility, being only able to map specific gestural features to MIDI Note messages, but enable glove musicians to implement commonly-used mappings quickly.

### 5.3.1. Input Mappings

Mapping input objects can combine multiple gesture features, but with some limitations. Only one movement or event parameter can be used per input object, but any number of qualifiers can be used. A simple example is shown in Figure 5.5, where the pitch angle of the left hand is used to control a MIDI CC value, and only when an open hand posture is made.

### 5.3.2. Output Mappings

Each mapping output object represents a destination for an output value generated by the input object connected to it. Values can be sent externally to Digital Audio Workstations (DAWs) or other applications as either MIDI or OSC, for ex-



MIDI CC (1 / 60)

OSC /control/60

Figure 5.6.: Multiple outputs for output value.

ample MIDI CC 60 and OSC "/control/60" (Figure 5.6); or can be used to trigger feedback in the gloves.

### 5.3.3. Instruments

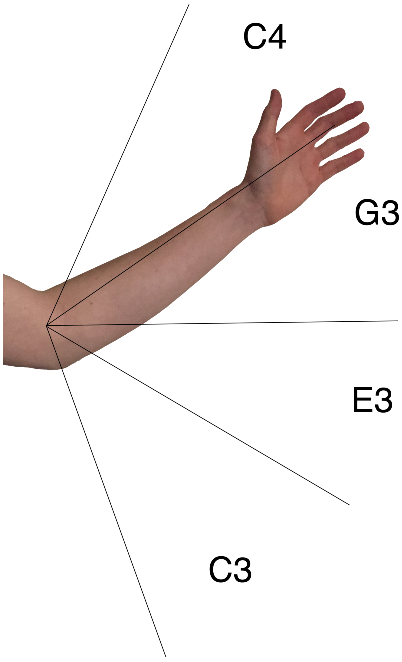
Glover includes two types of mapping *Instruments* that are useful macros for mapping gestural features to multiple MIDI notes at once. The instruments were added to Glover after feedback from glove users. Both instruments combine input and output mapping into a single Inspector Panel window (Section 5.3.5).

The *Note Matrix* (Figure 5.7) splits a movement axis (such as the pitch of the wrist) into a series of thresholds, which, when crossed, trigger notes in a chosen scale. The Note Matrix is only capable of to-many mappings, as if only one note is set as an output, no triggering thresholds are generated.

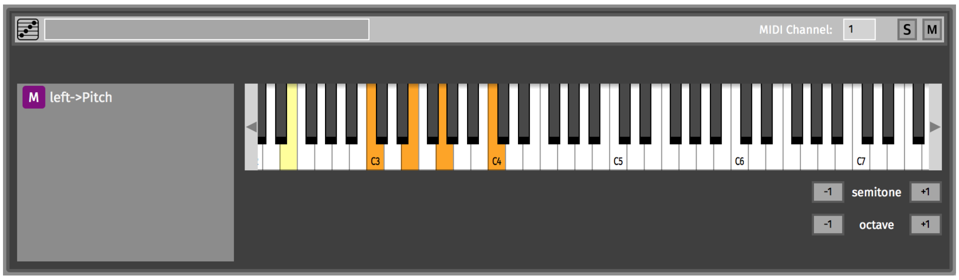
The *Chord Machine* (Figure 5.8) allows for multiple notes to be played at once (e.g. chords), triggered by selected qualifiers.

### 5.3.4. Mapping Organisation

The software also allows users to switch between multiple mapping designs, with each strategy being contained in a *scene*, which are in turn organised into *arrangements* (Figure 5.9). This allows musicians to *scene switch*, which is generally used to perform more musical material than would be cognitively or ergonomically possible at once. For example, one scene may be used to map events to trigger drum



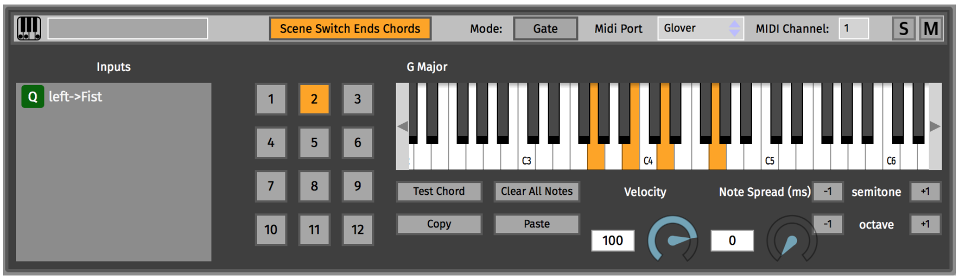
(a) Gestural representation.


MIDI Channel: 1

(b) Implementation in Note Matrix Inspector Panel.

Figure 5.7.: The Note Matrix splitting the pitch of the left hand into a C major chord

(a) Gestural representation.



Mode: Gate

Midi Port: Glover

MIDI Channel: 1

Q left->Fist

Velocity: 100

Note Spread (ms): 0

(b) Implementation in Chord Machine Inspector Panel.

Figure 5.8.: The Chord Machine being used to map a a fist posture to a C major chord, and an open hand to a G major.

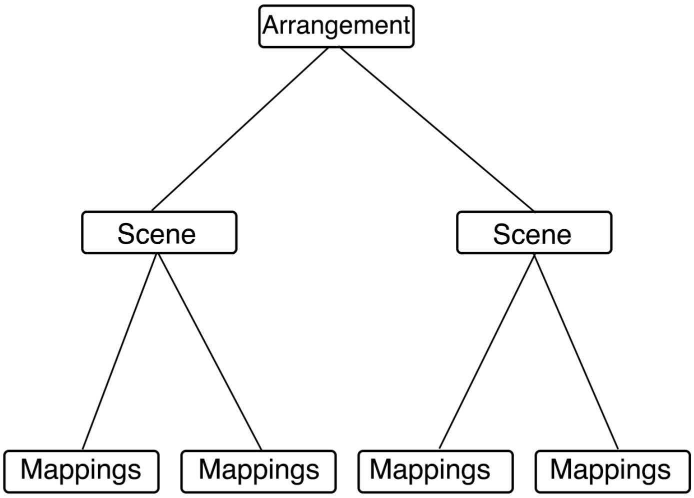


Figure 5.9.: How mappings are organised in a Glover project.

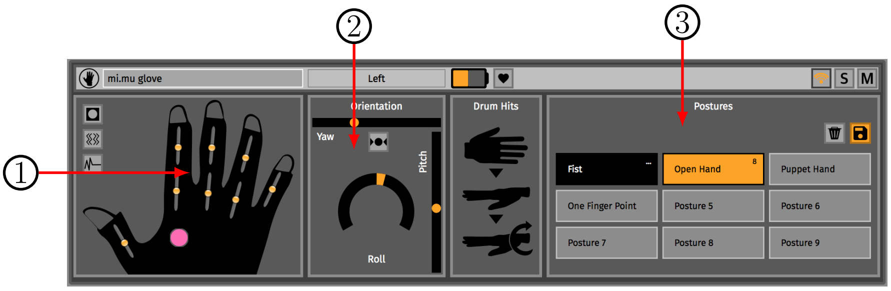
samples, while another may map movements to synthesiser parameters. Scenes are a flexible way of organising mappings: they can be switched between on-the-fly, and can either be always active, allowing controls to be used when other scenes are in focus, or set to be exclusively active only when that particular scene is in focus. Arrangements, comprised of one or more scenes, cannot be switched between on-the-fly, and are generally used for entire songs or pieces, with the gloves' calibration settings and posture training maintained between them.

### 5.3.5. The Inspector Panel

Most editing in Glover is done in the Inspector Panel (Figure 5.2 no. 8, Figure 5.10), which displays the details for mapping objects and device settings, as well as the details of the two mapping instruments.

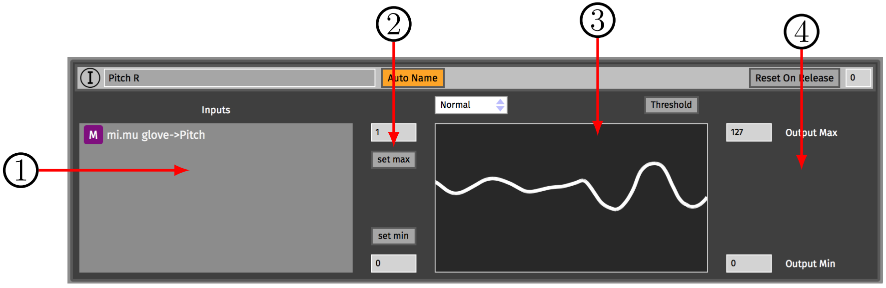
## 5.4. Mapping Affordances of Glover

The affordances of a piece of technology, or the perceived possible actions available to users, heavily influences user interactions (Norman, 1988). This also ap-



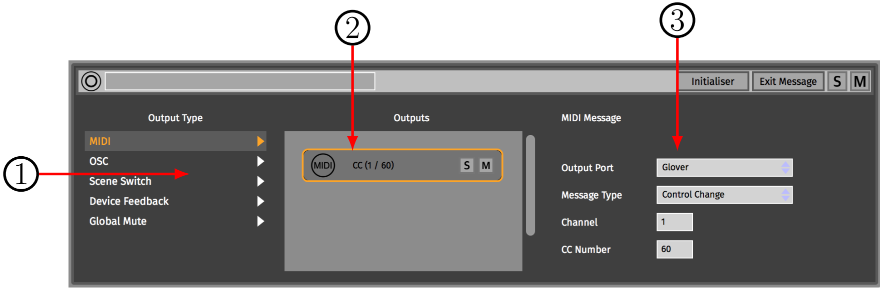
(a) Mi.Mu Glove Inspector.

1. Flex sensor feedback; button, calibration and vibration controls.
2. Orientation feedback.
3. Posture recogniser.



(b) Input Inspector.

1. Device inputs.
2. Minimum and maximum input controls for limiting range of input.
3. Feedback, normal/rate of change & threshold/continuous options.
4. Output value controls.



Output Port Glover
Message Type Control Change
Channel 1
CC Number 60

(c) Output Inspector.

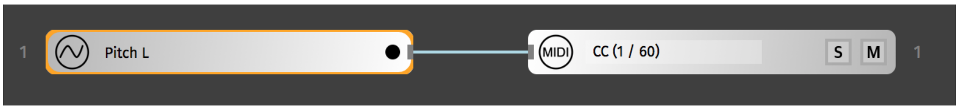
1. Output types available.
2. Current outputs selected.
3. Selected output controls.

Figure 5.10.: The Inspector Panel.

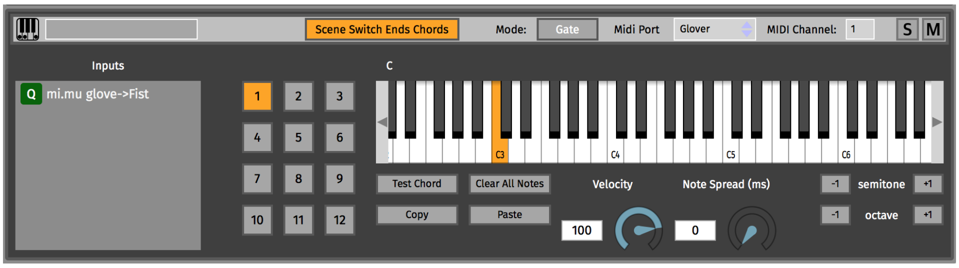
plies musical instruments: the layout of piano keys affords voicing chords in triads, with minor and major third intervals between each note, while the fretboard of a guitar affords greater use of fourth and fifth intervals. This phenomenon is compounded in digital instruments where the end-user designs and implements mappings, as potentially trivial design decisions made by the mapping software's designers could have a knock-on effect to the mappings created by end-users. As such, it is important to acknowledge that software used to create action–sound mappings will influence the mappings choices of DMI musicians.

Established literature on DMI mapping highlights the four mapping styles of one-to-one, one-to-many, many-to-one and many-to-many (Rovan *et al.*, 1997; Arfib *et al.*, 2002; Hunt, Wanderley and Paradis, 2003). To provide context and clarity around the mapping choices of Glove musicians in the empirical work of this dissertation, how those relationships are afforded in Glover is described in this section.

One aspect of mapping that must be determined is whether the possible combinations of Glover's gestural inputs constitute a *many* or *one* relationship. While an input object may contain multiple device inputs, the yaw parameter (a continuous movement) can be combined with fist posture (a qualifier). This could be considered a higher-level abstraction of a gesture feature, as the two device inputs represent two different parts of the user's body (the hand and the arm) that are acting together in one holistic movement. However, an equally valid input combination is the yaw movement of the right arm and the fist posture of the left hand. For such a case to be considered a holistic movement, it is possible to continue this abstraction process to encompass the user's entire body, rendering every instrument mapping as "one-to-" mapping, the *one* input being the user's entire body. For this reason, it was decided that each "device input" would be counted as a



(a) A one-to-one mapping using the patch-cord mapping.



Mode: Gate
Midi Port Glover
MIDI Channel: 1
mi.mu glove->Fist
Velocity 100
Note Spread (ms) 0

(b) A one-to-one mapping using the Chord Machine.

Figure 5.11.: One-to-one mappings.

single gesture feature, with their combination within one mapping option being a “many-to-” mapping.

### 5.4.1. One-to-one mapping

One-to-one mappings using the patch cord tools can be achieved by connecting an input object with a single device input to an output object with a single MIDI or OSC output, and are visualised in the traditional patch cord metaphor (Figure 5.11a).

One-to-one mappings are also possible using the Chord Machine mapping instrument, where a single note can be triggered using a single event or qualifier (Figure 5.11b).

### 5.4.2. One-to-many mapping

In the Input–Output objects, one-to-many mappings can be implemented by connecting one input object to multiple output objects. Alternatively, multiple output messages can be contained within an output object (Figure 5.6). Another form of one-to-many is to have two separate input objects with the same gesture feature

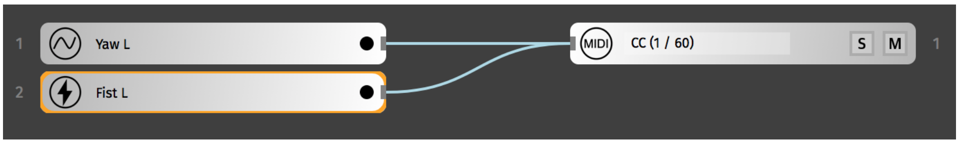


Figure 5.12.: An invalid many-to-one mapping.

input. With both the Chord Machine and the Note Matrix, one-to-many mappings are implemented by adding multiple notes to their piano keyboard interface triggered by a single gesture feature.

### 5.4.3. Many-to-one mapping

Many-to-one mappings are possible using Glover's patch cord interface. However, they do not follow the traditional visualisation of many-to-one, which would have two independent mapping input objects connected to a single mapping output object, for example Figure 5.12. In this implementation, the output simply alternates between the two inputs. Instead, many-to-one mappings resemble one-to-one mappings, with one input object connected to one output object, but with multiple gesture features in one input object (Figure 5.13).

Combined gesture features have the logical relationship of an AND gate, and only certain combinations of gesture features are valid in Glover: one movement and any number of qualifiers; one event and any number of qualifiers; or a combination of any number of qualifiers. As well as this, only one qualifier from each category, Directions, Postures and Button, can be used per hand at a time. This is due to the mutually exclusive nature of qualifiers, the hand can only be in one posture and one direction at any given time, and the button can only be up or down. Two postures from each hand can, however, be combined, for example the fist of the left hand and the open hand of the right.

Figure 5.13.: A valid many-to-one mapping.

### 5.4.4. Many-to-many mapping

Many-to-many mappings in the Input–Output interface are those that have an input object with more than one gesture feature connected to either an output object with multiple output messages or multiple output objects. Many-to-many mappings in the Instruments are those that have more than one gestural feature input and more than one note output.

## 5.5. Summary

This chapter has presented the DMI and end-user mapping software application, the Mi.Mu Gloves and Glover, which are used in the empirical research presented in the empirical work of this dissertation (Chapters 6–9).

Section 5.1 discussed the Mi.Mu Gloves, including the sensor layout and technical details of the gloves, as well as the adoption of the gloves beyond their initial practitioner. Section 5.2 discussed the Glover mapping software, how glove sensor data is presented to users as gestural features, and how Machine Learning (ML) techniques provide glove users with posture recognition tools. Section 5.3 discussed the how mapping is achieved in Glover, presenting the interfaces used to afford mapping design to users. Finally, Section 5.4 discussed how these interfaces allow users to create one-to-one, one-to-many, many-to-one and many-to-many mapping relationships.

The next chapter (Chapter 6) presents the beginnings of the empirical work, with a study into the mapping practice of the existing community of glove practitioners.

# 6. Mapping Practice of Existing Glove Users

The previous chapter (Chapter 5) presented the Digital Musical Instrument (DMI), the Mi.Mu Gloves, and end-user mapping application, Glover, examined in the empirical research of this dissertation.

This chapter presents initial empirical research conducted to examine the mapping practice of existing Mi.Mu Glove users, who performed a mapping design task. The results of this work found a marked difference in the mapping design behaviour of glove musicians with little performance experience, who designed mappings that reflected established embodied metaphors of music and movement, and those with much performance experience, who focused on creating ergonomic mappings that minimised performance mistakes. These findings influenced the focus of the work presented in Chapters 7 and 8, which examine the mapping practice of each of these two groups respectively.

The work presented in this chapter has previously appeared in the following publication:

Brown, D., Nash, C. and Mitchell, T. (2018) Understanding User-Defined Mapping Design for Mid-Air Musical Performance. In: *Proceedings of the 5th International Conference on Movement and Computing (MOCO)*. Genoa, Italy, June 2018. ACM. 27:1–27:8.

## 6.1. Introduction

To examine how glove musicians make mapping decisions, a group of existing Mi.Mu Glove musicians took part in a mapping task. These musicians have become familiar with using the gloves for mid-air musical interaction, and, as experienced users, already have established practices for creating mappings for the gloves.

Following on from previous research in user-defined gestures (Wobbrock, Morris and Wilson, 2009), participants were not expected to make identical decisions when designing mapping strategies. However, it was hypothesised that performers would make similar decisions based the underlying musical conceptual metaphors that drive our understanding of musical concepts (see Section 3.3). Whether musicians make use of musical metaphors to influence their mapping, and whether common metaphors exist between them, was explored. How these performers include expressive parameters in their mapping design was also examined.

## 6.2. Method

The study involved a group of five mid-air glove musicians, all of whom have owned a pair of gloves for significant time. The group were asked to individually perform a mapping exercise before participating in a group discussion.

### 6.2.1. Mapping Exercise

The participants were given a piece of monophonic music and asked to develop mappings that would enable them to perform it. The mapping needed to incorporate control for multiple notes and three expressive parameters: vibrato (pitch-bend); dynamics (volume control); and timbre (a low pass filter). While the note order and timing were fixed, participants were given the freedom to incorporate each of the expressive parameters as they wished.

The participants were provided with the necessary output parameters in their Glover projects, which were mapped to the MINI 3OSC preset of the Simpler synthesiser in Ableton Live, a Digital Audio Workstation (DAW) familiar to all of the participants.

The participants were required to decide upon appropriate input controls for these output parameters. They were also given the freedom to edit the output parameters if they felt there were more appropriate options, to allow the participants to design the mappings as naturally as possible, according to their personal tastes, experience and expertise.

The participants were provided with a score, annotated with the corresponding MIDI note values, and an audio example recorded from a quantised sequencer playing a piano synthesiser. One hour was allocated to complete the task.

The piece of music was written in A minor, with three main sections (Figure 6.1), arranged in a Rondo structure (AABBAACCAA):

*A:* a slow melody with small intervals between notes;
*B:* a slow melody with large intervals between notes;
*C:* a fast melody with small intervals between notes.

### 6.2.2. Group Discussion

After the mapping exercise, the participants performed with their mapping strategy to the group and discussed their designs. The group was encouraged to contribute their own feelings towards each performer's mapping. The participants' discussion was analysed using thematic analysis (Braun and Clarke, 2006).

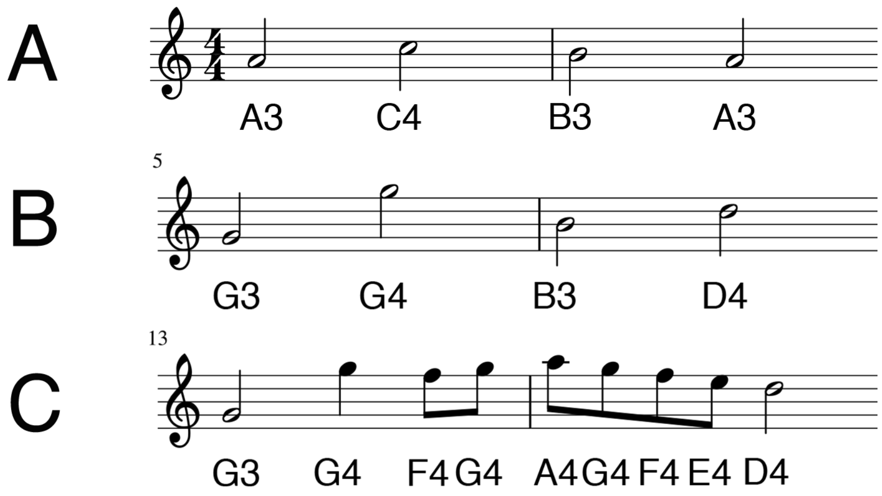



Figure 6.1.: The sections of the score.

## 6.3. Results

Five participants took part in the study. Each participant's mapping strategies are outlined in Table 6.1.

### 6.3.1. Participant Background

*A:* User for over two years. Has experience playing piano and guitar. Uses the gloves sometimes for composition and live performance. Regular use includes controlling synthesis, manipulating effects and controlling visuals. Typically uses the chord machine.

*B:* User for under three months. Has experience playing the violin, viola and keyboards. Uses the gloves for composition, and has never used them for live performance, but intends to. Regular use includes developing new composition strategies and "investigating new textures and musical objects". Typically uses OSC and MIDI CC messages to control inputs to self-developed Max/MSP and SuperCollider software.

*C:* User for between one and two years. Has experience playing the piano. Uses the gloves for live performance regularly and composes using them some of the time. Regular use includes solo musical performance, controlling visual effects and outboard synthesisers. Typically uses the chord machine, seven postures per hand, scene switching and button for calibration.

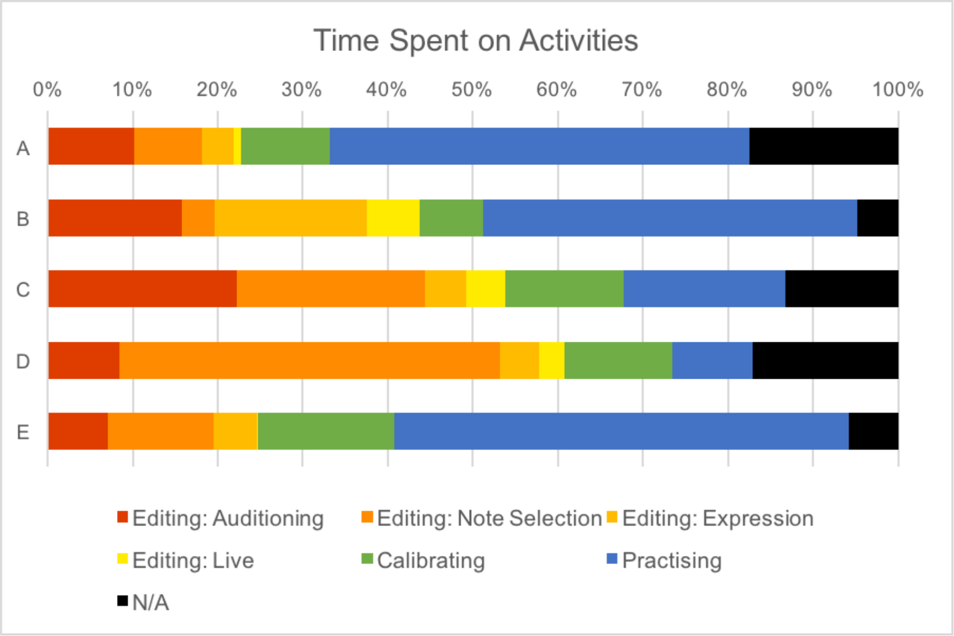
*D:* User for over two years. Has experience playing piano, oboe, bass, flute, violin, cello and synths. Uses the gloves for composition and performance often. Regular use includes standalone musical performance and use with other instruments and controllers. Typically uses "all" features.

*E:* User for over two years. Has no experience playing other instruments, and never uses them for live performance or composition. Regularly uses the gloves for development purposes.

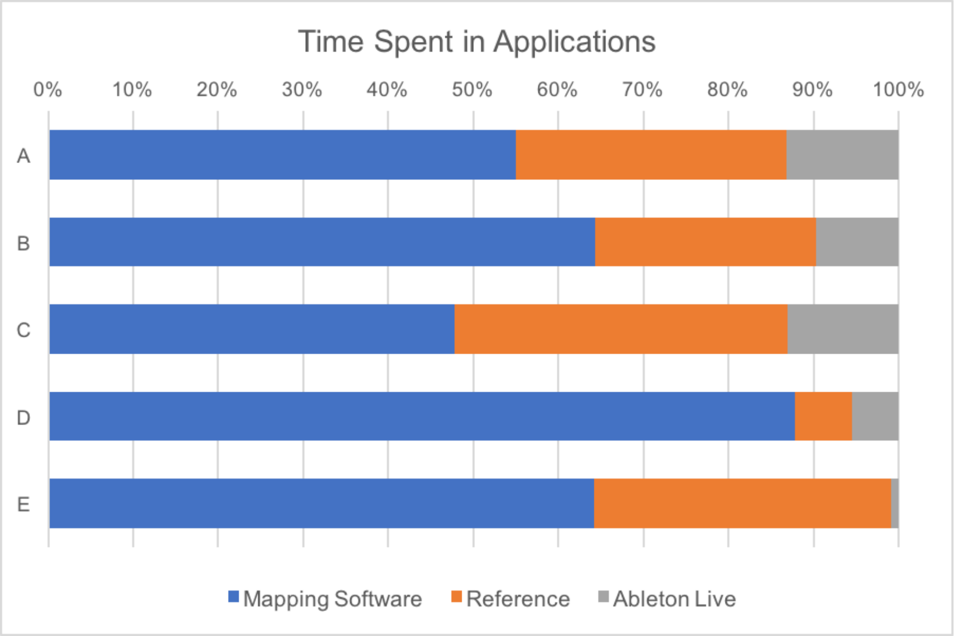
### 6.3.2. Exercise

During the exercise, each participant's interaction with the mapping software was recorded. These recordings were then analysed to find how each participant used their time during the task (Figure 6.2), measuring how long each participant spent in the mapping software, Ableton Live and on reference material (audio example and musical score), as well as the amount of time each participant spent: practising their performance; editing, either note selection mappings, expressive mappings, auditioning these mappings or editing settings in Ableton Live; or calibrating and setting up the gloves. Periods of inactivity or time spent on unrelated activities were marked as N/A (Not Applicable).

Participants A and E (who perform with the gloves the least) spent the majority of their time practising their performances (49% and 52% respectively). In contrast, the regular performers (C and D) spent less time practising (19% and 9% respectively) and the majority of their time editing (53% and 61%). Participant B, who



(a)



(b)

Figure 6.2.: Time spent during the exercise.

| *P* | *Note Selection* | *Expression* | *Comments* |
|---|---|---|---|
| *A* | All on RH.<br>Qualifiers: postures and direction (A and B sections).<br>Movement: Note matrix on pitch axis (C section). | All on LH.<br>Timbre: pitch axis.<br>Vibrato: roll axis.<br>Dynamics: no mapping. | Vertical spatial relation between direction qualifiers and notes.<br>Specific postures used for specific phrases. |
| *B* | All on LH.<br>Movement: Note matrix on pitch axis.<br>Movement: Octave intervals on yaw axis.<br>No qualifiers used. | Split between hands.<br>Timbre: LH average finger flex.<br>Vibrato: RH pitch axis.<br>Dynamics: no mapping. | Spatial grid-like representation of notes used. |
| *C* | All on LH<br>Qualifiers: postures and directions. | All on RH<br>Timbre: average finger flex.<br>Vibrato: roll axis.<br>Dynamics: pitch axis. | Circular motion for performing C section.<br>Focus on ergonomics and being visually appealing for performance. |
| *D* | All on LH.<br>Qualifiers: postures and directions.<br>Scene Switching: button to click through musical sequence (C section). | All on RH.<br>Timbre: average finger flex.<br>Vibrato: roll axis.<br>Dynamics: pitch axis. | Focus on ergonomics.<br>No relation between directions and notes. |
| *E* | Split between hands.<br>Movement: note matrix for pitch selection (mirrored on both hands).<br>Qualifiers: open hand posture for triggering. | All on RH.<br>Timbre: roll axis.<br>Vibrato: no mapping.<br>Dynamics: no mapping. | Vertical representation of notes.<br>Open hand "letting go" of the notes. |

Table 6.1.: Description of each participant's final mapping strategy.

mainly composes with the gloves, spent an approximately equal amount of time editing (43.4%) and performing (43.6%).

Each participant spent the majority of their time in the mapping application (Figure 6.2b). Interestingly, the the two most experienced performers spent both the least amount of time (C: 47.7%) and greatest amount of time (D: 84.8%) in the mapping application. These two participants also spent the greatest (C: 39%) and least (D: 6.5%) amount of time observing reference material. Four of the five par-

| *Expression* | *Control Parameter* | *Amount* |
|---|---|---|
| Vibrato | No Mapping | 1 |
| | Pitch Axis | 1 |
| | Roll Axis | 3 |
| Dynamics | No Mapping | 3 |
| | Pitch Axis | 2 |
| Timbre | Pitch Axis | 1 |
| | Roll Axis | 1 |
| | Average Finger Flex | 3 |

Table 6.2.: Expressive parameter mapping choices.

ticipants spent a notable amount of time in the Ableton Live application, which was spent either editing Ableton parameters or practising their performances.

### Note Selection

Participants A, C and D trialled multiple solutions for note selection mappings. Participant A swapped a Note Matrix for a Chord Machine for sections A and B of the piece, keeping a Note Matrix for section C.

Participant A spent most of their time practising their second note selection solution instead of continuing to make further edits. In contrast, participants C and D exhibited a continuous editing, auditioning, and practising cycle throughout the exercise. Participant C began by using a Note Matrix to perform section C of the piece, before switching to a chord machine; while participant D also began by using a note matrix for section C, but switched to a button and scene switch solution, clicking through the notes of the section.

### Expression

All but one participant spent <5% of their time editing the expressive mappings. Participant B spent 18% of their time on this task, and was the only participant to audition and trial multiple expressive mapping solutions. All of the other participants left the expressive mapping part of the task to the end of the hour, and did not experiment with more than one mapping solution for each expressive parameter. This suggests that the lack of expressive experimentation could be due to the time limitation of one hour, more time could have allowed participants to explore more solutions.

#### Ableton Live

Although an Ableton Live project with all the necessary MIDI and synthesiser settings was provided for the participants, all of the participants spent time editing the Ableton Live project. Participant A added extra effects processing to their project, while other participants edited synthesiser parameters such as the synthesiser's ADSR amplitude envelope.

#### Calibration

Each participant spent an average of 12% of their time conducting calibration tasks, revealing the lengthy setup time required to use the gloves. Also, all of the participants returned to recalibrate the gloves throughout the exercise. This included retraining the posture recogniser with fresh training data, setting minimum and maximum values for movement data, and refreshing WiFi connections to their gloves.

### 6.3.3. Mapping Designs

A description of each participant's final mapping design can be found in Figure 6.1.

#### Calibration

Four of the five participants set up the same series of mappings for resetting the gloves orientation, referred to as "set forwards". Each participant used either a

Figure 6.3.: An example of the "set forwards" mapping.

button or unusual posture (such as "pinky point") to trigger this orientation reset procedure, accompanying it with a pulse of the glove's vibration motor for feedback (Figure 6.3). This seems to have become a common procedure for glove performers. Participants noted that drifting from one's starting point while performing with the gloves happens regularly, and that this reset procedure has adopted by many members of the glove community.

### Bimanual Control

Of the five who provided complete mappings, three split note selection and expressive control between the two hands. This was done seemingly for cognitive purposes.

> "I don't have to worry about this [expressive] hand once I've got the muscle memory for this [note selection] hand."

Others cited a desire for *"independent control"* of notes and expression.

Of the other two, one (Participant E) mapped controls symmetrically, providing both hands the ability to select and trigger notes. However, this participant only mapped one expressive parameter, timbre, citing that mapping the others was too cognitively challenging.

> "I tried volume control, but I found it hard to control it consciously while controlling the notes."

The other (Participant B) mapped both note selection and expressive parameters to a single hand, claiming that it is *"quite interesting"* to map as much as you can on one hand. They also found difficulties controlling expression.

> "I tried mapping [average finger flex] to volume ... but it gets a bit uncontrollable."

They also commented that they were *"using the mouse for the rest of it,"* referring to the task of creating the mapping on their laptop.

### Mapping Expression

There was slight trend of participants making similar mapping choices for expressive parameters, although often no mapping was provided for expressive parameters. A breakdown is provided in Table 6.2.

When participants did create expressive mappings, a hand's average finger flex was the most popular control for timbral expression. One participant (C) commented that it "looks cool", suggesting that aesthetic considerations were an important factor. Another (B) commented that "there's a sort of symbiotic relationship between opening [their hand] equalling the filter", suggesting a metaphorical relationship, with the opening and closing of their fingers representing the opening and closing of the low pass filter.

Meanwhile, the roll axis was the most popular for vibrato control, with one participant commenting that "it seemed like a natural choice" and another that "it was sort of incidental".

### Metaphors

In an explanation of their mappings, it was found that four of the five participants expressed musical metaphors. During the discussion, the participants expressed a connection between spatial terms and musical pitch, supporting the idea that pitch has a strong schematic relation to space (Brower, 2000; Lerdahl, 1988).

> "It makes sense that C is physically the highest note."
>
> "Next octave is here [on the right], as I wanted some relation between where my hand is on the lateral plane and the [musical] pitch."
>
> "There is a kind of height relation for some of the notes."
>
> "Low notes low, high notes high."
>
> "You kind of want to play [the notes] like this [participant gestures up and down]."
>
> "Naturally it seems good to go up and down a scale."
>
> "The G obviously needs to come down [from the A]."

Also present were examples of dynamic controls, that reflected the UP–DOWN metaphor of dynamics (Wilkie, Holland and Mulholland, 2010).

> "I often do volume up and down."
>
> "I made it so it got louder and brighter as I raised my hand."

It is interesting to note that the participants who expressed the strongest representation of these metaphors in their mappings were the users who performed in front of audiences with the data gloves the least, instead using them for compositional or personal use.

### Ergonomics

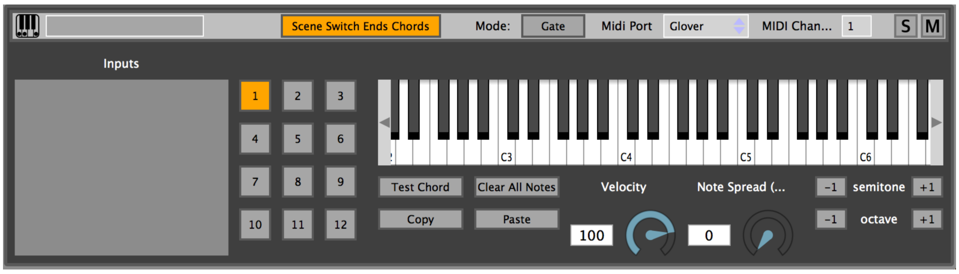
Although most of the participants used musical metaphors when discussing their mappings, some participants focussed on the playability of their note selection mappings.

> "It's annoying to do quick melodies so I used scene changing and just set each scene to [trigger each note in the sequence]. I used the easy, simple format of the song to program that sequence of notes as a sequence of scene changes."
>
> "I like how [other participants] thought about the musicality behind the notes, I didn't do that at all, I just put the postures where I saw them fit more for performance."
>
> "I did try with a Note Matrix, but you just don't have enough control."

These participants (C and D) focussed on the ergonomics of their control, focusing on how they would transition between note triggering postures, emphasising a need to be able to easily switch between them without accidentally triggering other notes.

> "I move from postures that use more unbent fingers to postures with bent fingers for reliable note triggering. For example, from one finger point to a fist. Then I change direction without changing posture so as not to accidentally trigger other notes."

This is reflected in the time spent by both participants auditioning and editing note selection mappings against practising their performances – if a particular musical section took too long to master with one mapping, these participants changed the mapping to something more playable, to the extreme case of exploiting unseen affordances of the mapping software to perform a musical section with the touch of a button.



Mode: Gate
Midi Port Glover
MIDI Chan... 1
Velocity 100
Note Spread (... 0

Figure 6.4.: The chord machine mapping tool.

## 6.4. Discussion

This study found evidence of contrasting mapping design behaviour between novice and experienced glove performers. Participants A and E (with little glove performing experience) first designed their mappings, spending little time experimenting with solutions, and then devoted the majority of their time to mastering the mapping solution they had defined. In contrast, the regular performers C and D took advantage of the dynamic nature of the gloves interaction, constantly updating their mapping solutions to aid their performance instead of dedicating time to extended practise. If a mapping is too difficult to master quickly, rather than dedicating time to practise, these participants changed their mapping solutions. These rapid feedback cycles of creating, auditioning and editing mappings reflects Progressive Evaluation in virtuosic music interaction, where expert users rapidly switch between these modes in creative tasks (Nash and Blackwell, 2014; Nash, 2011).

The participants who adhered most strongly to musical metaphors commented that they did so as it was “natural”, “made sense” and was “obvious”. This supports the notion that using metaphor leads to an intuitive interaction (Antle, Corness and Droumeva, 2009), however, by adhering to metaphor strongly and spending such little time editing their mappings, these participants perhaps failed to fully exploit the potential of the dynamic mapping possible with the gloves, requiring more time to be dedicated to practise.

In contrast, the participants with the most performing experience expressed very few or no metaphors at all in their mapping. These users instead focused on pragmatic, ergonomic solutions, and made fewer performance mistakes than the less experienced performers, who particularly struggled to perform the tasks C section (fast melody) accurately with their mappings.

The experienced performers' mapping practice highlights an interesting feature regarding user-defined control mappings in musical performance, in that, unlike acoustic instruments, in which the performer must master the movements necessary to play the instrument, the control mapping can be customised to suit the movements of the performer.

It was also found that many of the participants mapped note selection to one hand and expressive parameters to the other, reflecting the theory of bimanual action, where skilled manual tasks are divided asymmetrically between the two hands (Guiard, 1987), and has been observed in previous user-defined mid-air interaction research (Aigner *et al.*, 2012). The two participants who chose not to do this (B and E) had the most trouble controlling expressive parameters and note selection, suggesting that dividing different types of musical tasks between the hands could make for cognitively easier control of multiple musical parameters.

Many of the participants used the chord machine (Figure 6.4), designed to give glove users control over chords, to play single notes. This suggests that the affordances present in the chord machine tool provides a preferred mapping tool to glove musicians than the patch chord mapping tools, which reflect the traditional conceptualisation of mapping strategies. This observation presents an avenue for future research, looking at the effects of using different methods for presenting input and output mapping options to users.

For many expressive parameters the participants made no mapping choice at all, and from the exercise interaction data it was observed that four of the five participants spent very little time mapping expression. The lack of concentration on expressive parameters is perhaps down to the nature of the task: the participants were explicitly instructed to provide mappings for specific note selection (the target piece), while given freedom over their use of expressive parameters. The fact that participants were only given one hour to complete the task may have also had an effect. For instance, participant B spent a considerable amount of time auditioning and editing their expressive mappings, but they stressed that they “hadn’t had time to really explore” their choice to map vibrato to the pitch axis of their right hand. Given more time, other participants may have experimented further with different expressive solutions.

The amount of exercise time taken up by calibration tasks reveals that the gloves require a considerable amount of adjustment throughout a user’s interaction. Breaks in the participants’ time designing mappings would have disrupted their creative processes and and any periods of flow (Csikszentmihalyi, 1996). This suggests that calibration tasks should be removed from the users direct control to help them focus on mapping design, however, there is a balance to be achieved between abstracting such complexity away, which may benefit novice users, while still providing control over the precise workings of the gloves for expert users like participant D, who customises details such as their gloves’ IP addresses and UDP send/receive ports.

Although the focus of the task was to create performance mappings, it is interesting how many of the participants edited sound features of the Ableton Live synthesiser used in the exercise. This is likely down to each participant’s desire for the sound output to more closely match their own aesthetic taste: one participant commented “Ah, that’s better” after adding extra audio processing to the Ableton

Live synthesiser. Future work may give participants freedom over the sound output, allowing for a better reflection of their personal performances.

Interestingly during the group discussion, while the participants were encouraging towards each other, there was little discussion between the participants about the aesthetic choices each of them had made. This is likely due to the aesthetic distance the participant's felt towards the task, with one participant explicitly commenting that "this is not something I would ever perform in real life". This lack of intrinsic investment in the task likely had an impact on the mapping decisions made by the participants, reflected in the decision by experienced Glove musicians to focus almost exclusively on the ergonomics of their mappings.

## 6.5. Methodological Limitations

While this study has revealed some interesting insights into the mapping design processes of mid-air interaction musicians, there are important lessons learnt from this study to apply to future research. The length of time given (one hour) may not have provided some of the participants, particularly the more inexperienced users, with enough time to fully explore and experiment with possible mapping solutions. Additionally, participants commented that the target piece of the task did not reflect music they would choose to perform. This may have influenced how the participants approached the task, focusing on completing their mappings and giving a satisfactory performance instead of experimenting with mapping solutions.

Learning from this, the work in subsequent chapters (Chapters 7 and 8) examines mapping design in a non-instrumental manner, studying how users of the gloves design mappings in their own personal practice.

## 6.6. Summary

The work presented in this chapter found that glove musicians with more performance experience have a mapping design practice involving rapid iterations of auditioning and editing performance solutions, focusing on creating mappings that increase ergonomic playability and minimise error-proneness, while spending less time mastering the mappings they create.

In contrast, glove users with less performance experience design mappings that tend to adhere to established musical metaphors, particularly spatial relations between movement and sound. They also spend less time designing mappings, instead concentrating on practising performing with their mapping designs.

The next chapter (Chapter 7), examines the mapping practice of novice glove users, where the initial development of mapping design practice is examined over a month-long longitudinal study, while Chapter 8 looks at the mapping practice of expert glove musicians.

# 7. Development of Mapping Practice in Novice Glove Musicians

The previous chapter (Chapter 6) began the empirical research into the end-user mapping design practice of Mi.Mu Glove musicians, which is explored further in this and the following chapter (Chapters 7 and 8). Chapter 6 found a marked difference between *novice* glove musicians with very little or no performance experience, who designed "intuitive" mappings that adhered to embodied metaphors, and *expert* glove musicians with performance experience, who designed mappings that focused on ergonomics and error-minimisation. This chapter and the next will explore these two distinct groups in more depth, with this chapter focusing on novice glove musicians, where the initial development of glove musicianship and mapping design practice is examined.

## 7.1. Introduction

The work presented in this chapter addresses the research question: what factors influence and affect mapping design choices in the initial development of glove musicianship? To address this, a longitudinal study was undertaken, aiming to track the development of mapping practice over a meaningful time period. This was to allow for the research to investigate the development of serious mapping practice within the context of preparing for a musical performance, as opposed to only examining what a new glove user would design in an afternoon of play and experimentation.

Previous work into a musician's creative practice with new musical interfaces has called for moving beyond examining initial impressions and to instead examine how these instruments become integrated into a musician's practice, through the use of longitudinal studies (Johnston, 2011; Nash, 2011). This method allows researchers to investigate creative practice "in the wild", which provides observations grounded in the context of an individual's own experience (Kaye, 2009), revealing, in musical cases, how a musician's instruments and tools are used in the real world to facilitate creativity.

An important tool is indirect usage logging, which has been advocated in the study of human-computer interactions as part of longitudinal, observational research methods (Kaye, 2009), and in creativity focused research in particular (Nash, 2011; Shneiderman, 2007), where usage logging can minimise the psychological impacts of direct observation research techniques, allowing the research aspect of creative interactions to become invisible. For this study, usage logging (described in Section 7.2.2) has been incorporated into the Glover application.

While large numbers of participants for such studies provide for more statistically significant results (Nash, 2011), finding or recruiting a significant user base for Digital Musical Instruments (DMIs) can be difficult (McPherson and Kim, 2012), particularly as the interfaces are often experimental prototypes made by hand, or require expensive, specialist third party commercial hardware. However, meaningful insights can be gained through the detailed investigation of the creative practice of a small number of participants (Gelineck and Serafin, 2012; Collins, 2007; Collins, 2005).

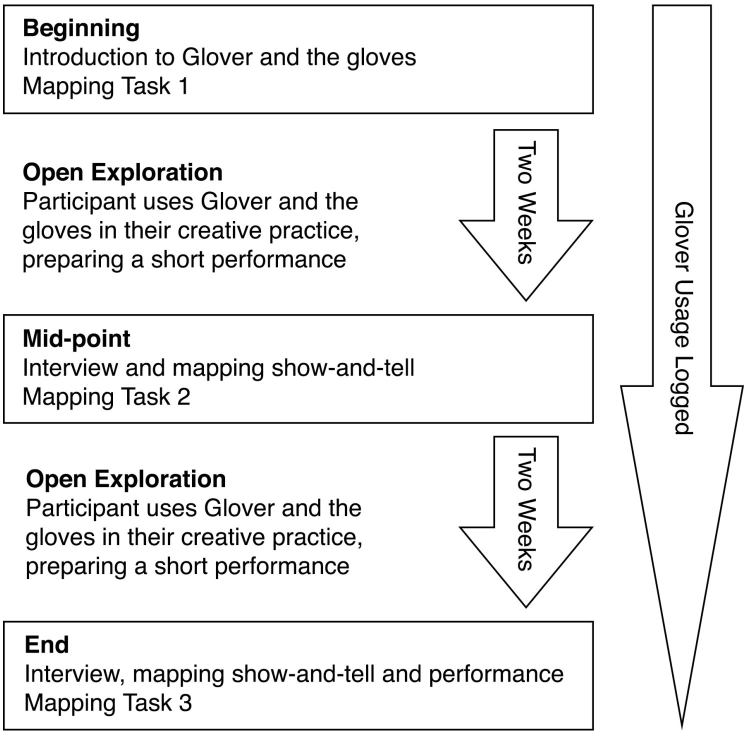


Figure 7.1.: The longitudinal method used to examine novice mapping practice.

## 7.2. Method

The longitudinal method is summarised in Figure 7.1. Participants were contacted and recruited through applications to the Mi.Mu Glove residency scheme, which lends Mi.Mu Gloves to artists for a short time period. While this provided a self-selecting sample, these participants had initial motivation to use the Gloves in their musical practice, and were thus an appropriate sample for the research. Each participant was given a pair of Mi.Mu Gloves and a copy of Glover (see Chapter 5) for one month.

Each participant was also provided with a technical instruction guide (see Appendix A), which assured that each participant received an equal level of technical support.

### 7.2.1. Open Exploration

The participants had the freedom to use the gloves as they wished, with the following stipulation: they must prepare a short musical performance (5 mins), to be performed at the end of the study period. Before beginning the study, each participant took a profiling survey to determine their previous musical expertise. During the study, the participants met with a researcher at regular intervals: the beginning, mid-point (two weeks in) and end of the study period; to take part in semi-structured interviews about their experience with the Gloves, and to discuss their mapping ideas and practice. The themes found from the Grounded Theory analysis of experienced glove musicians (Chapter 8) informed the interview questions. The interviews were analysed using Thematic Analysis (Braun and Clarke, 2006). During the month period, usage logging, described in Section 7.2.2, was used to record the participants' activity in the Glover software.

### 7.2.2. Usage Logging

To collect data on how the participants were using Glover during the open exploration element of this study, usage logging was incorporated into the Glover software.

As much interaction data as practically possible was collected to allow for flexibility in the analysis. A detailed account of a participant's use of Glover is logged, with each modification made to the glover file recorded. This is done by listening for any changes to the Extensible Mark-up Language (XML) structure that Glover uses to store the participant's modifications, recording every change made to the mapping design. Two "Save As" files are also logged, one for the beginning of the session and one at the end, allowing for the contextualisation of the changes made throughout a session.

| *Category* | *What is Recorded* | *Details Logged* |
|---|---|---|
| Glover | Save As Files | Beginning and end of session |
| | Device Settings | Added/Removed, Calibration |
| | Arrangements/Scenes | Added/Removed |
| | Mapping Objects | Added/Removed, Input and Output parameters |
| | Connections | Added/Removed, Inputs and Outputs connected |
| Activity | Mouse Active | On/Off |
| | Mouse Clicks | Coordinates |
| | Device Active | On/Off |
| | MIDI Output Active | On/Off |
| | Inspector | Component displayed in inspector |
| | Process | Application in focus (e.g. Glover or DAW) |

Table 7.1.: Breakdown of Usage Logging.

Keyboard and mouse activity and the user's currently viewed application is also logged, allowing for the observation of the levels of activity, as well as the use of related music applications alongside Glover, such as a Digital Audio Workstation (DAW). The streams of gestural input data and MIDI output data are not logged, as this would create too much information to be uploaded, but the logging notes when input devices or MIDI outputs are active. A breakdown of the data collected by the usage logging is provided in Table 7.1.

The logs are saved to file, compressed, and encrypted, before being sent over the internet for analysis. Previous work has discussed the difficulties in using standard data transfer protocols in remote logging, such as FTP and POP/SMTP email, as they are usually blocked by client-side security features (Nash, 2011). Using a similar solution to Nash, our logging files are sent as binary data in HTTP POST commands to a dedicated Node.JS web application, which then uses SMTP email to send the files to the researchers (Figure 7.2).

As the Mi.Mu Gloves require a WiFi connection, a user has no connection to the internet while using the Gloves. This means that the upload process could not be incorporated into Glover itself and simply triggered in the application's shut down procedure, as a WiFi connection to the internet may not have been established.

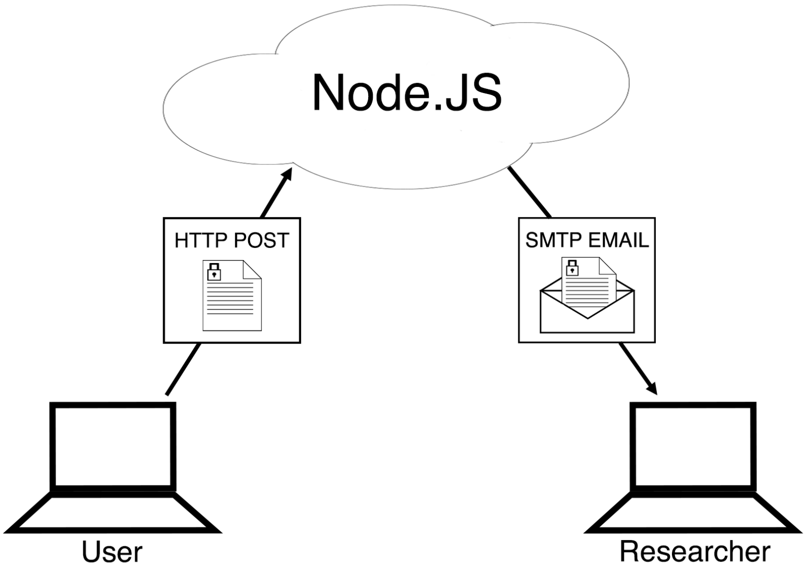
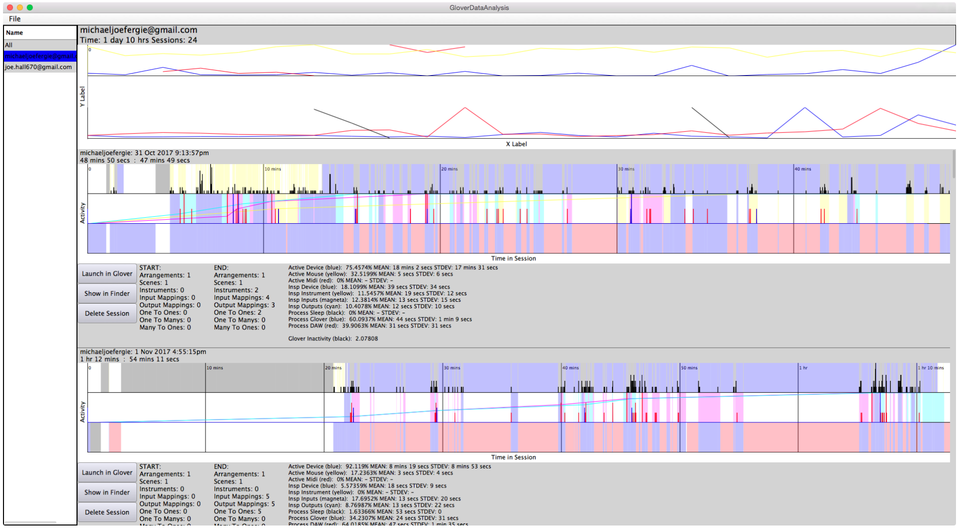


Figure 7.2.: The data upload process.

Holding up Glover's shut down procedure until a participant had reconnected to the internet would be inappropriate, as this would be noticeable to users and interrupt the experience of the software. Instead, a separate uploader application is launched as Glover shuts down, which runs as a background process until the user has reconnected to the internet and a successful connection can be made to the web application. If no connection is made, the logging files remain on the user's computer and are added to the next upload attempt.

### Data Analysis Software

To support the analysis of the logging data, a data analysis application was developed. The application provides overviews and visualisations of a participant's use of Glover, organised into each session (Figure 7.3). The application was designed to provide the logging data according to Shneiderman (1996)'s Visual Information Seeking Mantra: Overview first, zoom and filter, then details-on-demand; to facilitate effective analysis of the data.



michaeljoefergie@gmail.com
Time: 1 day 10 hrs Sessions: 24

michaeljoefergie: 31 Oct 2017 9:13:57pm
48 mins 50 secs : 47 mins 49 secs

START:
Arrangements: 1
Scenes: 1
Instruments: 0
Input Mappings: 0
Output Mappings: 0
One To Ones: 0
One To Manys: 0
Many To Ones: 0

END:
Arrangements: 1
Scenes: 1
Instruments: 2
Input Mappings: 4
Output Mappings: 3
One To Ones: 2
One To Manys: 0
Many To Ones: 0

Active Device (blue): 75.4574% MEAN: 18 mins 2 secs STDEV: 17 mins 31 secs
Active Mouse (yellow): 32.5199% MEAN: 5 secs STDEV: 6 secs
Active Midi (red): 0% MEAN: – STDEV: –
Insp Device (blue): 18.1099% MEAN: 39 secs STDEV: 34 secs
Insp Instrument (yellow): 11.5457% MEAN: 19 secs STDEV: 12 secs
Insp Inputs (magneta): 12.3814% MEAN: 13 secs STDEV: 15 secs
Insp Outputs (cyan): 10.4078% MEAN: 12 secs STDEV: 10 secs
Process Sleep (black): 0% MEAN: – STDEV: –
Process Glover (blue): 60.0937% MEAN: 44 secs STDEV: 1 min 9 secs
Process DAW (red): 39.9063% MEAN: 31 secs STDEV: 31 secs

Glover Inactivity (black): 2.07808

michaeljoefergie: 1 Nov 2017 4:55:15pm
1 hr 12 mins : 54 mins 11 secs

START:
Arrangements: 1
Scenes: 1
Instruments: 0
Input Mappings: 0
Output Mappings: 0
One To Ones: 0
One To Manys: 0

END:
Arrangements: 1
Scenes: 1
Instruments: 0
Input Mappings: 5
Output Mappings: 5
One To Ones: 5
One To Manys: 0

Active Device (blue): 92.119% MEAN: 8 mins 19 secs STDEV: 8 mins 53 secs
Active Mouse (yellow): 17.2363% MEAN: 3 secs STDEV: 4 secs
Active Midi (red): 0% MEAN: – STDEV: –
Insp Device (blue): 5.57359% MEAN: 18 secs STDEV: 9 secs
Insp Instrument (yellow): 0% MEAN: – STDEV: –
Insp Inputs (magneta): 17.6952% MEAN: 13 secs STDEV: 20 secs
Insp Outputs (cyan): 8.76987% MEAN: 13 secs STDEV: 22 secs
Process Sleep (black): 1.63366% MEAN: 53 secs STDEV: 0
Process Glover (blue): 34.2307% MEAN: 24 secs STDEV: 31 secs

Figure 7.3.: The Logging Analysis Software.

The software organises interaction data by participant, providing an overview on all of a participant's interactions with Glover, as well as information on individual sessions.

A session is defined from when a Glover file is opened to when it is closed. This means that if a user is using glover and a new file is opened, the previous file, which is automatically closed, is recorded as a session, and a new session is started for the new file.

Collecting and being able to analyse each participant's logging data during the study period allowed for any observations taken from the logs to inform the interviews with the participants, and gave a chance for these observations to be verified by the participants.

### Filtering Interaction Data

Studying the interaction data of early participants, it became clear that some data filtering needed to take place. First, the inspector data, which details which component is visible and being edited in the inspector window, was only relevant when the participant's focus was on Glover. Similarly, users spent long periods of time with Glover open but without any user interaction data (glove, mouse, and midi output) being active, often for periods of several hours. This too was skewing analysis data, as users were leaving Glover open while their computer was asleep, or simply leaving it open during long breaks. It was decided that data about time spent in different applications would be filtered by periods of definite activity, when either the glove input device, MIDI output or mouse and keyboard data were active, with a one minute window.

### 7.2.3. Task Analysis

While the participants were free to use the gloves as they wished during the study period, at regular intervals (beginning, mid-point and end of the month), participants also took part in set mapping tasks, where the participant was asked to design a mapping for a musical task. Learning from the mapping task of Chapter 6, the mapping of discrete and continuous parameters was separated into two tasks: a note selection task, designed to examine how the participants designed glove mappings to play discrete notes; and an expressive task, designed to examine how participants designed mappings for continuous expressive parameters. In both tasks, the participants were given 45 minutes to complete the task. Afterwards, the participants took part in brief interviews to discuss their mapping solutions.

#### Note Selection

For the note selection exercises, the participants were tasked with designing a mapping to perform a short musical sequence. The three musical sequences (Figure 7.4) were adapted from three pedagogical études for violin (Holstein, n.d.): Kreutzer's Étude No. 10 (Opus 42, No. 2), Wolfhardt's Étude No. 11a (Opus 38, No. 103), and Sitt's Étude No. 2 (Opus 32, No. 18). Learning from Chapter 6, the

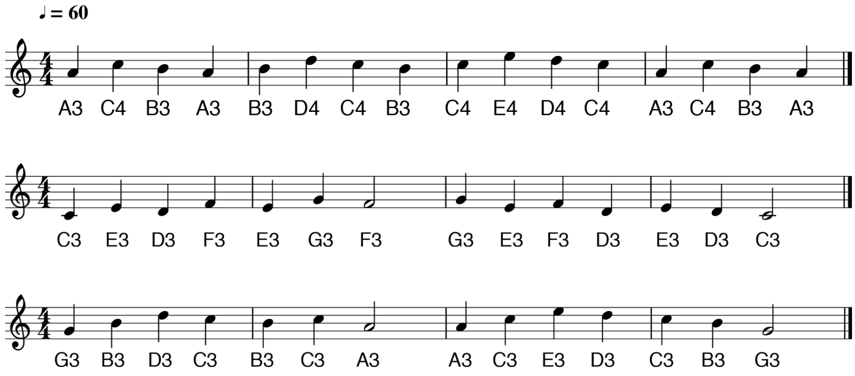


Figure 7.4.: The musical material for the three note selection exercises.

musical extracts were kept short to allow the participants to complete the exercise within the given time. The note-related MIDI output from Glover was mapped to a Grand Piano sampler instrument in Ableton Live 9.

### Expression

For the expression exercises, the participants were tasked with designing mappings for five expressive parameters, based on the five facets of musical expression described by Juslin (2003): Tempo, Articulation, Timbre, Pitch and Dynamics; for three one-minute musical extracts taken from established Theremin repertoire: Tárrega's *"Recuerdos de la Alhambra"*, Saint-Seäns' *"The Swan"* and *The Romance* from Shostakovich's *"The Gad Fly Suite"*. Theremin repertoire was chosen due to its established suitability for mid-air performance.

The musical extracts were synthesised in Ableton Live 9, with the accompaniment parts synthesised using the Wurli2 Soft Piano instrument, and the lead part (whose expressive parameters were controlled by the participant) synthesised using the Operator instrument. Control Change MIDI output from Glover was mapped to the following parameters.

- Tempo: Tempo for the Live set.

- Articulation: Attack and Release parameters of the Lead Operator instrument.
- Dynamics: Volume of the Lead Operator instrument.
- Pitch: Amount of vibrato on the Lead Operator instrument.
- Timbre: Filter Frequency of an Auto Filter applied to the Lead Operator instrument.

## 7.3. Participants

Eleven musicians ($F = 4, M = 7, \mu = 26.4, \sigma = 4.8$) took part in the study. Results from the profiling survey are presented in Figure 7.5. The participants tended to attend a moderate amount of musical performances, and to compose and perform music a moderate amount of time. Seven of the participants had experience playing piano or using DJ or similar music controllers, while five participants had experience playing guitar. All of the participants rated themselves as at least somewhat technically proficient.

All of the participants had at least some experience performing music live in front of audiences, of which the musicians provided the following descriptions.

> "Usually perform with a hybrid electronic/acoustic set, using guitars, keyboards and sequencers with effects and running these and my vocal mic through a sampler to further manipulate the sound."

> "I perform music for live performance arts, such as circus and contemporary dance. I sometimes perform with other musicians, but more often on my own, using live looping or pre-recorded loops and phrases to play along with during a show."

> "I play in the band called [REDACTED]. Pop, Rock, Electronic, J-Pop. Guitar, bass, drums and vocals. Backing tracks in Ableton Live. On av-

> erage we perform once month. I sometimes preform solo in the open mic, sometimes DJing on my iPad."
>
> "I used to perform quite often as a session guitarist a few years ago, in different bands with a variety of genres such as progressive metal, alternative, rock, fusion etc."
>
> "lipsyncing to slowed down music, combined with visual / physical actions … sometimes using technology e.g. MIDI controllers to modify the sound whilst performing. Recording and live looping … often solo, sometimes as part of a collaboration, sometimes I play guitar in a drag band. I perform at least once a month, sometimes once a week."
>
> "I use a system to perform live shows which relies on a desk, mic and use the Ableton Push to layout original tracks."

## 7.4. Observations: Open Exploration

Each novice spent varying degrees of time using the Glover software during the month-long study period. Total times spent ranged from 3h49m over 9 individual sessions to 22h41m over 44 individual sessions, with the mean time spent being 11h11m (SD = 6h22m). Due to this, the usage logging of the first 12 hours of each novices practice was used in the analysis, and novices with under five hours of total usage were omitted (N = 10). While this time-based data could have been normalised, it was decided that maintaining the data's real time information would provide for a more accurate representation of the time required to learn and develop creative practice with the Gloves system.

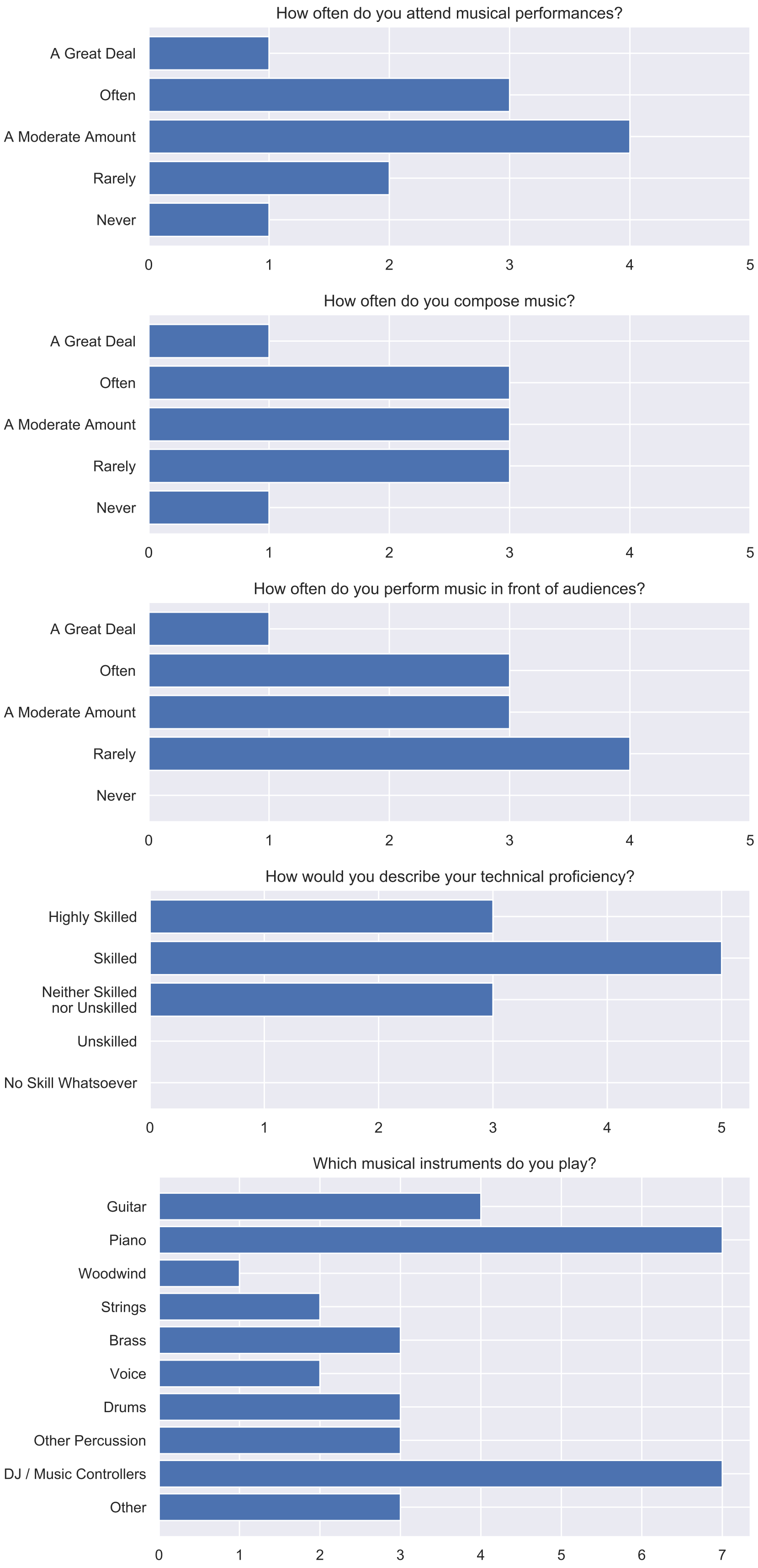


Figure 7.5.: Participant profiling survey results.

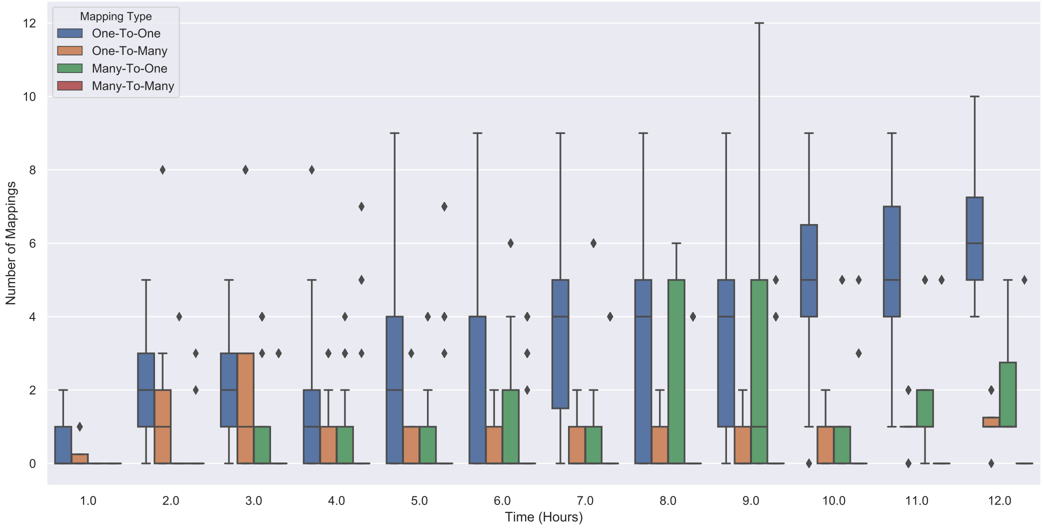


Figure 7.6.: Mapping relationships used in novice mapping designs.

## 7.4.1. Mapping Design

### Mapping Relationships

The participants' use of mapping relationships in Glover (described in Section 5.4) was observed through the usage logs, which revealed that the mapping relationships most used by the novice glove musicians were one-to-one mappings, with their use steadily increasing as the participant's spent more time using Glover (Figure 7.6). Many-to-one and one-to-many mappings were occasionally used, but their usage remained low throughout, with the median values for each remaining below two for the duration of the study. Meanwhile, many-to-many mappings saw almost no use. This suggests that the novices found the most simple mapping relationship, one-to-one, to be the most useful in their mapping designs.

### Mapping Interfaces

The use of mapping interfaces (see Section 5.3) was also observed. Instrument interfaces in mapping design remained fairly constant throught the study period, while the amount of input-output objects used increased throughout the time

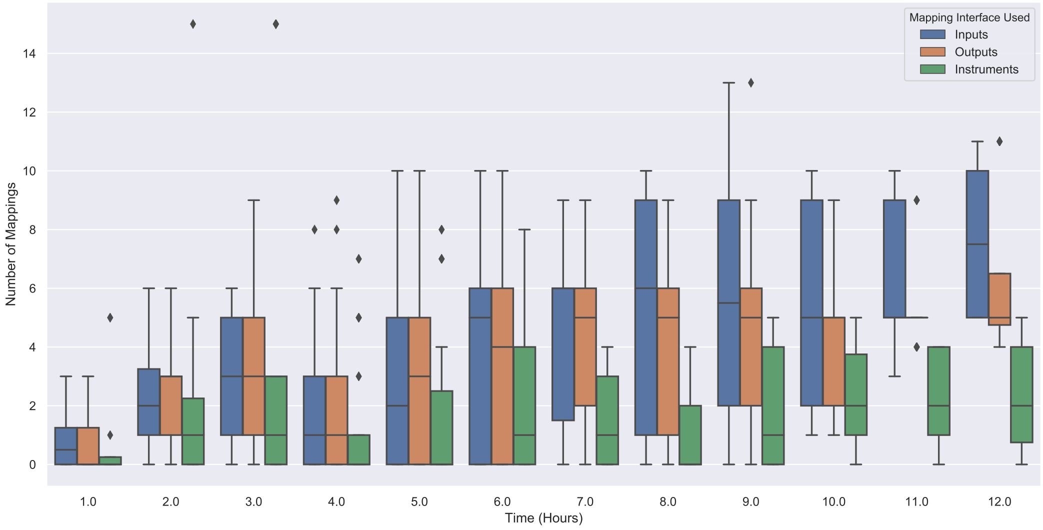


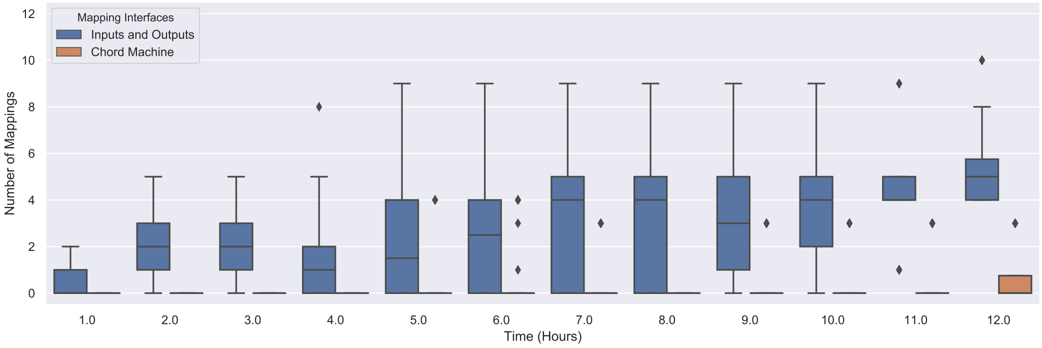
Figure 7.7.: Use of mapping interfaces in novice mapping designs.

spent with Glover (Figure 7.7). Combining this with the observations of mapping relationships (Figure 7.6), suggests that the novices tended towards implementing simple, one-to-one mappings using the Input–Output object interfaces. Figures 7.8 and 7.9 show that the Input–Output objects were used almost exclusively for both one-to-one and many-to-one mappings, while Glover's Instrument interfaces were seldom used.

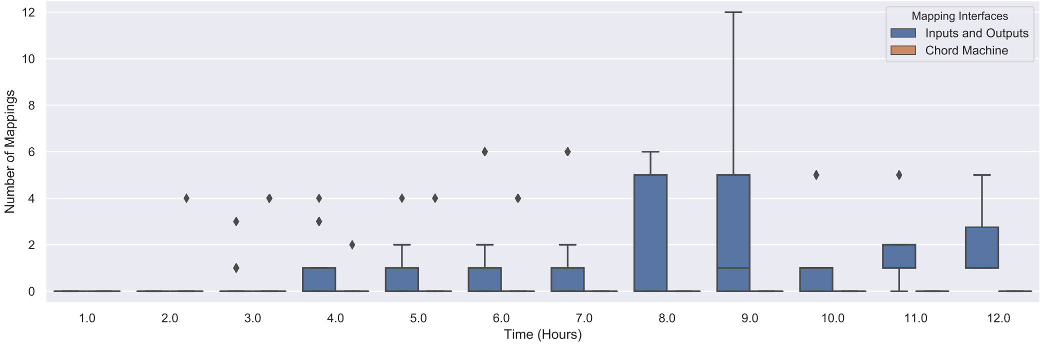
### Gesture Features

The participants' use of the types of gesture features available to be mapped in Glover (see Section 5.2.1 for more detail): *Events*, one-shot triggers; *Qualifiers*, state-based controls including Posture Recognition, and *Movements*, continuous signals; was also examined.

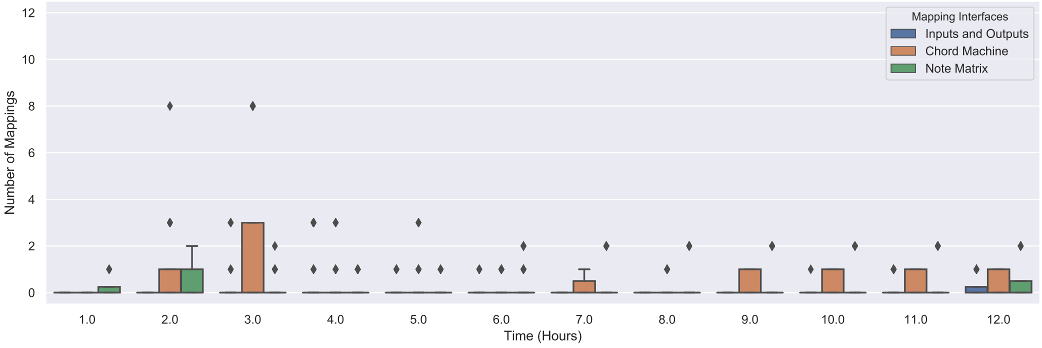
Movements and Qualifiers, specifically Postures, were found to be the most popular gestural parameter in mapping designs, with the participants' use of both steadily increasing over time (Figure 7.10). Events were rarely used, and were only used in meaningful amounts after 9 hours of Glover use, while Postures made up the majority of qualifiers used.



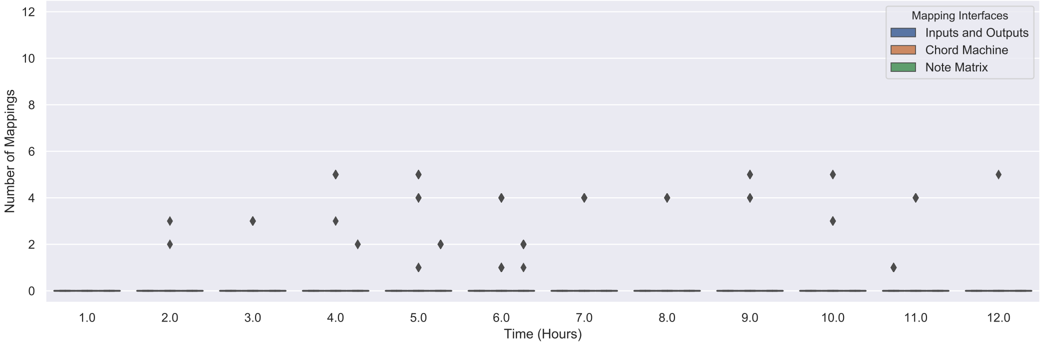
(a) One-to-one mappings.



(b) Many-to-one mappings.

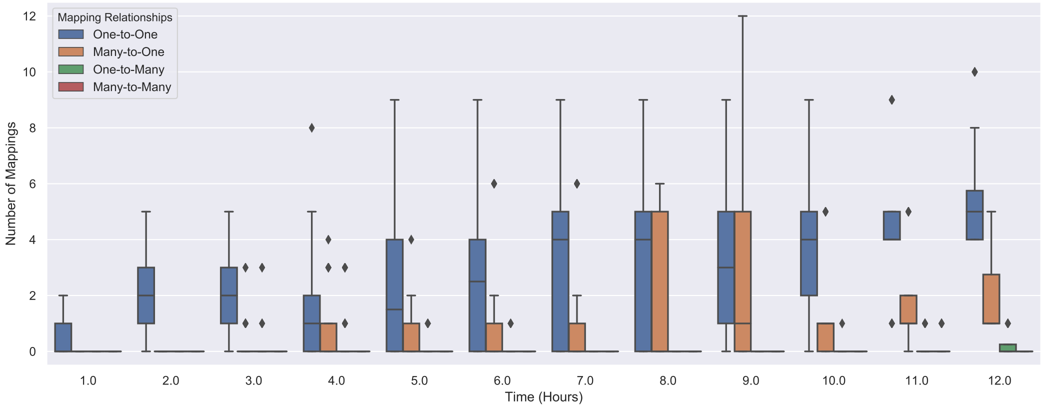


(c) One-to-many mappings.

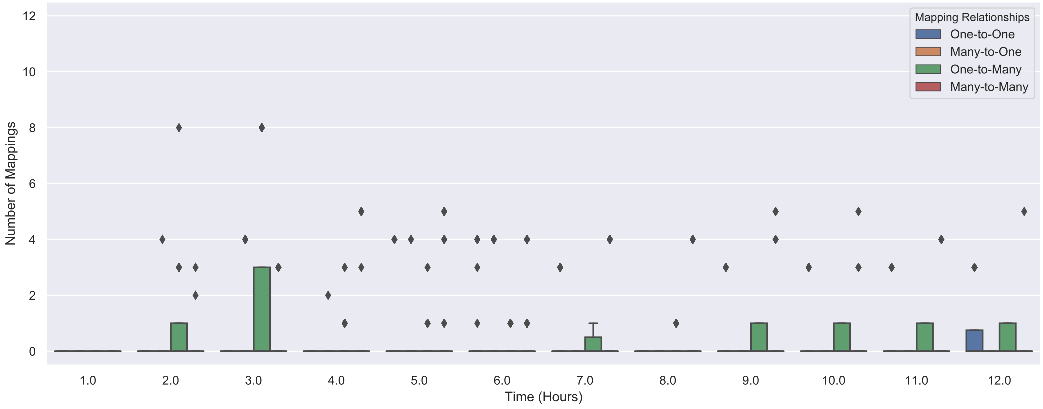


(d) Many-to-many mappings.

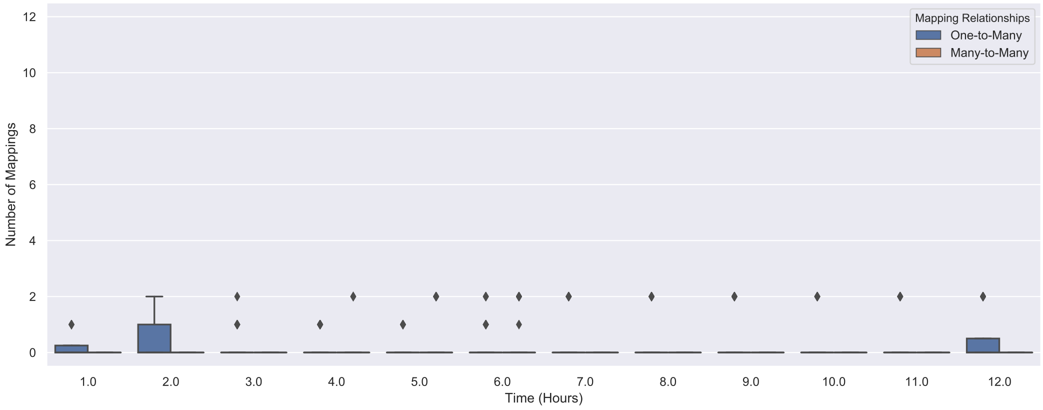
Figure 7.8.: Use of mapping interfaces by mapping strategy categories.



(a) Input–Output Objects.



(b) Chord Machine.



(c) Note Matrix.

Figure 7.9.: Use of mapping strategy categories by mapping interface.

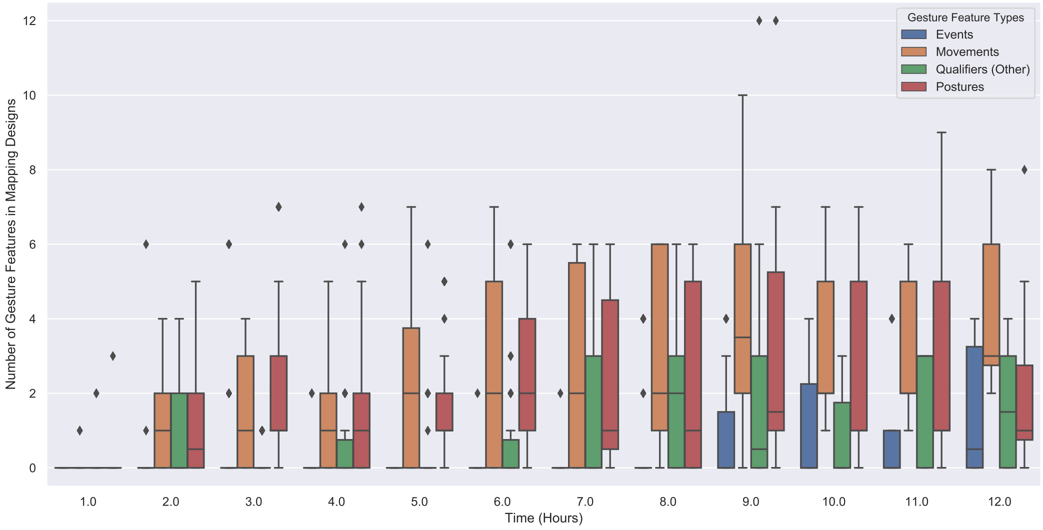


Figure 7.10.: Gesture feature types used in mapping strategies.

### MIDI Output

The participants use of MIDI outputs (Figure 7.11) steadily increased over time, with discrete, note-based outputs: Note On, Note Off, etc., being more commonly used than continuous, expressive outputs: Control Change or Pitch Bend.

## 7.4.2. Activity

At the first few hours of glove practice, the novices spent the majority of their time in Glover over a DAW (Figure 7.12). Time allocated to DAWs grew over time, with participants beginning to spend roughly 40% of their session in their DAW after the eight hour mark.

The use of Glover's Inspector Panel during a session decreased over time, from 80% of session time to 30-50% after eight hours of use. The device window was in focus for the majority of the time spent in the inspector, particularly during the first four hours of use. After nine hours, the input inspector saw the most use.

The time the gloves were actively used during Glover sessions fluctuated between 40-70%, growing to 50-80% after eight hours of use (Figure 7.14). MIDI

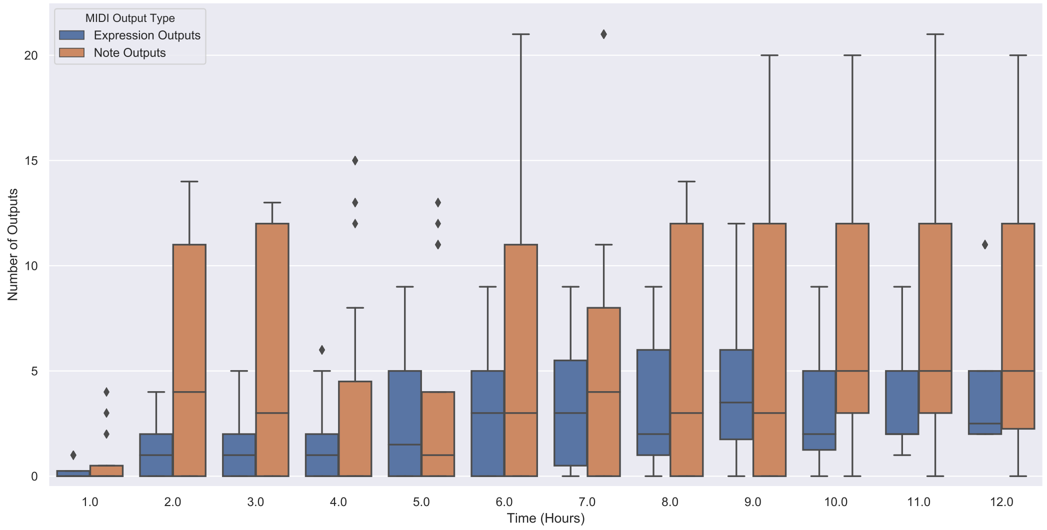


Figure 7.11.: MIDI output types used in mapping strategies.

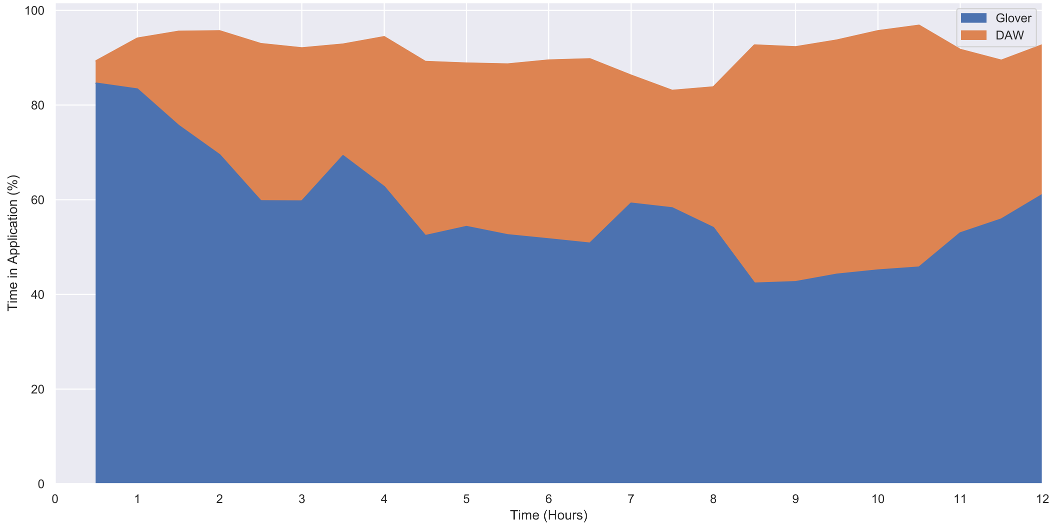


Figure 7.12.: Percentage of time spent in relevant applications during a Glover session.

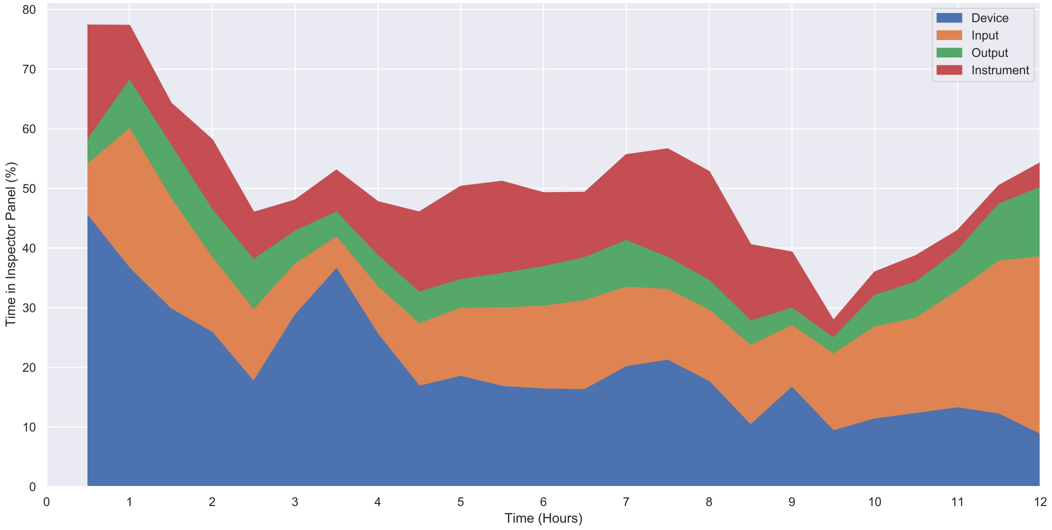


Figure 7.13.: Percentage of time spent in the Inspector Panel while using Glover.

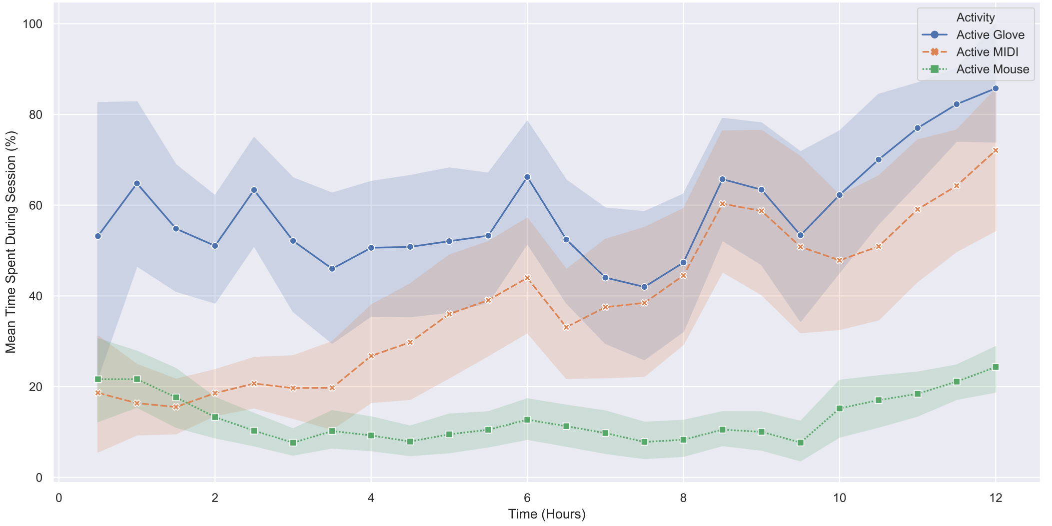


Figure 7.14.: Rates of activity during Glover sessions.

output activity steadily increased over time, matching the rate of glove activity after the seven hour mark.

### 7.4.3. Thematic Analysis

Presented below are the major themes found from the Thematic Analysis of the interviews with the novice glove musicians.

#### Leveraging Existing Musical Practice

The novice musicians incorporated their existing instrumental expertise into their glove practice, and would design glove mappings that were influenced by the instruments they had experience playing, often designing "air" versions of established instruments such as pianos or violas.

> "I mentioned I'm trying to play it like a piano... The chords on the left hand just because in my piano playing career that's all I can do [with the left hand]"
>
> "we just started off by saying 'lets make a viola"'.
>
> "So the general idea I have is I have taken some plug-ins I have to make a MIDI guitar sampler ... I make an air guitar essentially just to test, I play notes and chords with the gloves."

#### Augmenting Existing Practice

The novices would use also their existing instrumental practice by augmenting an instrument with the gloves. This was commonly done by guitar players, where the gloves would be used as a substitute for an expression pedal, to control expressive effects such as delays and pitch shifts.

> "But my intention would be while I do this [glove control] with my right hand I can tap on the guitar."
>
> "I was thinking why can't I actually explore incorporating the guitar and the gloves as well."
>
> "I wanted to play guitar, and control effects through the gloves."

### Accidental Triggering and Problems with Postures

The novices, similar to the experienced practitioners (discussed in Chapter 8), had issues with accidental triggering. This mostly occurred with mappings that involved the posture recogniser, which caused frustration and disruptions. The novices had difficulties training postures that were reliable, and had recognition difficulties when moving between postures that were very similar, for example where only the position of a single digit was different.

> "I kept turning it on and off by mistake, and other things were happening like I was turning the reverb on and not noticing that the drums sound terrible."
>
> "I just based it on the fist and open hand, because they were the ones I knew I could do without accidentally triggering something else."
>
> "[The problems are] mainly the postures interacting with each other, you can accidentally hit a posture on your way to another one."
>
> "And there should be another one done with [his thumb tucked underneath palm] but I'm finding it difficult to switch between this posture and this one [flat palm and tucked thumb]."
>
> "All the postures just mess up and it just ends up becoming a mess, having to reconfigure the postures, everything changes, it just becomes a mess."

> “One more thing that I’ve found a bit tricky to do is sometimes when I program the gloves they don’t necessarily program the way I do the posture.”

The novices also found it difficult to perceive whether the errors were due to them or were attributed to the posture recogniser.

> “[the posture recogniser] is quite prone to errors but a lot of that might just be user error rather than anything else.”

The novices would either start to avoid using postures in their glove practice in favour of the other discrete controls available, such as the direction parameters, or would change their postures to ones with greater difference between them, sometimes only making use of two postures: a closed fist and open hand.

> “I only found a couple of postures that really work for me when I’m stressed out, fist and palm [open hand].”
>
> “I just based it on the fist and open hand, because they were the ones I knew I could do without accidentally triggering something else.”

Some novices persevered with their troublesome postures due to aesthetic and creative reasons: using that particular posture was aesthetically interesting to them and they were determined to incorporate it into their performances.

> “I want to use that posture as I kind of like, I can’t exactly explain why, but I’ve been a bit fixated on this [posture] and this [posture].”

### Audience Perception

The visual relationship between actions and sounds was important, with the novices’ focus being that the audience was able to perceive that they were controlling music with the gloves. This led the novices to make use of larger movements through using the orientation parameters, and also lead them to use metaphors of existing instruments or established relational metaphors of actions and sounds, such as UP AND DOWN IS LOUDER AND SOFTER.

> "I wanted the controls the gloves had to be obvious to the audience, and I wanted that to be the strongest feature of the performance."
>
> "The important factor is that you interact with the audience and the audience can see that you are doing something. The fact that I can do this and contribute to the audio-visual relationship is important to me, because people come to an event to see a performance."
>
> "I'm trying to make something that is intelligible for an audience straight away."

> "My intention is to make something that is very noticeable, the important thing is that the audience can see that I'm controlling the music ... my first idea is that's going to be a noticeable thing that happens."
>
> "It requires quite a big movement, a movement over quite a long range? I was instantly drawn to using the orientation mappings."

> "As you play it in the shape that you might hold a viola it will look like, from the audiences perspective, that you're playing an air viola."

> "My second idea for this intro would be to program the pitch [glove parameter] to alter the dynamics of the samples, so if I were to go down here the sample would be very low, and if I go up here the sound will grow."

> "You've got the sound in your hand and when you turn it to face the floor, it gets quieter. That just seemed to make sense."

One novice discussed an idea for a gesture or movement that would not be fully realised by the gesture parameter, and used an embellished expressive movement to express their aesthetic intent. This novice discussed an embodied metaphor of peeling velcro, as well as a metaphor of invisible "sound objects".

> "What I wanted to happen was to lift up my arms really slowly and to have the sound of velcro peeling away, and then I was hoping to build a performance using lots of different samples ... using the gloves to make it appear that objects I was moving in particular ways had these sounds attached to them."

### Ergonomic Controls

The novices were conscious of the ergonomics of their gestural parameters, particularly from novices who were using the gloves to augment their existing instrumental practice, (e.g. using a guitar and the gloves simultaneously). These novices gave thought into which gestural parameters would be most appropriate, and how to leverage the existing shapes and movements of their instrumental playing into glove controls, such as the posture of one's hand when holding a guitar pick.

> "The other thing I spent a lot of time on is when you hold a pick, some people would do this [shifts hand], so I was thinking you could map in different picking shapes. You have to be very aware of how you're playing with the glove on."

The novices also considered ergonomics through the difficulty of performing the movements necessary for their mappings, often prioritising mappings that were "easy", which lead to iterative processes of auditioning and editing mappings.

> "The other thing I combined it with was gripping [going from open hand to fist] because that's extremely easy."
>
> "The major motif, that's not two motions [triggering two notes], that's using the chord machine and spreading the notes out, so its a single motion for two notes. Definitely cheating."
>
> "The tune had ten notes, so I thought lets just put the other five notes on the other hand, and this meant that I used the most common notes on the left hand and the less common notes on the right hand ... I had instances where I'd try to trigger a note with my third finger, and I trigger my little [finger] as well, and then I started to work on harmonies and found that I had to do little finger left and ring finger right, and then the opposite hand, and so I swapped the notes, and then I play the tune again ... so there was a period of 'ok this is now here', and then I would carry on discovering how best to do the note layout."

### Minimising Mistakes

A common consideration in mapping design was minimising the potential for performer-related mistakes. For instance, the novices would adjust the musical material they were performing, and minimise the notes available in a specific scene, to only use notes in the same tonal key.

> "I wanted to keep everything quantised to one key so that it was safe wherever you are, it would be much more challenging with a chromatic scale."
>
> "All of the notes compliment each other so you can do what you like with it."

> “ I thought it would be best to keep everything in one key so there was no place you could go where it would be wrong.”

> “As I say you can program in the notes that you want so you can’t really make mistakes as such, fourths compliment each other ... nothing will really sound out of place.”

As well as peformer-related mistakes, the novices also focused on minimising system-related mistakes. The novices would make mapping decisions based on the perceived reliability of the gestural controls.

> “I have generally been using more [directions], because they seem a bit more predictable. Up, down, the bigger movements. There’s a little too much margin for error within them so it’s a bit more difficult, that can be frustrating when you hit the notes you didn’t want to hear.”

> “My intention was to use the technology in a simple, robust way, and not get lost in the technicalities of it.”

> “I suppose I was clever about what I chose to use and how to use it [...] I was being sensible rather than dangerous.”

### Understanding Gestural Controls

A common theme from the novices was their development of understanding of the gestural parameters available and the movements that they referred to. There was often misunderstandings around what gestural parameters meant, for example, the novices often confused the orientation parameters to have translational effects.

> “I tried to map the tempo to the pitch parameter, but it just didn’t work the way I thought it would in terms of what that gesture did to the sound.”

### Issues with MIDI Mapping

MIDI and MIDI mapping to a DAW was causing problems for many novices, which became a barrier to mapping progress. Sometimes, this would be due to a lack of understanding, with the novice having difficulties working it out.

> "The pan for instance, every time I MIDI mapped anything the setting would go from half way up and it would have too much effect, and I couldn't work out how to MIDI map it so that it would start at zero and work its way up."

Often, the process of MIDI mapping from Glover to a DAW caused frustration for the novices.

> "I think the mapping, the MIDI learning aspect of the DAW I'm using [Ableton] is frustrating, because by the nature of how it works, Glover's spitting out everything that's going on, all the active MIDI messages, so you have to solo them, then go to Ableton, click the thing you want to learn, you know there's a lot of back and forth there its a bit tedious. But I don't know it there's a better way of doing it or not but that's the way I found to do it, and that was frustrating"

### Exploiting System Limitations

The participants discovered and utilised system limitations within their control mappings. For instance, one participant found that when using the extreme ends of the yaw control it cross over the maximum and minimum thresholds and would snap between $-1$ and $+1$ output values, which the participant used to create "glitchy" auditory effects.

> "I found that if I moved too far [on the yaw axis] the output would jump from the top to the bottom values. I quite liked it, it made these really glitchy sounds."

### 7.4.4. Final Performance Mappings

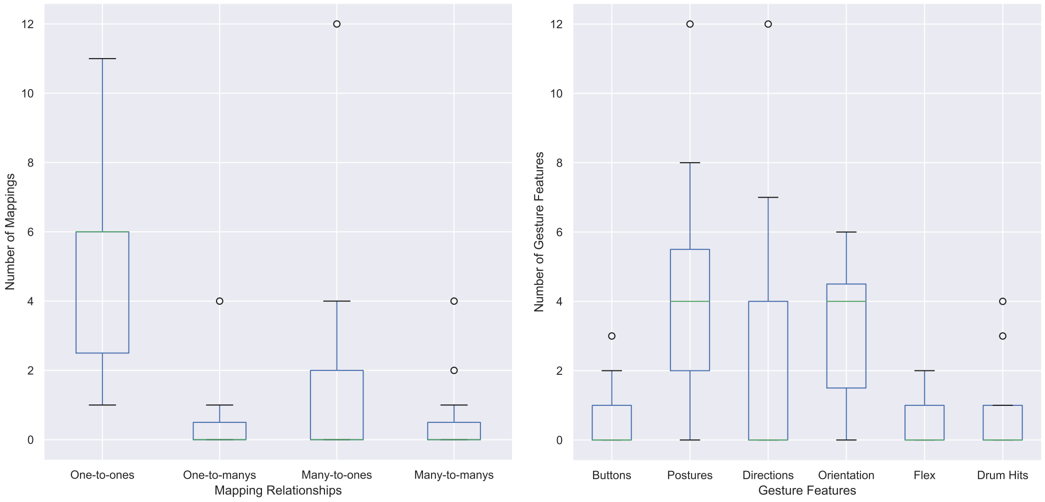
The mapping strategies used by the participants in their final performances is presented in Figure 7.15. One-to-one mapping relationships were used the most often, with postures and orientations proving to be the most used gesture features. The Note Matrix and Chord Machine were rarely used, with most mappings being implemented in the input–output objects. Continuous expression outputs and discrete note-based outputs were used in roughly equal measure.

## 7.5. Observations: Task Analysis

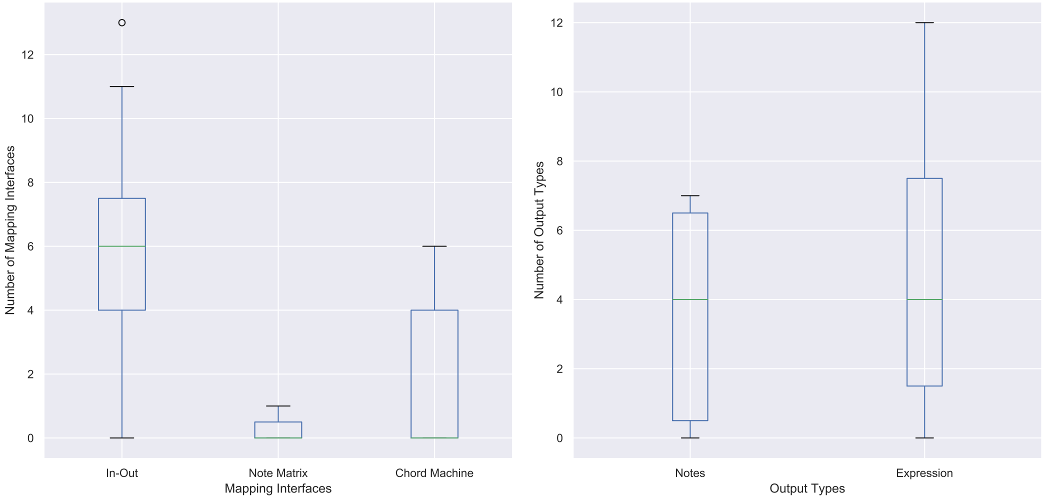
Five of the novices performed the note selection tasks, and five novices performed the expression tasks.

### 7.5.1. Mapping Design Choices

To better compare the mapping strategies between note exercises and expressive exercises, the mapping design features were scaled to the amount of mappings necessary to complete the exercise. For example, the third note selection exercise consisted of six notes, and so a minimum of six MIDI note outputs were required to complete the exercise. Meanwhile, the expressive exercises required gestural parameters to be mapped to five MIDI CC outputs in order to complete the exercise. To allow for easier comparison, each exercise was scaled by the minimum number of output mappings required to complete the exercise, to give a value of



(a) Mapping relationships. (b) Gesture features.



(c) Mapping interfaces. (d) Types of MIDI output.

Figure 7.15.: The novices' use of Glover for their final performance mapping strategies.

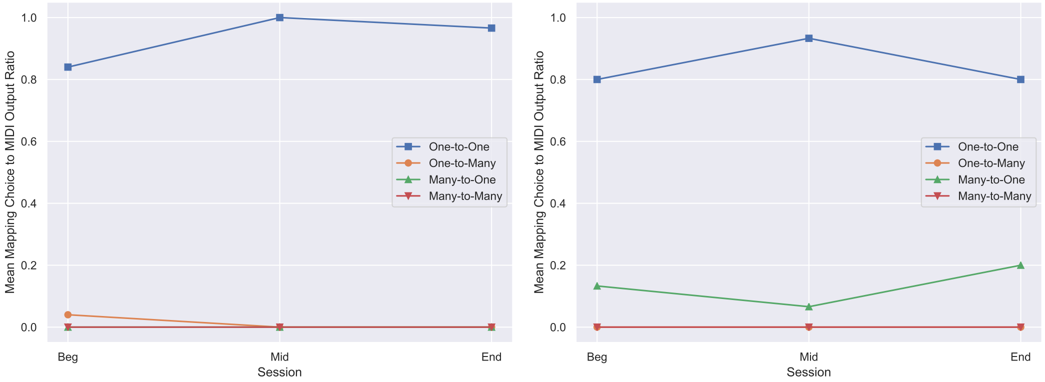
1 for minimum completion. Figure 7.16 shows the mapping choices made by the novices in the three mapping tasks (beginning, middle and end).

For the note selection exercises, nearly all of the mappings were implemented in simple, one-to-one mapping relationships (Figure 7.16a). The gesture features used (Figure 7.16c) were predominantly postures and directions, with postures being more popular at the mid-point, while directions became more popular at the end. Movements (made into discrete parameters by using the thresholding function or through use in the Note Matrix) and events were unpopular choices, that saw a small amount of use in the first exercise, but were not used in later sessions. In the novice's mapping interface use (Figure 7.16e), the note matrix was abandoned after the first task. The majority of note mappings were implemented in the Input–Output objects, while the Chord Machine became more popular in the latter two sessions.

In the expressive exercises, simple one-to-one mappings were the most popular mapping relationship to use, but many-to-one mappings were also used, particularly in the final task (Figure 7.16b). Regarding gesture features, movements were the most popular, being used for nearly all mapping solutions. Events were used to an extent in the first task, but were not used in the later tasks. Meanwhile, the use of postures in mappings grew over the tasks. All of the mappings implemented were done so in the Input–Output objects.

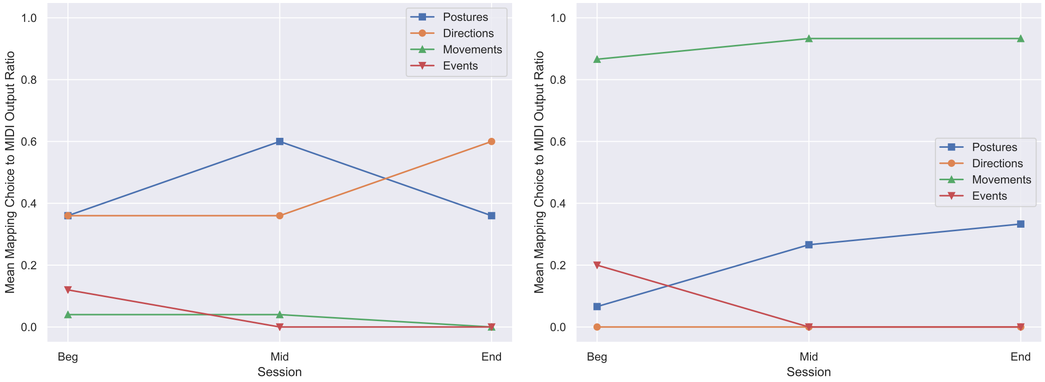
### 7.5.2. Development of Mapping Choices

To examine how much the novices' mapping strategies changed over the three mapping tasks, the difference in mapping behaviour between the session tasks was examined. To do so, the difference in the use of individual gesture features (postures, directions, movements, and events), mapping interfaces (input–outputs, chord machine, and note matrix) and mapping relationships (one-to-



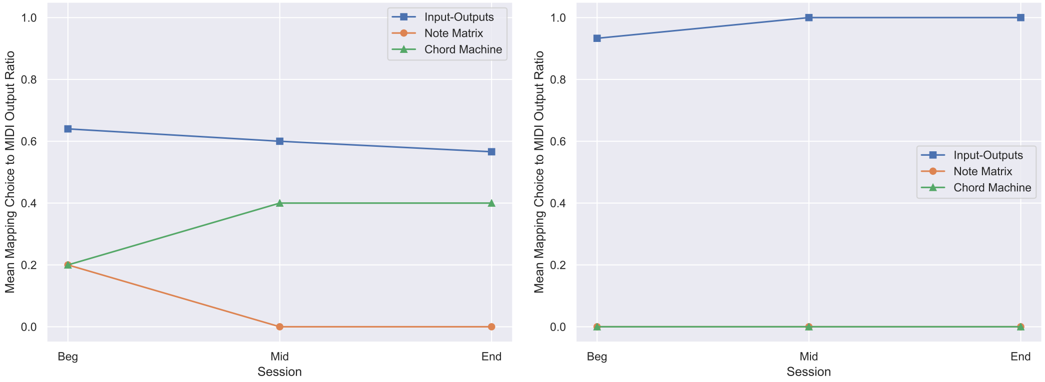
(a) Mapping choice (Notes).

(b) Mapping choice (Expression).



(c) Gesture choice (Notes).

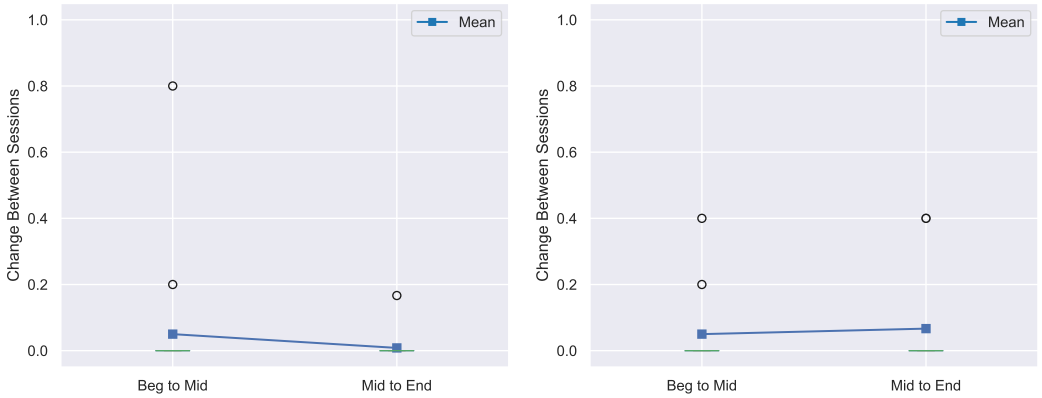
(d) Gesture choice (Expression).



(e) Interface choice (Notes).

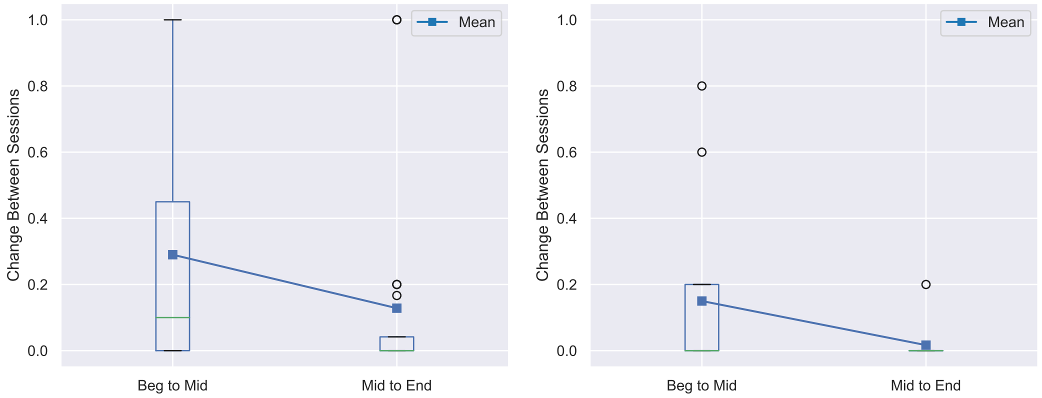
(f) Interface choice (Expression).

Figure 7.16.: Ratio of mapping choices to minimum necessary MIDI outputs over the three mapping tasks (L: Note Selection, R: Expression).



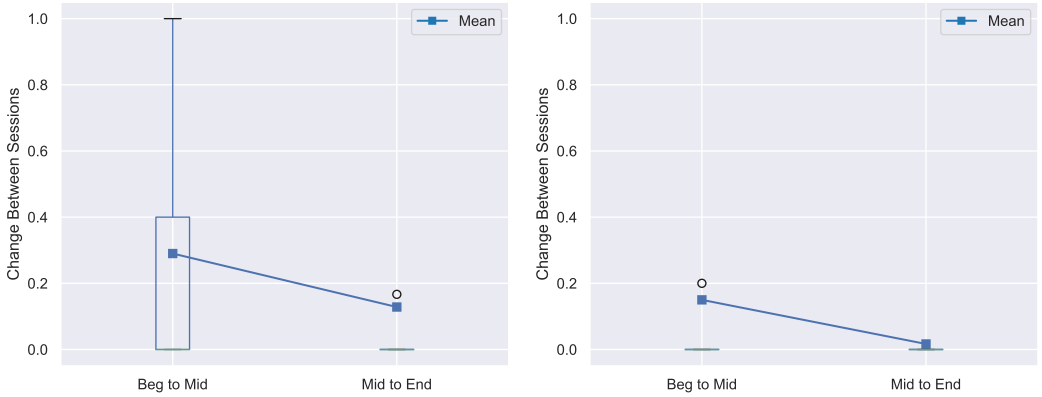
(a) Mapping Relationships (Notes).

(b) Mapping Relationships (Expression).



(c) Gesture Features (Notes).

(d) Gesture Features (Expression).



(e) Interface Choice (Notes).

(f) Interface Choice (Expression).

Figure 7.17.: Amount of change in mapping choices between beginning to mid-point and mid-point to end-point mapping tasks (L: Note Selection, R: Expression).

one, one-to-many, many-to-one and many-to-many) between the mapping tasks of each individual novice was measured (Figure 7.17). For example, for a note selection task with five MIDI note outputs, if a novice used five postures in the first exercise and five directions in the second, this would be measured as a 1.0 change for posture use (5 to 0) and a 1.0 change for direction use (0 to 5). If the participant used five postures in the first exercise and 3 postures and 2 directions, that would be measured as a 0.4 change for posture use (5 to 3) and a 0.4 change for direction use (0 to 2).

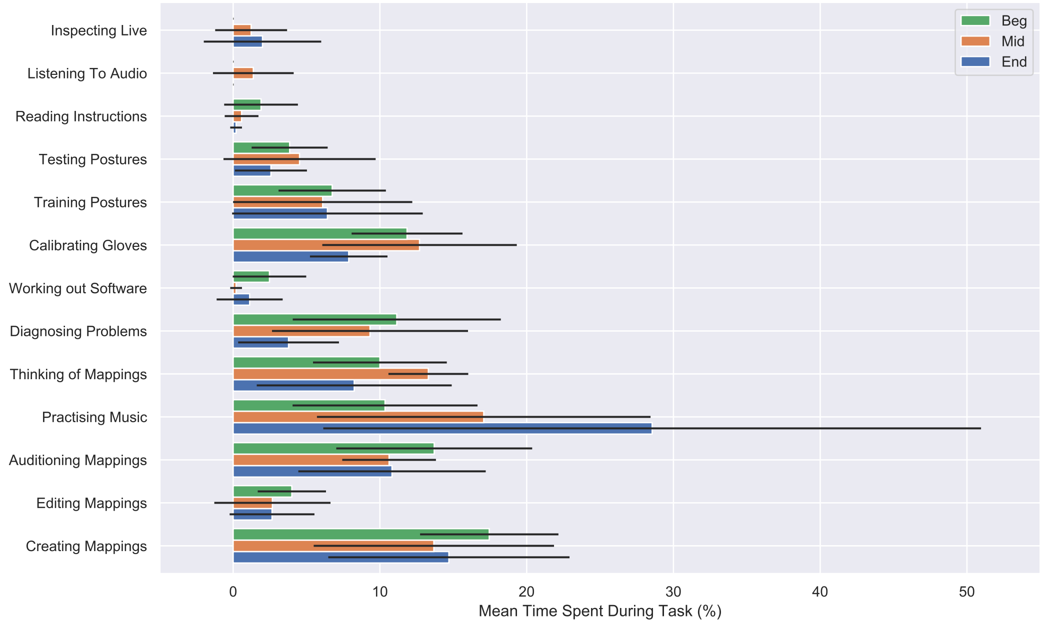
For the note selection tasks, there was an observable amount of change in the participants' choices between the beginning and mid-point tasks, and very little change in choices between the mid-point and end tasks, for the participants' use of gesture features and interfaces. The participants' use of mapping relationships remained fairly constant over the three exercises, showing very few changes between the beginning and mid-point and mid-point and end sessions.

The participants' mapping choices remained constant between the expression exercises, with very little change in the participants' mapping choices between these tasks. There was no observable difference between the beginning and mid-point and mid-point and end exercises for mapping relationships or interfaces, but for gesture features there were some changes between the beginning and mid-point tasks and no changes between the mid-point and end tasks.

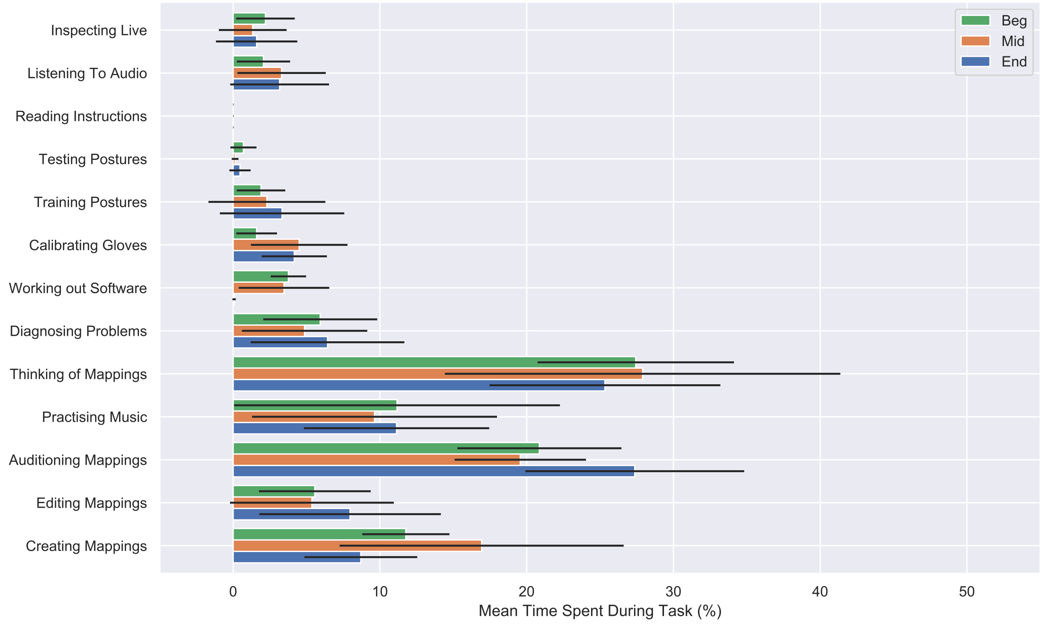
### 7.5.3. Time Spent

The participants were observed as they performed the mapping task, and the time that the novices spent on the mapping tasks was coded into a set of subtasks:

1. *Creating Mappings.* Implementing new mappings in Glover.



(a) Note Selection Task.



(b) Expression Task.

Figure 7.18.: Breakdown of time spent during mapping tasks.

2. *Editing Mappings.* Attenuating existing mappings, such as setting minimum and maximum output values.
3. *Auditioning Mappings.* Testing the newly implemented or attenuated mappings.
4. *Thinking of Mappings.* Thinking of new mappings to create and implement.
5. *Practising Music.* Practising the musical material with their implemented mappings.
6. *Calibrating Gloves.* Glove calibration tasks such as setting forwards and setting finger flex minima and maxima.
7. *Training Postures.* Adding new posture examples to the posture recogniser.
8. *Testing Postures.* Testing the posture recogniser outside of its use in mappings.
9. *Diagnosing Problems.* Periods in which the participant is diagnosing issues with their mappings and Glover.
10. *Reading Instructions.* Referring to the technical manual provided.
11. *Working out Software.* Exploring the features and affordances of Glover or searching for a specific feature.
12. *Inspecting Live.* Inspecting the Ableton Live session for the task.
13. *Listening to Audio.* Listening to the reference material.
14. *Other.* Any other task that does not fit in the above categories.

A break down of how the participants spent their time performing the mappings tasks is provided in Figure 7.18. Between the sessions (beginning, middle and end) in the note selection tasks, more time was dedicated to practising music in the later tasks, while less time was dedicated to auditioning mappings, calibrating gloves, and diagnosing problems. Overall however, the time the novices dedicated to each sub-task did not change between tasks.

The differences between the note selection tasks and the expression tasks were more pronounced, with the expression task novices dedicating much more time to thinking of and auditioning mappings, with less time spent calibrating, diagnosing problems, practising the musical material and training and testing the posture recogniser.

### 7.5.4. Interviews

From the short interviews after the mapping tasks, the following themes emerged through Thematic Analysis (Braun and Clarke, 2006).

#### Note Selection

Throughout the note selection tasks, the participants discussed their lack of understanding between whether the issues they faced were attributed to the system error originating from the gloves or to their own performer errors, and their difficulties with accidental triggering and using the posture recogniser.

> "I'm just not sure if there's a problem with the precision of the gloves, or if its an issue with me needing to be more precise and practise those movements"
>
> "I don't really know what was going on with the postures, because I couldn't get them to register."
>
> "there's certain considerations that you don't think about, passing through notes was a big issue."
>
> "I could go back and see if I could get the postures working, but it was difficult to find postures distinct enough that they weren't going to get confused"

> "Again its frustrating I can't seem to use the postures, I'm sensing that I'm doing something wrong."

The participants discussed how they designed mappings that reflected existing instruments, particularly with how the notes should be laid out reflecting a piano keyboard.

> "I feel that although it might not be as expressive in a performance, its a much easier analogy as a keyboard player, it feels much more natural to play the notes with these postures, it feels like I'm just playing the piano."
>
> "With only five notes, we tend to think of music as a keyboard left to right, so at first I did left and right, then left, right and then up but it doesn't go with this idiom of everything going from left to right, so it made more sense to put up between left and right so you're moving between them."

The note selection participants discussed how their mappings developed over the three tasks to focus on control and ease of playing.

> "In the end it ended up being just using one hand and trying to figure out the ones that don't interact with each other too much, very distinct movements between the directions worked well."
>
> "The way I structured it was helpful as the first three notes were on this hand and the other three on this one, so with the way the melody went it made it easier to play, and I could get it down faster than last time."
>
> "I just wanted to get something that would work. It was more about efficiency than aesthetics."

### Expression

In the expression tasks, the participants were much more focused on the aesthetics of their mappings, with novices designing mappings to reflect the movements of conductors. These participants also discussed their desire to design mappings that reflected the theremin, but did not do so due to the incompatibility between the gloves gestural parameters and the theremin's control movements.

> "I was trying to create an analogue with how you'd usually see conductors, and in my mind up and down is usually tempo and amplitude."
>
> "I was thinking of doing something like a conductor."
>
> "The only other way I would want to do it is the theremin way, I'm sure I could do that but it would be complex [to map]. I think its because the piece is quite thereminy [sic], I had that analogy in my mind."
>
> "I think there's an obvious comparison [to the theremin]. As I went through the session I tried to move away from that to something that would work better for [the gloves]."

These participants were also influenced by existing musical interactions, designing mappings based on rotary knobs or 2D parameter controllers common in DAWs.

> "Maybe because I'm used to seeing it with knobs, but it made sense to have the filter cut off being roll."
>
> "Focusing on the right hand, I started off thinking the brightness was the easy one being a filter because its got that sweep its better to control it with the roll."

> "I mapped volume on the vertical pitch parameter and tempo to yaw, basically so I could control time and volume in one graph. I was picturing it as a 2D graph."
>
> "Certainly the vibrato and the brightness really do make sense to me, maybe because I'm so used to seeing it, as a pair of dials you can twist."

The participants also discussed how their mapping solutions changed little over the three tasks due to their solutions feeling "natural" and intuitive. The changes that these participants made towards the later sessions was around increasing the controllability of their mappings, by using many-to-one mappings with qualifiers added to their continuous movement controls.

> "My mappings were really similar to what I did last time. They made so much intuitive sense that to do anything else would be unintuitive and a bit forced."
>
> "It feels like I should do the same things as it's the same expression parameters."

> "The other thing is when you have so many parameters going on you sometimes control more than the ones you want to, so I wanted an approach where I could be a bit more selective by adding a qualifier and have more specific control."

## 7.6. Discussion

### 7.6.1. Open Exploration

In the context of their personal glove practice, the novices were concerned about maintaining control in their mapping design, and prioritised designing mappings that minimised performer-related mistakes and provided ergonomic gestural controls. The novices would also minimise errors by simplifying the musical material they intended to perform for their final performance piece, such as only using musical material that consisted of a single key. The usage logging revealed that over time, mapping designs tended towards using the input–output mapping interfaces to implement simple, one-to-one mapping relationships, with one-to-one relationships being the most popular mapping relationships used (Figure 7.6). This supports the themes of control and error minimisation that were elicited during the interviews.

The novices' use of mappings based on metaphors of existing instruments, such as guitars and violas, reflected themes around embodiment (see Section 3.2), specifically the embodied idea of "what is it like to *do* the music?" (Cox, 2016). In this case, when thinking of mappings for specific musical material, the novices are imagining how they would *do* it. This was mostly done at the beginning of the novices time with the gloves, and was used as an initial learning exercise, to help them learn the affordances of the gloves and Glover.

The novices were also concerned with the transparency of their mapping designs (Fyans, Gurevich and Stapleton, 2009b; Fels, Gadd and Mulder, 2002), and designed mappings that maintained a coherent visual, embodied link between their movements and sound output. This was motivated around making sure their audience would be able to clearly perceive that their actions were controlling sound, and influenced the novices' use of metaphor in mapping design. Unlike the experienced glove musicians (see Chapter 8), the novices were less concerned about expressing their own aesthetics in the visual action–sound relationships, with only one novice discussing a mapping that expressed a novel visual metaphor.

Similar to the experienced glove musicians (Chapter 8), the novices were also concerned with the reliability of their gestural controls, and had issues with accidental triggering, particularly with the posture recogniser. These problems limited the novices mapping choices, with many novices feeling that they could only reliably make use of two postures: fist and open hand. The usage logging results revealed that despite these reported issues, postures made up a large proportion of the gesture features used by the novices in their mapping designs (Figure 7.10), and were more popular than other Qualifier choices. This suggests that while unreliable, the postures remained a useful mapping option for the novices.

This continued perseverance with the posture recogniser is indicative of the control affordances of the gloves, and the focus in the novice's mapping designs on the audience's perception of their mappings. The novices, motivated to design mappings with high audience transparency, mapped expressive output parameters to large, noticeable movements through the orientation-based gesture features. The postures, which use the gloves' flex sensors, are qualifiers that compliment the orientation parameters drawn from the gloves' Inertial Measurement Units (IMUs), as they can be used without restricting the range of the orientation's movement. In contrast, the direction qualifiers, also drawn from the IMUs, would heavily restrict the orientation parameters in a many-to-one mapping, as they would only be usable within the direction's active lobe.

An interesting observation was how the participants were finding and exploiting limitations of the Gloves system, such as running over the ends of control ranges to create glitchy effects. This reflects observations made in similar research, which found that musicians pushed the designs of a DMI and exploited unseen affordances (Zappi and McPherson, 2014a).

### 7.6.2. Tasks

In the task-based exercises, the novices used simple controls to express the musical material, with one-to-one mappings being the most popular mapping relationship used in both note selection and expressive mappings tasks, and being exclusively used in the note selection tasks.

In their solutions for the mapping tasks, the participants in the note selection tasks prioritised the ease of playability, using reliable controls that minimised accidental triggering. The expression participants focused more on the visual aesthetics of their mappings, using metaphors of existing music interactions and instruments, only beginning to consider controllability in the latter sessions. This move towards controllability is reflected in the mapping choices (Figure 7.16), where the use of many-to-one mappings and postures increased in the last task.

The task analysis results showed that the participants quickly established consistent mapping behaviour, with a drop in the amount of changes in mapping decisions between the beginning and middle sessions and the middle and end sessions being observed in the note selection tasks, while very little change was observed between any of the expression tasks (Figure 7.17). The lack of change in the expression tasks, revealed in the post-task discussion, was attributed to how the participants felt their initial mapping solutions felt natural and intuitive, reflecting embodied principles (Section 3.2), and that they were reluctant to deviate from them.

No significant differences were observed in how the participants allocated their time during the mapping tasks over the three sessions. However, there was an observable difference in time allocation between the expression and note selection tasks. The expression task novices allocated more time for creatively thinking of new mappings, with more time spent by the note selection novices performing technical tasks such as calibrating the gloves, training and testing the posture recogniser, and diagnosing problems. There was also a difference in the time spent practising the musical material in the expressive and note selection tasks, with the expression tasks allocating less time to practising their mapping solutions.

This is likely due to the nature of the two tasks, as a task involving participants having to perform a series of notes means that there is an explicitly "correct" way of performing the musical material: they must adhere to the musical *text* (Taruskin, 1995). As well as this, the novel nature of the instrument may lead the participants to wanting to demonstrate and convince the audience that they were in control. Meanwhile, in expression-based tasks, the participants were tasked with deviating from this musical text, with the participant having artistic control over the manner in which they did so. This difference in the nature of the tasks most likely influenced the priorities in the participants' mapping design, particularly with the note selection participants' focus on minimising mistakes. Mistakes in note-based tasks are explicitly defined by the musical material, while mistakes in expression tasks are often undefined, with musical expression open to performer interpretation.

While the priorities in mapping between the two tasks differed, both groups used metaphors of existing instruments and music interaction interfaces, which revealed the participants thinking in an embodied way around "what is it like to *do* the music" (Cox, 2016), as the participants discussed metaphors of typical DAW knobs and graphs, as well as theremins and pianos.

### 7.6.3. Limitations

A limitation of this study is that the total time the participants used the gloves over the study period varied greatly, and a participant's use could have been distributed unevenly across the month-long study period, with a participant potentially using the gloves for more time in the first two weeks or the last two weeks. This could have had an effect on the mapping choices exhibited in the mid- and end-point mapping tasks.

While the longitudinal approach allowed the research to investigate mapping design practice beyond the initial exploration phase of interaction, the month-long study period is still a very short time frame to examine the development of musical practice. A smaller number of participants over a longer time period could have provided more insights into creative mapping design with the gloves, similar to previous research that examined the compositional practice of a single composer over three years (Collins, 2005).

## 7.7. Summary

In this chapter, the early development of mapping practice with the Mi.Mu gloves has been examined. A longitudinal, mixed-methods study was conducted, following and observing the development of glove mapping designs over a month-long period. Usage logging and interviews were used to examine mapping in the context of a musician's own practice, and mapping design was also measured through a series of mapping tasks conducted throughout the study period. From this, the following findings were made.

1. Novice glove musicians predominantly used simple mapping strategies, with one-to-one mapping strategies being the most used mapping relationship in their mapping designs.
2. Novice glove musicians tended to think in embodied ways around "what is it like to *do* the music" when designing mappings, evidenced through their use of metaphors in mapping design.
3. The novice glove musicians were influenced by the audience's perception of their mapping strategies, and aimed to provide transparent mappings.

4. Novice glove musicians leveraged their existing musical practice in their glove performances, through using the gloves to augment their existing instruments, and through designing mappings around metaphors of their instruments.
5. The controllability and reliability of mappings are a major influencing factor on novice mapping design, and novice glove musicians prioritised minimising performer-related mistakes through using ergonomic, "easy" mappings, and minimising system-related mistakes caused by accidental triggering and posture recognition errors.
6. Accidental triggering and mistake minimisation were mainly issues when mapping discrete notes, while expressive parameter mapping afforded more time to be allocated to creative mapping design.

These findings, along with findings from Chapters 6 and 8, contribute towards a set of heuristics for mid-air mapping design, presented in Chapter 10. The following chapter (Chapter 8) examines the mapping design practice of expert glove musicians.

# 8. Experienced Musical Practice with the Gloves

In Chapter 6, a marked difference between novice and experienced glove performers was observed. Chapter 7 explored the interactions of novice glove musicians in greater depth, finding that novices tend towards designing simple one-to-one mappings based on embodied metaphors of instruments and established metaphors of music and movement (Wilkie, Holland and Mulholland, 2010). This chapter explores the interactions of experienced glove musicians in greater depth. Four expert-level glove musicians, who all use the gloves in their professional musical practice, are interviewed about their mapping design practice. Based on the findings in Chapter 6, the hypothesis for the work presented in this chapter is that the experienced glove users would report prioritising ergonomic control over adherence to embodied musical metaphors.

The work presented in this chapter has previously appeared in the following publication.

Brown, D., Nash, C. and Mitchell, T. (2018) Simple Mappings, Expressive Movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity*. 29 (2–3), pp. 129–148.

## 8.1. The Interaction of the Experienced Users

As discussed in Section 2.10, a typical Human-Computer Interaction (HCI) evaluation often takes place using participants who are largely unfamiliar with the system being tested, biasing the analysis to initial usability (Siegel, 2012). Although this is an important aspect for user interfaces and Digital Musical Instruments (DMIs), it is equally useful to evaluate the practice of an experienced user, whose insights reveal much about an interaction than novice users. This is particularly important with interfaces for musical expression, where analyses of expert performers are critical in determining the true expressive and virtuosic capabilities of an instrument.

To allow for an exploration into glove musician's interaction in the context of their own creative practice, a qualitative approach was taken. Qualitative, observational approaches are used extensively in User Experience (UX) research, and while it can lack quantitative precision, the approach allows research to investigate phenomena within real-world contexts (Mackenzie, 2013; Kaye, 2009).

## 8.2. Method

Grounded Theory (Glaser and Anselm, 1967) was used to examine the musical practice of glove musicians. Four experienced glove musicians took part in an unstructured interview focusing on the factors that affect their mapping design and the use of the gloves in their musical practice. The interview data was concurrently collated and analysed to facilitate theoretical sampling. The four musicians approached were those whose professional practice with the gloves includes live performances for large audiences. Qualitative research is highly dependent on the interpretations and perspective of the researcher (Elliott, Fischer and Rennie, 1999), and consequently, it is important to expose the background of the researcher. The lead researcher is not a practising glove musician, but does possess intimate knowledge of the associated systems and software. This background could influence the perspective of the analysis, which could potentially focus on aspects of the system such as reliability or usability issues, and less focus may be given to aesthetic or artistic considerations. By acknowledging this potential bias, the researcher hopes to be aware of its effects in the analysis and adapt accordingly. The analysis builds on an understanding of embodied interaction (Section 3.2), where knowledge is considered to be built from an individual's bodily experiences in the world. This position has been influenced by HCI and DMI literature, in which embodiment is a common epistemological position (Dijk and Hummels, 2017; Cox, 2016; Kaye, 2009; Leman, 2008; Dourish, 2004).

### 8.2.1. Participants

Established glove musicians who all use the gloves in their professional practice were approached to take part in this research. Four musicians agreed to participate.

The four participants have all been using the gloves in their professional performance practice for several years, each performing at national and international tours and events. The group are a strong community, and each musician is well known to the others. They also meet regularly (every six to twelve months) to share their work and provide each other with support and feedback. Throughout the interviews, the musicians referred to each other's work, so each musician has been assigned a letter: A, B, C and D. In addition to occasional performances with the gloves, Musician C works extensively as a facilitator, designing and developing mappings for others' performances. All four musicians have been tied to the development of the gloves and its software to varying degrees, providing at times significant design input and feedback, as well as suggesting and designing mapping features such as the chord machine and note matrix instruments.

## 8.3. Results

### 8.3.1. Simple Mappings

The musicians were found to use simple, one-to-one and few-to-one mappings that minimise the potential for performer-related errors. Often referred to by the musicians as “practical” mappings, they provide the musicians with control over their musical content that can be mastered with little effort. For instance, Musician A routinely uses combinations of open hand and fist postures, coupled with the directional lobes of movement to quickly facilitate auditory feedback.

> “...practical, make it work quickly, use the different directions and opposite postures so fist and open hand or something... that’s like the quickest way to do a lot of different things.” – Musician A.

Simple mappings are also used due to the pressures of performing in front of large audiences, where they are used to minimise performer error.

> “I had to figure out a way of mapping everything that I was going to teach them, that was, interesting for people to watch, and then interesting for them to play, but dead simple...it’s not that [name] couldn’t do that if they had the time, they just did not have the time. I literally had 45 minutes to teach them the song and then they performed it in front of 6000 people.” – Musician C.

Simple mappings are also used to make control relationships obvious to audiences. For instance, Musician B described how they have simplified their mappings as audience members are unable to accurately perceive more complex control relationships. They found there was little point in making mappings overly complex when they could communicate their musical intentions using mappings that are easier to perform.

> “That’s another thing, it’s interesting when tech people think they know how it works, and I’ve had quite a few people think that I’m launching lots of clips, so when I’m doing the violin thing I’m just triggering a sample and then miming to it ... there’s times for having a backup clip that I could launch if I needed to fall back, because sometimes it doesn’t come across, and the audience isn’t part of that conversation, no matter how hard I make it for myself.” – Musician B.

### 8.3.2. Expressive Movement

While the musicians use simple excitation mappings, they embellish their excitation movements with theatrical ancillary movement. This ancillary movement has no effect on the sound parameters being controlled, but is used by the musicians to express aesthetic intentions, and to make performances more engaging for their audiences. The incorporation of ancillary movement comes from the performance context of the musicians’ practice, and was often referred to as “performance theatrics”. Making their performance movements more visually engaging was a priority for both Musicians A and B, who also discussed their work or desire to work with choreographers.

> “...but also exaggerating certain movements... first of all you’re on a stage so people are looking at you, and before I did anything of the choreography it made me really aware of how I moved on stage and felt not that super comfortable about it, because I’m not a trained dan-

> cer and suddenly I had to do movement that I didn't necessarily like. So thinking more about the choreography and making it more exaggerated and theatrical made me way more confident performing because now at least I knew what I was doing and I was sure it would look cool because we thought about it, you know." – Musician A.

This aesthetic consideration in the design of mappings and consequently movement is something that has become more prominent in the musicians' practice as their Glove performance has developed. Both musicians A and B remarked that in their early mapping practice they would use the "next available" control, and their focus was on creating performable solutions. For example, the musicians would cycle through the most ergonomic hand postures: fist, open hand, one finger point, two finger point. As their practice with the gloves has developed, they have moved towards considering the visual aesthetics of their performances.

> "There's a video of me on the day that I came up with it, and I just have one glove and I just go two finger point, fist, two finger point, fist, and just pitch, that's all, and that works exactly the same way as I do it now but with this and the turning around and putting my hands to the side for no reason" – Musician A. "[T]hings would be quite small, I used to think in terms of the next available thing, like if I'm pointing up then I could point forwards." – Musician B.

This development reflects the advancement of the musicians' creative practice, and a move from mapping for functional control to a more abstracted, aesthetically driven approach. Musician A discussed their use of ancillary movements to develop more aesthetically engaging performances. They discuss how in one musical phrase the last gesture has no musical effect, but is performed due to the perceived movement in the music.

> "...there's one thing that I do that doesn't trigger a different chord, I do [gestures] this then this, and the two finger point actually doesn't trigger anything. I don't know why I do it, but it's just because in the music it feels like something changes so I feel stupid if I don't change my posture, but it's not triggering anything." – Musician A.

While the musicians' mappings are simple, the use of theatrical ancillary movement allows them to express their aesthetic intentions and provided engaging performances. Musician C, who believes that something was "aesthetically lost" when mappings were simplified, remarked that Musician A's use of ancillary gesture added to their performances.

> "Something is aesthetically lost, for sure. That's why I think it's so cool what [Musician A] is doing as they're actually probably doing simple things, but they're incorporating them into a choreography that makes them seem and look and feel more subtle." – Musician C.

Similarly, Musician D remarks how it was the simplicity of Musician B's performances that provided an engaging performance.

> "When I saw [Musician B] for the first time, they were very specific in the things that they did, and one of the things that really caught me was that I was trying to do too many things, why am I doing so much? I could really par down the pallet and be just as impactful." – Musician D.

### 8.3.3. Metaphors in Mapping

In their expressive movements, the musicians often used visual metaphors of lyrical or musical material. The metaphors would often correlate with established metaphors relating music to spatial properties, such as space being used to conceptualise musical pitch: UP AND DOWN IN PITCH IS UP AND DOWN IN SPACE (Wilkie, Holland and Mulholland, 2010; Lerdahl, 1988).

> "This feels to me, terms of pitch, the audience will perceive that the chord does indeed change, it goes down, but it's not a big change, so you need a bigger gesture to get to the next chord. So, if the interval's further away you need to make the posture bigger." – Musician B.

Other times, gestural metaphors representing lyrical content were used. Musician A designed mappings that reflected the meaning of their lyrics, for example, Musician A used a metaphor of OPENING HANDS IS OPENING EYES. In this mapping, open hand postures and the up direction on both hands were mapped to a MIDI trigger, performed by the musician in front of their eyes.

> "...it's about my friend waking up from a coma, which actually happened like a year ago, so the chords that I trigger, the first time I trigger the chords I go like this [open hand gesture in front of eyes] because represents their eyes opening." – Musician A

These metaphors aid the musicians in creating visually engaging performances through their simple mapping strategies. Musician D reflected on a performance of Musician B, who incorporated a visual metaphor of RELEASING A FEATHER IS RELEASING A NOTE into a simple mapping of an open hand posture qualified by directions triggering MIDI note on messages.

> "I saw them have five or six chords in a space and it was just a piano sound, and when they opened their hand in a zone it would let out a chord, and it looked like they were letting them off like releasing a feather or something, into the air, and it was so beautiful. I knew

> that it was just forward, up, left, right but it looked like 'I'm going to put it into that wind bit over there, and I'm going to let the feather go over there' and then they were singing with it and it was just really gorgeous." – Musician D.

The musicians' use of metaphor has also developed from being based on the interactions of traditional instruments to more abstract metaphors reflecting the relationship between movement and music (Cox, 2016; Johnson and Larson, 2003). For instance, Musician B used mappings that used interaction metaphors of existing instruments in their early performances such as guitars and violins. For example, for a guitar mapping, Musician B would map postures on the left hand to note selection, which would then be triggered by a Drum Hit event on the right hand. This approach afforded both them and their audience transparency (Fels, Gadd and Mulder, 2002), but they now find this approach to be limiting creatively, and they have moved to thinking more abstractly about how their movements relate to the music itself.

> "I think it would be good to work with a choreographer at some point. For me the mileage had run out [with their previous show]. But it was great for what it was at the time, to go there and play those shows and be like here's my invisible guitar, here's my invisible drum kit and whatever, that sets it up as a gestural thing that people can understand. If I started from day one with all this abstract stuff then what is it? Nobody knows, there's no way in. But that has to evolve. I'm thinking about what is the movement of the music. It's the first time I've gone the other way around and thought 'what would it be?'." – Musician B.

Figure 8.1.: The “secret finger” posture

### 8.3.4. Accidental Triggering

A hindrance to the musicians use of aesthetic mappings is the prevalence of accidental triggering. For instance, due to the snap-to-nearest-class behaviour of the machine-learning-based posture recogniser, it often mistakes the relaxed hands of the wearer as an “open hand” posture. If the musicians map any controls to an open hand, they can often be triggered unintentionally. The musicians have developed several strategies for dealing with accidental triggering, one being the avoidance of certain controls. For instance, Musician A avoids using “open hand” postures in their mappings, and instead uses another posture they call “secret finger” (Figure 8.1), a posture similar but subtly different to an open hand. “I do [secret finger] when I don’t want people to pay attention to my posture because it’s almost no posture, and I usually do this one when I kind of want to do it with an open hand but I just need the control.” – Musician A.

One strategy that the musicians employ when they wished to use the open hand posture for metaphoric or aesthetic purposes is to use "hidden" controls alongside the open hand posture (with a Boolean AND relationship) to ensure intended control. For instance, Musician B described an instrument using a metaphor of pushing a note between directional lobes. To stop accidental triggering, they use an additional qualifier for each control: the down direction of their other hand. This extra qualifier is considered "hidden" as it does not form part of the performance movement, and the intention is that it goes unnoticed by the audience.

> "When I did the pushing the synth thing there's a qualifier on the other hand so it's only active when I'm pointing down." – Musician B.
> "So I want to play with open hand, right open hand forwards, which I will map, but that's going to trigger all the time, so I'm going to have to have a sneaky other thing, like only when my left hand is down or something." – Musician A.

Similarly, the name "secret finger", a posture shared by Musician's A and D, suggests that it is the musicians' intention that the exact nature of the control is not perceived by the audience. Musician A does use open hand postures when they wish to express a specific aesthetic intent, such as the mapping used to represent a lyric about opening eyes. However, Musician A was acutely aware that this mapping choice is vulnerable to accidental triggering and described how they immediately return to a "fist" posture once the phrase has been performed to minimise risk.

> "I definitely don't make too much dependent on open hand. In that one with the opening eyes thing I do trigger stuff with open hand but once I've done it I immediately go back to fist just to make sure I don't trigger it again." – Musician A.

### 8.3.5. Reliability in Gestural Controls

The issues with accidental triggering leads to the musicians needing to balance their aesthetic intentions with reliable controls, and influences the way musicians create simple, “practical” mapping strategies. For instance, when Musician A designed mappings that are used to express their aesthetic intentions through metaphors, it is important that the controls not only did this, but could be reliably triggered. For example, a mapping that used a fist posture on each hand coupled with inwards directions that when performed represented two wires being connected.

> “I sing ‘rewired’ [gestures bringing hands together] because this is the things that come back together. It works and its practical.” – Musician A.

This leads to the musicians considering the “robustness” of their mapping choices, which is mostly done when the musicians consider posture controls. The behaviour of the software’s Posture Recognition algorithm causes it to recognise and trigger “pass through” postures: postures that the hand unavoidably “passes through” as it transitions from one posture to another. For example, if a musician is moving from a fist to an open hand posture, if they move in such a way that the index finger starts to extend before the other fingers, the posture recogniser may briefly register a first finger point posture (Figure 8.2). This vulnerability causes the musicians to consider the kinematics of their hands and their choice of postures carefully; what they frequently referred to as the “robustness” of their posture choices.

> “Puppet hand isn’t a very stable posture to be doing that with either. It’s not very robust in terms of the likelihood of it happening during other things, you know, when you’re gesticulating, there’s normally

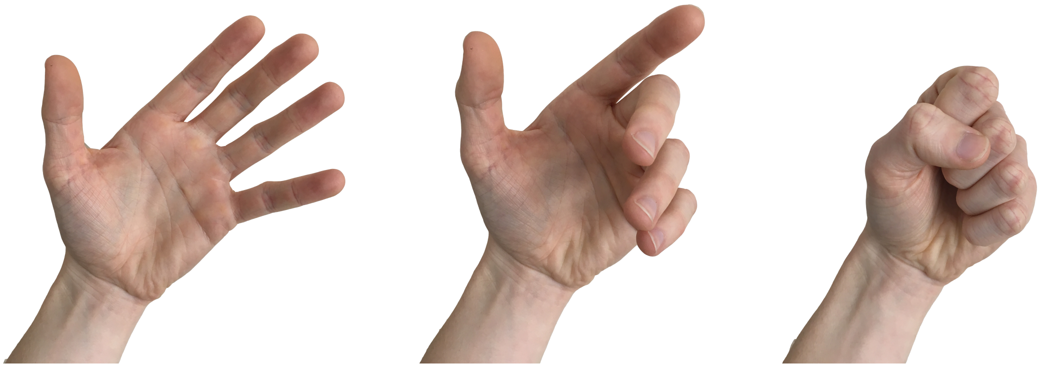


Figure 8.2.: Example of a "pass-through" posture: as the hand moves from "Open Hand" to "Fist", a "One Finger Point" posture is recognised.

> a puppet hand in there. You think about hands in a really different way." – Musician B.

While most of the musicians reported that they often change their posture choices to more robust, practical set that minimises these types of errors, Musician C reported that they practise with the posture recogniser to develop the proprioception necessary to master their posture choices. Furthermore, Musician C also added how they can gain reliable control by practising their movements rather changing their posture choices.

> "...and I've spent enough time with the software, a lot of time with the software, so I feel like I've had a chance and I continually have a chance to build for myself quite subtle and robust posture changes." – Musician C. "If it's just me I'll persevere and I'll practise and I'll practise and I'll practise." – Musician C.

This motivation could be due to the nature of Musician C's personal Glove practice, as they do not perform in front of audiences to the extent that the other musicians do, and therefore they may not have the same motivation to develop mappings that mitigate performance errors.

### 8.3.6. Personal Aesthetics

Mapping design has become a very personal expression of aesthetics for the glove musicians. Enabling musicians to design their own mappings means that choices vary wildly between practitioners.

> "What we're doing is different enough from each other that we've all invented our own standard way of doing things. I inherited a little bit from what we were developing together with [Musician D], but I know [Musician A] does things their way, [Musician B] does things their way, they also inherited things from us, because we taught them initially how to use them, but I'm sure they've developed their own workarounds." – Musician C.

This lack of shared practice is interesting considering the closeness of the glove community and given that new users are often introduced to the gloves by experienced users demonstrating and sharing elements of their own practice. This suggests that mapping is a very personal creative endeavour for the musicians. For instance, Musician A expressed a reluctance to perform using mappings designed by others, as they saw the design of their mappings as a dimension of their musicianship. Musician C also remarked on the importance of designing one's own mappings to provide engaging and distinct performances.

> "I feel like playing with the gloves is such an expression of how I see and feel music? So there's almost no point in copying someone else's movements or sound–gesture relationships because playing with them is part of the expression, totally, in how you use them." – Musician A. "I feel like that the ability to spend time with your own mappings, and create your own mappings, is really important for making something that is really engaging visually." – Musician C.

One exception to the lack of shared mappings is Musician C's work with a collaborator. The collaborator's glove musicianship is purely in performing with them, and they do not have the same creative investment in mapping design. The collaborator wanted to perform a cover of Musician D's material, and expressed a desire to perform it using mappings designed by Musician D.

> "I had kind of figured out how she could start the song, and then [collaborator] decided they wanted to have [Musician D] do it for them."
> – Musician C.

While aesthetic mapping practice has become very individual, there has been a development of standard practice around technical aspects of glove mappings, such as using the buttons on the glove to initialise the glove's orientation parameters, as observed in Chapter 6. As these mappings are related to solving system related issues (for instance, Musician C advocates for "kill all notes" control on the left-hand button), there is no personal aesthetic investment in the controls, and they are freely shared and copied between musicians.

The desire for personal customisation extends to the hardware interface and the low-level workings of the mapping software. For example, Musician C desires detailed control over the posture training process, such as the ability to remove sensors from the algorithm. In their current practice, Musician C "frees" sensors from the algorithm by providing it with enough varied training examples so that classification result becomes unaffected by the position certain fingers (in this case the thumb).

> "Almost any posture, if I program the postures in a way that the thumb is independent and doesn't add to the posture, I can move the thumb around." – Musician C. "I want more degrees of freedom. I want to

> be able to choose for myself which sensors are the ones that are contributing the postures that can be used as triggers … if you could say I don't want that to be in the posture recognition algorithm, because if I move that particular sensor, usually it's my thumbs, if I move that sensor I don't want that to mess up my postures, I want to be able to really use my fingers a lot more, in a more nuanced way." – Musician C.

## 8.4. Discussion

These findings suggest that glove performers achieve expressive performances using simple one-to-one and few-to-one mappings to minimise the risk of performance error, while embellishing these simple mappings with theatrical ancillary gestures. The musicians use of simple mappings suggests that it is not the complexity of their mappings that facilitates their musical expression, going against the argument that simple mappings lead to musical toys that musicians quickly grow tired of (Hunt, Wanderley and Paradis, 2003). For these musicians, expressive performance can be achieved through simple action–sound mappings that facilitate theatrical movement and their personal ideas and aesthetics of action–sound relationships.

However, the musicians' ability to express their personal aesthetics is hindered by issues with accidental triggering caused by the snap-to-nearest behaviour of the posture recogniser, with the musicians' mapping decisions being influenced by the need to consider the robustness of their mapping choices.

An interesting finding in this research is the importance of a musicians' personal mapping strategies. While it might be expected that mapping practice would be similar between musicians due to their frequent collaborations and sharing of ideas, this research has found that mapping practice is an incredibly personal endeavour, with glove musicians considering mapping design to be an important aspect of their creative practice. Providing musicians with the ability to define their own ideas around action–sound mapping forms an important part of musical expression with the gloves.

Another point of interest is the desire of some of the musicians to work with choreographers to aid them in developing visually sophisticated and expressive movements. This focus on expressive movements was influenced by the musicians' aim to provide engaging performances for their audiences, and highlights the importance of the performer-audience relationship in mapping design; both the minimisation of performer errors and the use of expressive ancillary gesture come from the desire to provide a good performance, and such factors are more important for musicians in the context of live performance than in other domains, such as composition (Fiebrink *et al.*, 2010). This is particularly highlighted by the personal glove practice of Musician C, who does not have the same error minimisation priority of the other musicians, being more willing to spend time mastering difficult mappings in their personal practice, while in their mapping design for other performers, simplicity and audience engagement remain important factors. New Interface for Musical Expression (NIME) mapping design literature advocates for the use of metaphors in mapping design to facilitate both musician and audience engagement (Hunt, Wanderley and Paradis, 2003; Fels, Gadd and Mulder, 2002). The Glove musicians used metaphors in their mapping design, mainly using metaphors in their ancillary gestures to communicate meaning, for example musician A's "opening eyes" metaphor. This again was mainly influenced by wanting to provide engaging performances. The focus on end-user mapping in the context of professional performance reveals mapping design influences that might not be apparent in laboratory-based studies (Caramiaux *et al.*, 2014a; Françoise, 2013), which find the importance of designing mappings with embodied metaphors, but do not touch upon factors raised in this study around reliability, an important aspect of mapping design for these musicians as they are aware that their mappings are being designed for a live performance. Chapter 9 examines this aspect in detail.

The aspects that influence expert end-user mapping design correlate with aspects of the Cognitive Dimensions of Music Notations (Nash, 2015): The use of simple action–sound mappings relates to *Visibility*, the clarity of how music is visualised in notations; expressive ancillary movement relates to *Role Expressiveness*, the visual aesthetic properties of the notation; the use of metaphor relates to *Closeness of Mapping*, the way in which the representation of the music aligns with how it is described; and accidental triggering relates to *Error Proneness*, or the ease at which unintentional mistakes can be made.

## 8.5. Summary

Through investigating the mapping practice of four experienced glove musicians, the work presented in this chapter has revealed a series of factors that influence experienced end-user mapping design for music performance. Primarily, the musicians focus on creating simple mappings that reduce the possibility of performer error, focusing on developing expressive, performative ancillary movement, with the underlying aim of these factors being the desire to provide engaging performances for their audiences. By studying the creative mapping practice of experienced mid-air musicians, an understudied group of DMI practitioners (McPherson and Kim, 2012), this research contributes novel insights into expressive DMI mapping design. The findings from this research can be summarised as follows.

1. Experienced musicians can use simple mapping solutions while providing engaging performances through expressive movement.
2. It is important for DMI mappings to reflect a musician's personal interpretations and aesthetics of music and movement, and that the mapping design process is an important part of creativity with the gloves. Therefore, DMIs instruments that permit end-user customisation and mapping personalisation empower musicians to express their own personalised action–sound relationships.
3. Experienced musicians use metaphors in their mapping design to communicate aesthetic intent and to provide engaging performances. The use of metaphor has been advocated in the design of music interaction (Fels, Gadd and Mulder, 2002; Wessel and Wright, 2002), which this research supports.
4. A major barrier in musical practice with the gloves is accidental triggering. The behaviour of the posture recogniser and occurrence of "pass-through" postures force the musicians to move away from expressing their personal aesthetics and focus on the "robustness" of their posture choices. DMIs could minimise accidental triggering through the use of excitation controls that avoid the need for performers to pass through other body states that trigger excitation controls, and through avoiding controls that rely on gestures that are similar to a musician's relaxed body state.

The findings from Chapters 6, 7 and 8 contribute towards a series of design heuristics for mid-air DMIs, presented in Chapter 10.

Following on from findings from this chapter and the previous chapter (Chapter 7), which found that system-related error and accidental triggering causes serious issues for musical practice with the gloves, the next chapter (Chapter 9) will examine the effects of system reliability on a musician's ability to acquire skill with the gloves.

# 9. Effect of System Error on Performer Skill Acquisition

In the previous empirical chapters (Chapters 6–8), the mapping practice of glove musicians was examined to determine the factors that influence end-user mapping design for music performance. Aspects that routinely affected both novice and expert glove musicians were problems with accidental triggering and system-related misclassification errors from the user-trained posture recognition system in the Glover software. These errors caused frustration, and disrupted novice users' development of understanding around the gloves system, as well as their ability to implement mapping ideas in Glover; while the expert musicians had developed a series of techniques to cope with the issue. In this chapter, system-related error is examined to determine its affects on a musician's ability to acquire musical skill with the gloves. These issues of system reliability, in part caused by Machine Learning (ML), are applicable to many Digital Musical Instruments (DMIs), as ML and classification techniques are common tools in the mapping design process (Macionis and Kapur, 2018; Scurto, Bevilacqua and Françoise, 2017; Fiebrink, Trueman and Cook, 2009).

## 9.1. Introduction

Three factors of musical skill acquisition: movement smoothness (González-Sánchez *et al.*, 2019; Goebl and Palmer, 2013; Palmer *et al.*, 2007), timing error and note error; were measured as deliberate system error was introduced into the gloves system as participants repeatedly practised a musical task, with the system error designed to mimic classification errors. It is hypothesised that higher rates of system error will have greater negative effects on the participants' ability to acquire skill, and will also negatively affect a participant's confidence in the gloves' reliability.

## 9.2. Skill Acquisition

The evaluation of one's musical improvement, or the acquisition of musical skill, is a major part of traditional instrument pedagogy, and has been explored at length in studies examining musical ability (Duke, Cash and Allen, 2011; McPherson, 2005; Drake and Palmer, 2000; Hodges and Nolker, 1992). In traditional pedagogical examinations (ABRSM, 2019, for example), musical ability is measured qualitatively: an experienced instrumentalist observes the examined student perform pieces and technical exercises, and grades them according to a series of specifications.

While for many contexts qualitative measurements are appropriate, it is often desirable to quantify improvements in musical ability to enable statistical comparisons in dependent variables. One method used in skill acquisition studies is to measure the speed at which a performer is able to perform a specified piece (Goebl and Palmer, 2013). However, such a measurement can be inappropriate for musical tasks, as tempo is integral to musical expression (Juslin, 2003).

In these cases, useful metrics include measuring a performer's deviation from the musical metre, or their musical timing accuracy (Duke, Cash and Allen, 2011; Drake and Palmer, 2000); and measuring a performer's pitch accuracy (Drake and Palmer, 2000), which is often used for singing analysis (Nakano, Goto and Hiraga, 2006; Watts, Murphy and Barnes-Burroughs, 2003). Another is to measure the improvement in the efficiency of a performer's movements, often defined as an improvement in *movement smoothness* (Caramiaux *et al.*, 2018; Palmer *et al.*, 2007; Drake and Palmer, 2000), a measurement determined from jerk, the third derivation of motion.

Movement smoothness has been used to study improvements in movement in many contexts, including sports (Choi *et al.*, 2014), medical rehabilitation (Rohrer *et al.*, 2002) and music performance (Duke, Cash and Allen, 2011; Palmer *et al.*, 2007). Movement smoothness is tied to the idea that efficiency is important to skilled movement; the less extraneous motion that exists in a movement (and thus a smaller jerk-cost), the more skilful the performer of that movement is. A criticism of movement smoothness is that efficiency in movement can be linked to Taylorist movement economics (Gertler, 2009), which aimed to find the most efficient ways of moving to maximise production, and leaves little room for an individual's personal expression in movement. However, developing efficiency in movement is integral to developing skill in instrument performance, with more efficient movements allowing experienced musicians to perform faster with fewer inaccuracies (Goebl and Palmer, 2013), and in the most prominent creative movement domain: dance (Lepecki, 2006).

## 9.3. Perception of Error

An initial study was conducted to determine at what rate of system error impacts the performer's subjective perceptions of the instrument. Six conditions of system error rate were examined: 0%, 1%, 2%, 5%, 10% and 20%.

### 9.3.1. Method

Each participant was asked to perform a short musical task (described in Section 9.4) 20 times in each of the system error conditions. The participants were exposed to each condition in a random order. After each condition, the participants were asked to respond to the following questions:

1. On a scale from 0 – 10, where 0 is not responsive at all and 10 is completely responsive, how responsive would you say the gloves were to your actions?
2. On a scale from 0 – 10, where 0 is no control at all and 10 is complete control, how much control did you feel you had?
3. On a scale from 0 – 10, where 0 is not at all accurate and 10 is completely accurate, how accurately did you feel the gloves were responding to your actions?

### 9.3.2. Results

Five participants took part in the study. The movement smoothness and note selection error were also measured, and presented in Figure 9.2.

Regarding the user perception responses (Figure 9.1), there was a trend in all three aspects of perception towards more positive responses with smaller amounts of system error, with a significant drop-off in perceived accuracy, responsiveness and sense of control once any system error had been introduced.

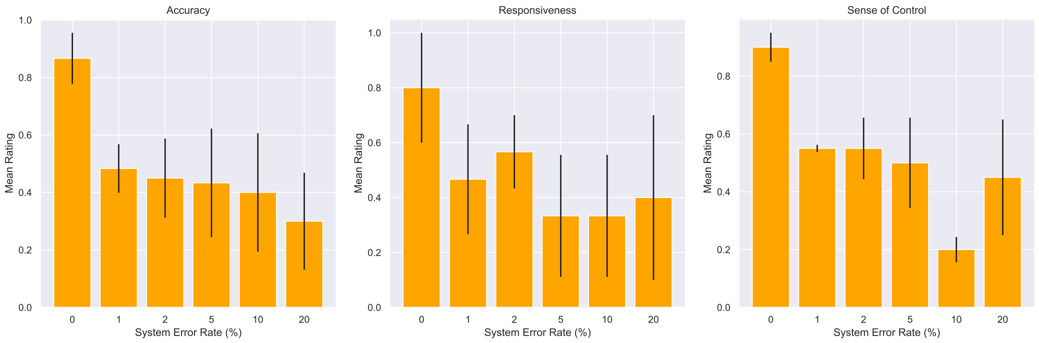


Figure 9.1.: User perception results.

| | 0% | 1% | 2% | 5% | 10% | 20% |
|---|---|---|---|---|---|---|
| 0% | 1 | 0.708 | 0.667 | 0.0788 | 0.0782 | *0.0312 |
| 1% | 0.708 | 1 | 0.93 | *0.0267 | *0.0156 | *0.00593 |
| 2% | 0.667 | 0.93 | 1 | *0.0284 | *0.0182 | *0.00711 |
| 5% | 0.0788 | *0.0267 | *0.0284 | 1 | 0.69 | 0.891 |
| 10% | 0.0782 | *0.0156 | *0.0182 | 0.69 | 1 | 0.526 |
| 20% | *0.0312 | *0.00593 | *0.00711 | 0.891 | 0.526 | 1 |

Table 9.1.: P-values for two-tailed t-tests comparing performer error in the six system error conditions. * indicates a statistically significant difference.

Movement smoothness scores were median-normalised by scaling each participant's scores by their median score across all conditions. A single factor ANOVA applied to the movement smoothness scores revealed that there was no statistically significant difference $(p < 0.05)$ on movement smoothness in the six conditions of system error rate: $F(1989, 5) = 0.89, p = 0.49$.

However, there was a statistically significant difference $(p < 0.05)$ in the rate of performer note selection errors in the six conditions of system error rate: $F(5, 230) = 3.21, p = 0.008$. A matrix of two-tailed t-tests shows that the significant difference exists between lower error-rate conditions of 1% and 2% and the higher error-rate conditions of 5%, 10% and 20%.

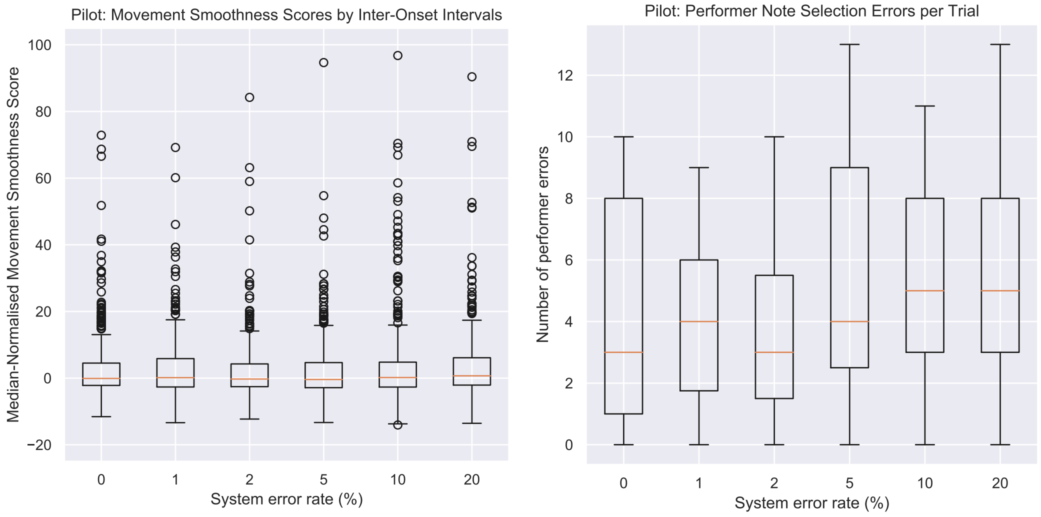


Figure 9.2.: Movement smoothness and note selection error results for perception study.

## 9.4. Method

To examine the effects of system error on the development of proprioceptive skill, intentional system error was introduced at three levels into the glove's signal chain: a control, where no deliberate error was introduced into the system, and two conditions where error was introduced: one at 5% (one in every twenty notes the performer played would be incorrect) and one at 10% (one in every ten notes incorrect). These conditions will be referred to as NE (no introduced system error), LE (low amount of system error: 5%), and HE (high amount of system error: 10%). These choices were influenced by the perception results, which suggested that performer note error was effected by these rates (Table 9.1).

### 9.4.1. Task

Participants were asked to perform a short piece of music (Figure 9.3) in time to a metronome, repeated 50 times. For the first four times round, the reference melody was played alongside the metronome. No participants that took part in

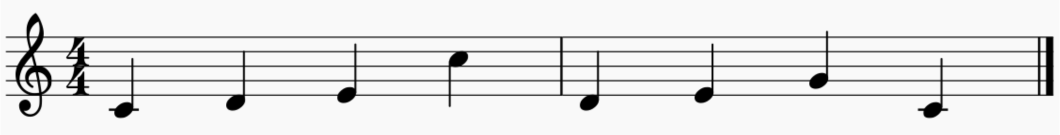



Figure 9.3.: The musical score for the task.

the study described in Section 9.3 were involved. An entirely new cohort of participants were recruited.

As the participants were being asked to perform a repetitive task 50 times, concerns were raised around problems with fatigue. For this reason, each participant only performed one condition. To mitigate the potential for one condition having participants with more musical ability than another condition, before participating in the experiment the participants completed a Musical Sophistication Index (MSI) questionnaire (Müllensiefen *et al.*, 2014). The MSI scores were then used to sort participants into each condition, so that each group had a similar spread of musical ability.

### 9.4.2. Mapping Strategy

The mapping strategy for piece used a "point and grab" interaction metaphor, and consisted of five directional lobes (point): up, down, left, right and forwards, each mapped to a note in the task's musical material (Figure 9.3).

To avoid adding additional system error, the machine learning part of the Glover software, the posture recognition system, was not used in this mapping. Instead, the "grab" element of the mapping was mapped to a threshold in the middle finger proximal flex sensor.

The mapping only used one glove, with all participants required to perform the task using the right-hand glove. The mapping strategy was kept simple, and its design was influenced by findings from a previous chapter (Chapter 7), which

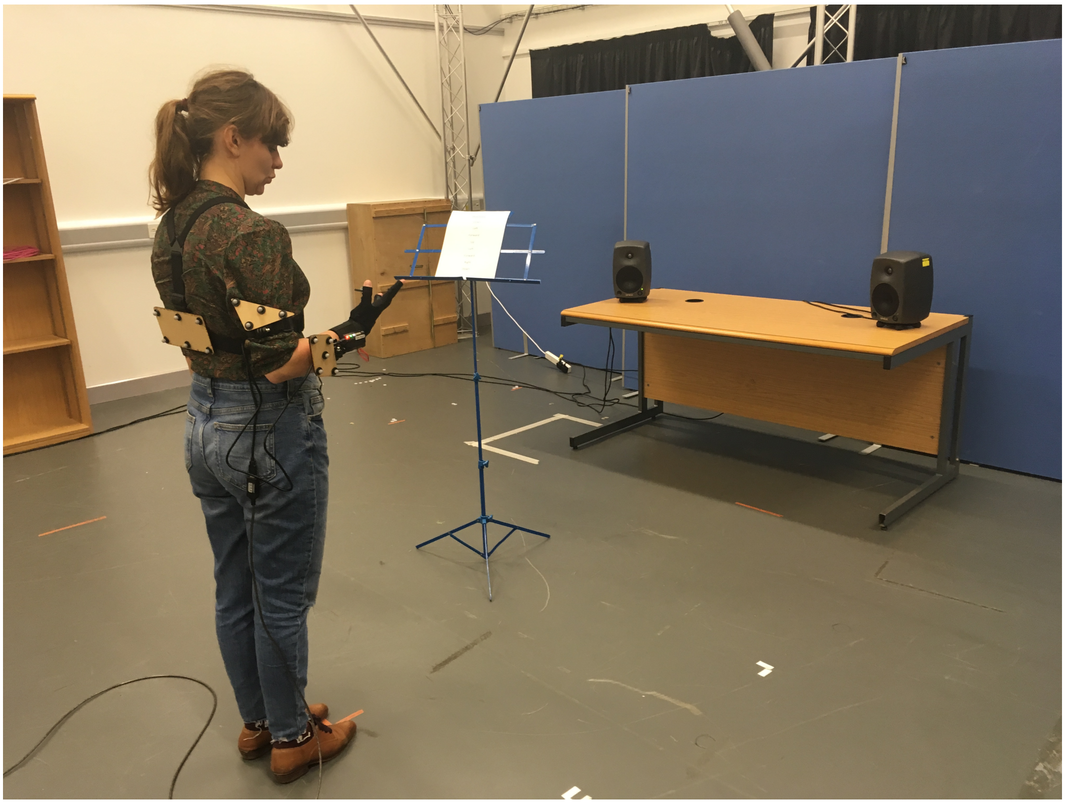


Figure 9.4.: Experimental Setup.

found that the use of directional lobes were a commonly used method of performing similar tasks.

### 9.4.3. System Error

The system error introduced into the gloves system was designed to mimic the misclassification errors regularly experienced by glove musicians, and also inherent in machine learning systems in general. At a probability determined by each conditions level of system error (0%, 5% and 10%), as a participant triggered a note, the intended pitch was altered to a pitch randomly selected from the other pitches present in the target piece (Figure 9.3).

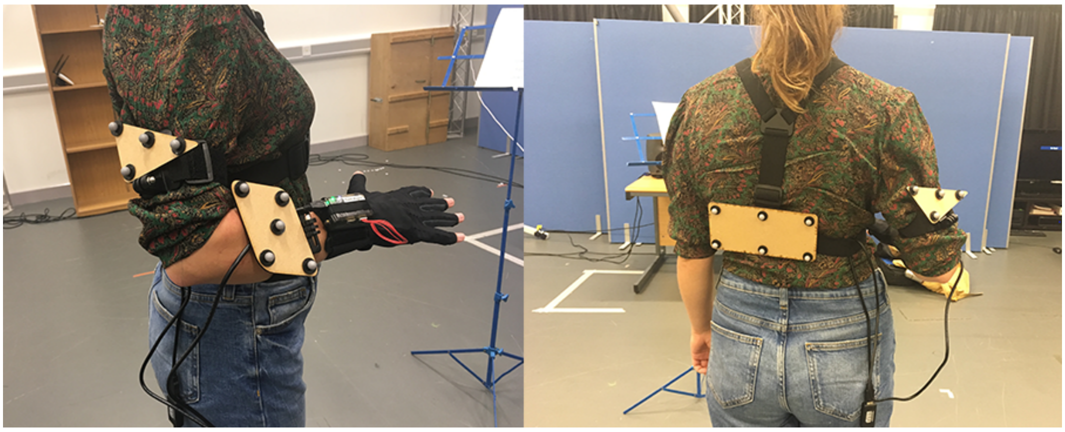


Figure 9.5.: Motion capture marker placement.

### 9.4.4. Experiment Setup

Figure 9.4 shows the experimental setup. Participants were asked to put on the glove and with motion capture markers, which were placed on the participant's back, lower arm and upper arm (Figure 9.5). The task and glove mapping strategy were explained to the participants, and they were then given a few minutes to familiarise themselves with the mapping strategy. Participants were provided with a score, denoting order of the point-and-grab directions required to play the piece. A pair of monitors provided auditory feedback, with the synthesiser controlled by the participant panned to the left monitor and the metronome and reference track panned to the right.

### 9.4.5. Data Collected

Movement data was recorded using a Vicon T40s motion capture system. Three markers were used to record the positions of the participants' back, upper arm and lower arm. Secondary back-up movement data was recorded from the accelerometers and flex sensors in the glove.

Music-related data was collected in MIDI format. This was made up of the notes performed by the participant and the system's MIDI output. This allowed for the

| Timestamp (relative) | Label | MIDI message output |
|---|---|---|
| -132 | SYSTEM_CORRECT | Note on C3 |
| -132 | USER_PERFORMED | Note on C3 |
| 0 | REFERENCE | Note on C3 |
| 1000 | REFERENCE | Note on D3 |
| 1011 | SYSTEM_CORRECT | Note on E3 |
| 1011 | USER_PERFORMED | Note on E3* |
| 1958 | SYSTEM_CORRECT | Note on E3 |
| 1958 | USER_PERFORMED | Note on E3 |
| 2000 | REFERENCE | Note on E3 |
| 2991 | SYSTEM_ERROR | Note on G3** |
| 2991 | USER_PERFORMED | Note on C4 |
| 3000 | REFERENCE | Note on C4 |
| 4000 | REFERENCE | Note on D3 |
| 4053 | SYSTEM_CORRECT | Note on G3 |
| 4053 | USER_PERFORMED | Note on G3* |
| 5000 | REFERENCE | Note on E3 |
| 5016 | SYSTEM_CORRECT | Note on E3 |
| 5016 | USER_PERFORMED | Note on E3 |
| 5919 | SYSTEM_CORRECT | Note on G3 |
| 5919 | USER_PERFORMED | Note on G3 |
| 6000 | REFERENCE | Note on G3 |
| 7000 | REFERENCE | Note on C3 |
| 7101 | SYSTEM_ERROR | Note on D3* |
| 7101 | USER_PERFORMED | Note on E3** |

Table 9.2.: Example of one trial of MIDI data. *Examples of performer error. **Examples of system error.

separation of performer-based errors from system-based errors, and the examination of the effects of system error on performer error. For example, if the musician intended to play C3, but the system decided to make an error and output D3, both the performer's intended note and the system's outputted wrong note were recorded. These values were recorded alongside a reference performance, which represented a completely accurate rendition of the piece.

As well as quantitative data about the participant's movement performance during the task, qualitative measurements around the participant's subjective experience of the task were taken in the form of a series of Likert scale questions.

To more effectively determine each participant's general improvement over time, a moving average was calculated for each participant's movement smoothness, note selection accuracy and timing accuracy scores, averaging over the previous 10 trials. Additionally, as this research is interested in the relative improvement as opposed to absolute smoothness scores, the mean scores were offset by the initial average.

### 9.4.6. Movement Smoothness

Movement smoothness was calculated using the same method as previous work (Caramiaux *et al.*, 2018): an integrated squared jerk value was measured along each axis of movement for the inter-onset intervals between performed note onsets. The cumulative values for each inter-onset interval performed in a single trial gave each trial an overall movement smoothness score. Higher values represent a larger jerk-cost, and thus a less-smooth movement.

To allow for meaningful averages to be calculated for each condition, median normalisation was applied to each participant's movement smoothness scores.

### 9.4.7. Performer Note Selection Error

To examine participant note selection error, the user performed Note On information was compared against a reference of the correct rendition (see Table 9.2). For example, in the sample provided in Table 9.2, the participant performed three wrong notes: the second (E3 instead of D3), fifth (G3 instead of D3) and eighth (E3 instead of C3).

### 9.4.8. Performer Timing Error

Performer timing error was also examined. A timing error score was assigned to each performed trial by measuring the cumulative difference between the time stamp in the reference MIDI data and the nearest note onset performed by the participant (see Table 9.2).

## 9.5. Results

45 participants took part in the study, 14 in the NE condition, 16 in LE condition and 15 in the HE condition. The distribution of participant MSI scores are shown in Figure 9.6. A single factor ANOVA showed that there was no statistical difference between groups: $F(42, 2) = 0.29, p = 0.75$.

### 9.5.1. Movement Smoothness

Movement smoothness was calculated using movement data from the motion capture system, using the method described in Section 9.4.6. The movement smoothness results are presented in Figure 9.7. These results suggest that for the NE and HE conditions, movement smoothness steadily improved after the 20th trial, while movement smoothness improved for the NE condition faster than in the HE

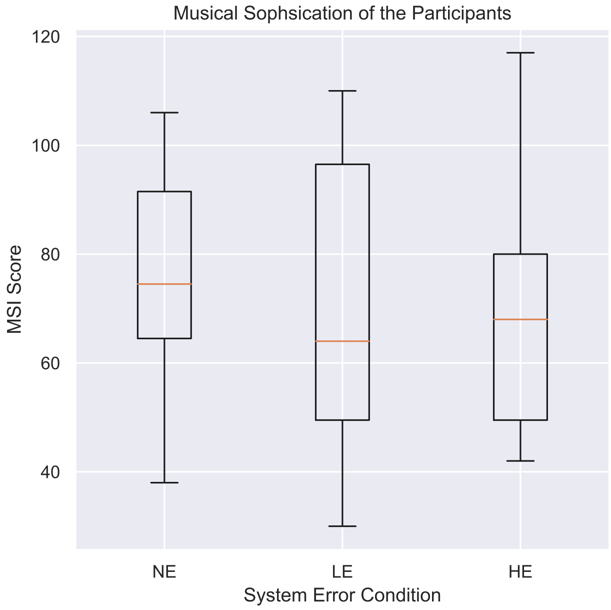


Figure 9.6.: Disribution of MSI scores for the three conditions.

Figure 9.7.: Change in Movement Smoothness scores over the course of the task. Lower scores represent smoother movements. Tails represent standard error.

Figure 9.8.: Note selection error rate over the course of the task. Lower scores indicate fewer performer errors. Tails represent standard error.

condition, NE smoothness levelled out around the 40th trial, and the HE condition continued to improve until the end of the task, equalling the improvement in movement smoothness achieved by the NE condition. Meanwhile, the LE condition results suggest little movement smoothness improvement over the course of the task, initially worsening until the 30th trial, then showing improvement at a similar gradient to the HE condition, before beginning to worsen again after the 40th trial.

### 9.5.2. Performer Note Selection Error

The performer note selection error results are presented in Figure 9.8, which suggests that the NE and LE trials initially showed equal levels of improvement until the 25th trial, when improvement in the LE condition plateaued. The results of the NE condition suggest continued improvement before plateauing around the 35th

Figure 9.9.: Performer timing error over the trials. Lower scores represent less performer error. Tails represent standard error.

trial. Performer note error for the HE condition improved over the course of the task, but with HE error rates initially being worse than both the NE and LE conditions. The HE condition began to show better error rates than the LE condition after the 30th trial, before worsening after the 45th.

### 9.5.3. Performer Timing Error

The performer timing error results are presented in Figure 9.9. The results suggest that for the NE condition, timing accuracy initially improved, then remained fairly stable throughout rest of the task. Timing accuracy for the LE condition started at a similar rate to the HE condition, but these results suggest a significant worsening after the 15th trial, before beginning to improve after the 30th trial, eventually becoming similarly accurate than the NE condition after the 40th trial.

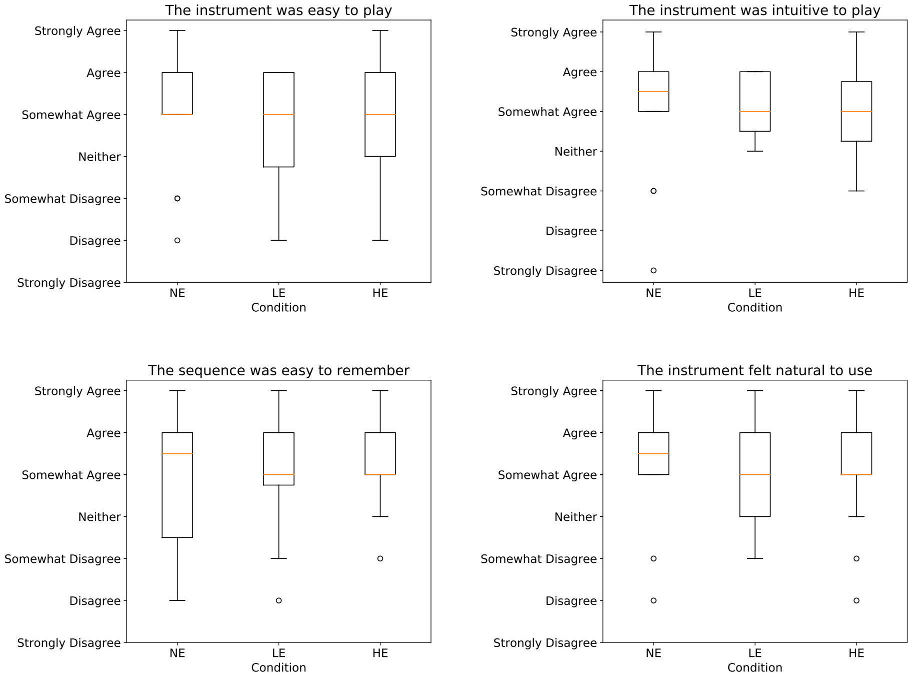


Figure 9.10.: Ease and Intuitiveness Likert Responses.

The HE condition results suggest an initial improvement, performing similarly to the NE condition, but became worse after the 40th trial.

## 9.5.4. User Perception

The user perception statements were grouped into four categories: Ease and Intuitiveness (Figure 9.10), Motivation and Engagement (Figure 9.11), Participant's Performance (Figure 9.12) and System Reliability (Figure 9.13). No statistically significant difference was found in any of the user perception statements (Table 9.3). Open-ended responses from participants were also examined for each of the three conditions.

### No Error Comments

The participants in the NE condition discussed how they felt fatigued after performing the task, and that the length and manner of the task made focusing on

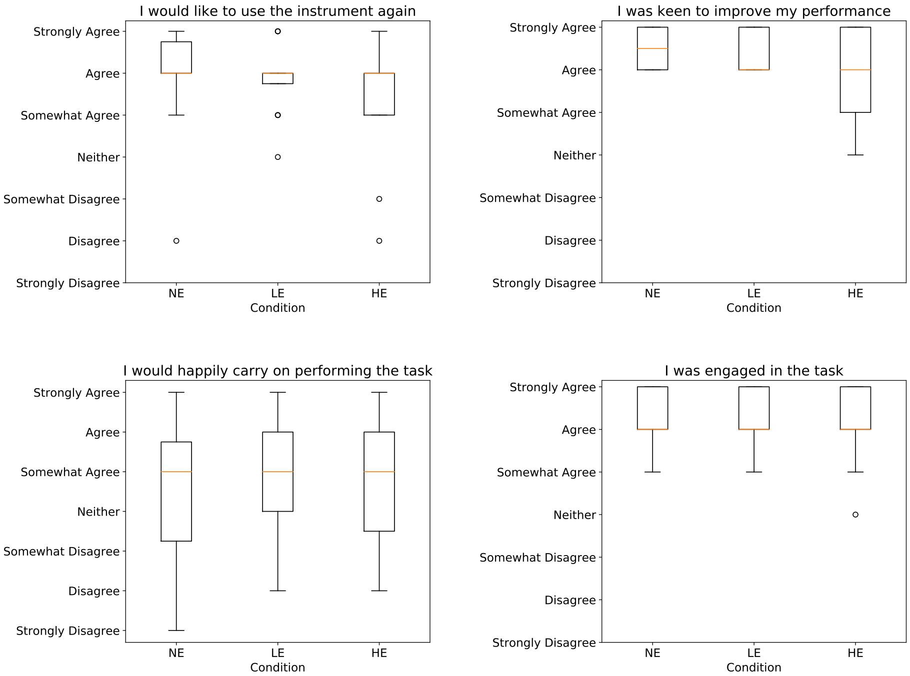


Figure 9.11.: Motivation and Engagement Likert Responses.

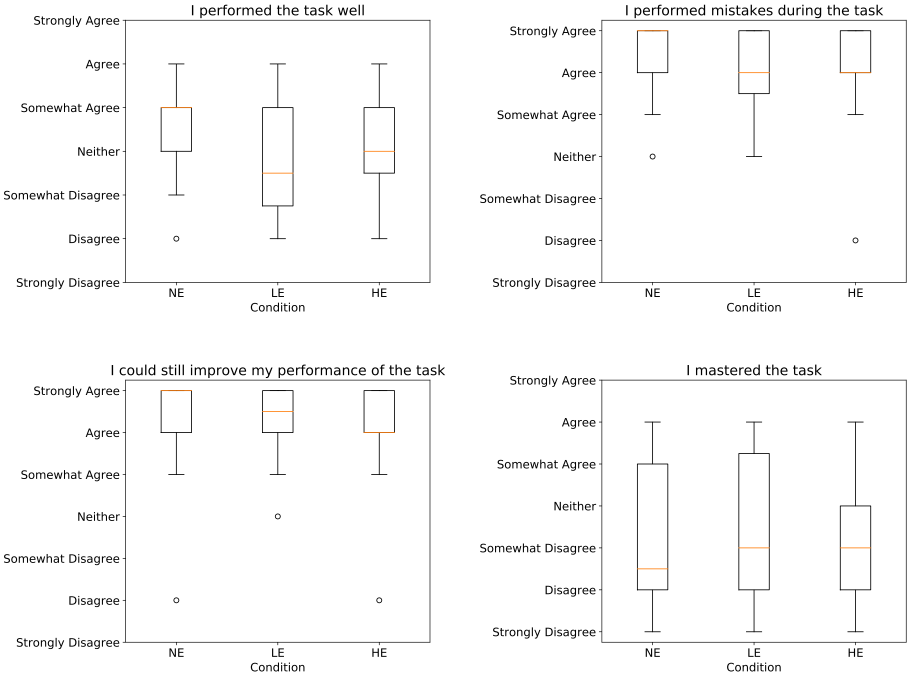


Figure 9.12.: Participant Performance Likert Responses.

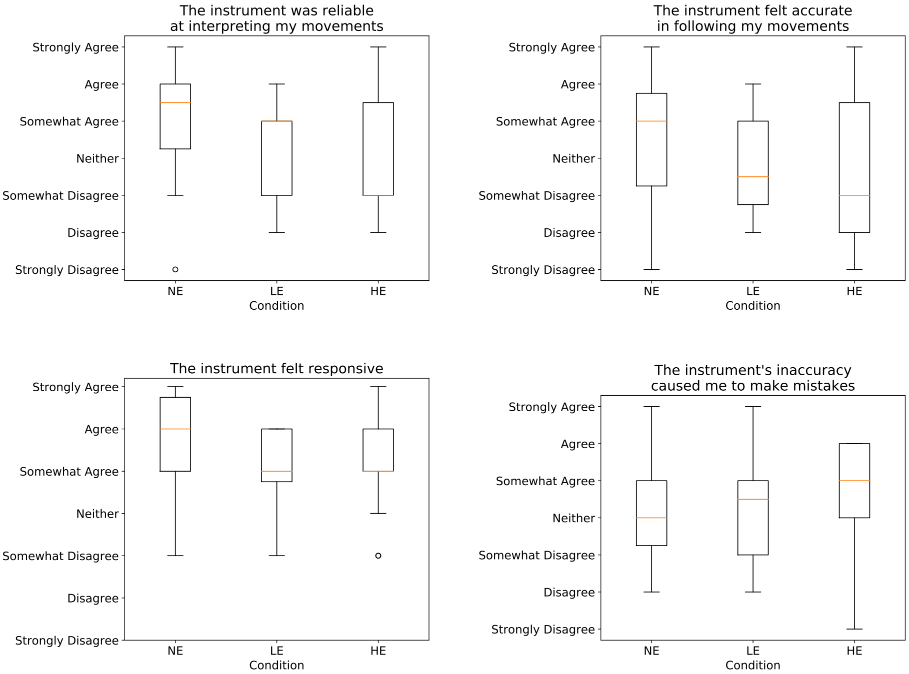


Figure 9.13.: System Reliability Likert Responses.

| | | **Median Responses** | | | **ANOVA Results** |
|---|---|---|---|---|---|
| **Category** | **Statements** | No Error | Low Error | High Error | (significance: $p < 0.05$) |
| Ease and Intuitiveness | The instrument was easy to play | somewhat agree | somewhat agree | somewhat agree | $F(2,42) = 0.208,\ p = 0.813$ |
| | The instrument was intuitive to play | somewhat agree | somewhat agree | somewhat agree | $F(2,42) = 0.039,\ p = 0.96$ |
| | The instrument was frustrating to play | neither | neither | neither | $F(2,42) = 0.041,\ p = 0.959$ |
| | The sequence was easy to remember | somewhat agree | somewhat agree | somewhat agree | $F(2,42) = 0.665,\ p = 0.52$ |
| | The instrument felt natural to use | somewhat agree | somewhat agree | somewhat agree | $F(2,42) = 0.214,\ p = 0.808$ |
| Motivation and Engagement | I would like to use the instrument again | agree | agree | agree | $F(2,42) = 0.588,\ p = 0.559$ |
| | I was keen to improve my performance | agree | agree | agree | $F(2,42) = 2.878,\ p = 0.0674$ |
| | I would happily carry on performing the task | somewhat agree | somewhat agree | somewhat agree | $F(2,42) = 0.101,\ p = 0.903$ |
| | I was engaged in the task | agree | agree | agree | $F(2,42) = 0.288,\ p = 0.751$ |
| Participant Performance | I performed the task well | agree | agree | agree | $F(2,42) = 1.669,\ p = 0.201$ |
| | I performed mistakes during the task | strongly agree | agree | agree | $F(2,42) = 1.193,\ p = 0.825$ |
| | I recovered from my mistakes well | somewhat agree | neither | somewhat agree | $F(2,42) = 1.202,\ p = 0.31$ |
| | I could still improve my performance of the task | strongly agree | strongly agree | agree | $F(2,42) = 0.044,\ p = 0.957$ |
| | I fully mastered the task | disagree | somewhat disagree | somewhat disagree | $F(2,42) = 0.234,\ p = 0.792$ |
| System Reliability | The instrument was reliably interpreting my movements | somewhat agree | somewhat agree | somewhat disagree | $F(2,42) = 1.647,\ p = 0.205$ |
| | The instrument was accurately following my movements | somewhat agree | somewhat disagree | somewhat disagree | $F(2,42) = 0.916,\ p = 0.408$ |
| | The instrument felt responsive | agree | somewhat agree | somewhat agree | $F(2,42) = 1.158,\ p = 0.324$ |
| | The instrument's inaccuracies caused me to make mistakes | neither | somewhat agree | somewhat agree | $F(2,42) = 0.021,\ p = 0.979$ |
| | I recovered from the instrument's mistakes well | neither | neither | somewhat agree | $F(2,42) = 0.337,\ p = 0.716$ |

Table 9.3.: User perception Likert responses.

the task difficult. Some participants discussed how it took them time to develop an understanding of the control mapping, while one participant discussed their lack of being able to perceive where errors were coming from.

> "After a while of using the gloves this way, my arm became tired."
>
> "My forearm and shoulder are quite sore."
>
> "It was hard to remain interested in just repeating the same tune 50 times."
>
> "The main problem I was making was with the last down note as I was bring my arm down with my wrist pointing slightly up, once I figured out I had to be aiming towards the floor, it made it a lot easier."
>
> "I couldn't work out if I was doing something different or making mistakes or if this was an error in the motion capture - but it was enough to make me grimace briefly!"

### Low Error Comments

The LE condition participants discussed the gloves inaccuracies, with some participants unable to perceive where the error was coming from, and that the system error was causing them to lose confidence in their own ability. One participant commented that they initially felt that the error came from them, they realised it was system-related towards the end of the task. The LE participants also discussed feeling fatigued.

> "I couldn't work out whether it was me or the gloves making mistakes."
>
> "The inaccuracy of the gloves made me doubt my own competence and made it more difficult to remember the sequence."
>
> "I enjoyed the session. However in the end I got tired a little bit."

### High Error Comments

The majority of the HE participants' comments were related to their perception of system error. Many of the participants reported perceiving the system as being the cause of error, while some were unsure about the origin of error.

> "It felt as if the notes were being changed. And that some positions were removed."
>
> "I wasn't sure whether it was me making errors."
>
> "It seemed unreliable at times, while the left and up directions worked quite well, the right forward and down directions were rarely interpreted right."

## 9.6. Discussion

The purpose of this study was to examine the effects of system error on a performer's ability to acquire skill with the Mi.Mu Gloves. Movement smoothness, note selection accuracy and timing accuracy were analysed, as well as the perception and experience of the participants.

In both movement smoothness and timing accuracy, the low error condition produced the worst performances. This is a striking result, and requires further examination and replication in future studies before it can be assumed to be true in all cases. However, these results could be attributed to the ability of the participants to perceive that the source of the error was coming from the system, as discussed in the upcoming subsections.

The findings have implications on the wider DMI field, where ML tools are widely used in DMI design (Macionis and Kapur, 2018; Vogl and Knees, 2017; Fiebrink and Cook, 2010).

### 9.6.1. Skill Acquisition

For movement smoothness, a 10% rate of misclassification style system errors have little effect on movement smoothness compared to the control condition. This suggests that the participants were realising that the errors they were experiencing originated from the system and not their own actions, and began to adapt to the condition. A 5% system error rate has a significant effect, with participants unable to improve their movement smoothness effectively over the course of the trials. These results suggest that participants in the 10% condition were able to perceive that the system errors were originating from the system, and were able to adapt, although their movement smoothness did not improve as early as the control condition.

The results for note selection error show that while error decreased across all conditions, a 10% rate of system error had no significant effect on the improvement of a performer's note selection errors over the control condition. Meanwhile, a 5% rate of error caused a significant effect, with participants unable to improve their note selection accuracy beyond the 25th trial. These results support observations of the movement smoothness results, suggesting that the high error condition was more easily perceivable, allowing the participants to adjust and adapt, while the low error condition caused disruption to skill acquisition due to participants being less able to perceive the source of error.

Regarding performer timing error, no improvement was observed in the control condition. Meanwhile, a 5% rate of system error was found to initially have a significant negative effect, but improved significantly during the latter half of the task. The 10% system error rate initially performed well, with participant's performing with less timing error than the control condition, but the condition rapidly deteriorated towards the end of the task. This suggests that fatigue or frustration could have played a role in the rate of performer timing error: to maintain low levels of performer note selection error and smoother movement while the system was making significant amounts (a 10% rate) of pitch-related error would require more concentration than for the control condition. These results suggest that in such conditions, timing accuracy was sacrificed to maintain note selection accuracy. Contrastingly, for the 5% condition, where an improvement in timing accuracy was observed around the same point as pitch accuracy worsened (around the 25th trial), pitch accuracy was sacrificed for timing accuracy.

### 9.6.2. User Perception

The user perception study found that even 1% of system error had a negative impact on a musician's sense of control, system accuracy and responsiveness. The qualitative responses from the main study found no significant perceptual difference between the three conditions. However, as each participant only performed the task in a single condition, direct comparisons between each group's perceptions are more difficult to make, and the more useful user perception results can be drawn from the perception study, where each participant performed in each condition. However, having each participant only perform in one condition for the main skill acquisition study was an appropriate choice, with participants reporting on the physical strain of the task, a common problem in mid-air interactions (Hincapié-Ramos *et al.*, 2014). Requiring each participant to perform 150 trials (3 conditions 50 times each) would have been overly strenuous.

## 9.7. Summary

In this chapter, the effect of system error on a performer's ability to acquire skill in performing with a mid-air DMI was examined. A performer's movement smoothness, note selection and timing error were measured as they performed a musical task with the gloves, while randomised system error, designed to mimic misclassification errors, was introduced into the glove's mapping chain. The findings show that lower system error rates negatively effect a musician's ability to acquire skill with a mid-air glove musical instrument, with motor skill, note and timing accuracy all disrupted. Larger rates of system error had little effect on skill acquisition, suggesting that participants were more able to perceive the system errors and adapt accordingly.

This chapter concludes the empirical research presented in this thesis. The following chapter presents a discussion and summation of the contributions made within this dissertation.

# 10. Discussion

## 10.1. Introduction

This dissertation presents work that investigates the end-user mapping design practices of musicians using a glove-based mid-air Digital Musical Instrument (DMI). The work presented looks at end-user mapping in the musical performance practice of novice and experienced glove musicians, and investigates the development of mapping design practice and motor skill. The qualitative and quantitative research methods used have been drawn from established User Experience (UX) methodologies, as well as recent work in the study of skill acquisition in the New Interface for Musical Expression (NIME) field. The research question this dissertation investigated was as follows:

*How do end-users of a glove-based music controller design action–sound mapping strategies for musical performance?*

To address this question, the following sub-questions were explored.

1. What factors influence mapping design decisions for glove-based music performance?
2. How does mapping practice develop as glove musicians gain experience?
3. How do these factors impact a musician's ability to acquire motor skill with glove instruments?

## 10.2. Overview of Thesis

Chapter 2 presented literature relating to mid-air interaction, mapping in DMIs, and end-user mapping design. The chapter covered implicit and explicit mapping techniques, the nature of action–sound mapping parameters and the qualities of mapping design, as well as existing end-user mapping software.

Chapter 3 presented literature relating to gesture and embodiment in music and music interaction, with a focus on conceptual metaphor theory. Semiotic and ecological theories are also covered.

Chapter 4 presented a systematic review of NIME literature from the The International Conference on New Interfaces for Musical Expression (NIME), The Sound and Music Computing Conference (SMC) and The International Computer Music Conference (ICMC), which analysed their relation to UX evaluation themes. Literature around the methodologies of Usability and UX design related to the NIME field is also covered.

Chapter 5 presented the Mi.Mu Gloves and Glover mapping software, the DMI used for the empirical work presented, in detail. The technical specification of the gloves and the software's mapping affordances were discussed

Chapter 6 presented a study that examined mapping design with a group of existing Mi.Mu Glove users. A mapping task was given to a group of glove users, which found that glove users with more experience performing live with the gloves prioritised ergonomic considerations and minimising performance mistakes, while those who did not perform live with the gloves used mappings that adhered to established embodied musical metaphors. The more experienced performers also spent more of the task time editing and iterating over mapping designs than practising their mappings.

Following on from observations made in Chapter 6, Chapter 7 presented a study into the initial development of glove musicianship. Eleven musicians were given a pair of Mi.Mu Gloves, and their development of performance practice was tracked over a month-long period. This study found that novice glove musicians designed mappings through the embodied manner of "what is it like to *do* the music?" designing mappings that use metaphors of established music interaction interfaces and instruments, and focus on their mappings' transparency to the audience. When designing for a specific task, the ergonomics and easiness of these control mappings were also an important factor in the novices' mapping designs, and their mapping design was hampered by system-related errors caused by posture recognition training issues, leading to accidental triggering.

Chapter 8 presented an empirical study study that examined the mapping practice of expert glove performers. Grounded Theory analysis of interviews with expert glove musicians revealed that the experienced glove musicians use simple mapping relationships in their glove performances, but embellish these mappings with performative ancillary movements that expressed their aesthetic intentions. Accidental triggering and system-related errors also impact their mapping designs, with the musicians developing gestural strategies to mitigate these issues.

Chapter 9 presented a study on the effect of accidental triggering on a musician's ability to acquire proprioceptive skill with the Mi.Mu Gloves. This study introduced system error into the glove mapping chain that mirrored the accidental triggering issues observed in previous chapters. This found that smaller rates of system error caused significant disruption to a musician's ability to acquire motor skill, and has a significant effect on note selection error and movement smoothness. Higher levels of system error had a significant effect on performer timing error, but no effect on motor skill acquisition or note selection error.

## 10.3. Main Findings

From this work, the following findings about end-user mapping design with a mid-air glove instrument have been made. A summary of the empirical findings is presented below in Table 10.1.

### What factors influence mapping design decisions for glove-based music performance?

#### Simple Mappings

This research has found that glove musicians make use of simple relationships in their mapping designs, and mainly use one-to-one and many-to-one mappings (Sections 7.4.1 and 8.3.1). This would seem to contradict established mapping literature which advocates for complex, many-to-many mapping relationships (Dobrian and Koppelman, 2006; Hunt, Wanderley and Paradis, 2003). However, the musician's conceptualisation of simple mapping is very much at the perceptual level of mapping, between what has been referred to in this work as gesture *features* (Section 2.5.4) and sound parameters. Thus the use of simple mappings can perhaps be explained by the nature of the input interface. In the case of a classic example that advocates for complex mapping, the "accidental theremin" (Hunt, Wanderley and Paradis, 2003), a systems level approach to mapping is taken, with input parameters being a set of one dimensional sliders mapped to specific parameters of synthesis. Meanwhile, for glove controllers, while the gestural parameter may be a single axis of orientation, the actual human hand and arm movements that go into actuating that parameter are incredibly complex.

This research has found that for glove controllers, where mapping is conceptualised at a perceptual level, mapping complexity does not need to be present in

**Mapping Practice of Existing Glove Users** (Chapter 6).

- Experienced glove musicians allocated more time to iterating over their mapping designs, displaying rapid feedback cycles of creating, auditioning and editing (Section 6.3.2).
- Novice glove musicians allocated less time to mapping design, spending more time practising with their mappings (Section 6.3.2).
- Experienced glove performers designed action–sound mappings that prioritised ergonomic control and minimisation of performer errors (Section 6.3.3).
- Novice glove performers designed action–sound mappings that adhered to established conceptual metaphors of music (Section 6.3.3).

**Development of Mapping Practice in Novice Glove Musicians** (Chapter 7).

- Novice glove musicians used simple one-to-one relationships in their personal mapping practice (Section 7.4.1).
- Novices tended towards using Posture and Movement gesture features in their mappings (Section 7.4.1).
- Novices designed mappings based on conceptual metaphors of existing instruments and music interactions (Section 7.4.3).
- Novices designed with the audience's perception of their mappings in mind (Section 7.4.3).
- Novices experienced issues with accidental triggering (Section 7.4.3).
- For mapping tasks, novices quickly developed an established mapping method (Section 7.5.2).
- For note-based mapping, novices focused on ergonomic control and error minimisation in their mapping designs (Section 7.5.4).
- For expression-based mapping, novices focused on visual aesthetics and metaphors (Section 7.5.4).

**Experienced Musical Practice with the Gloves** (Chapter 8).

- Expert glove musicians use simple mappings that minimise performer-related errors (Section 8.3.1).
- Experts embellish simple mappings with performative ancillary movement (Section 8.3.2).
- Experts designed mappings around novel interaction metaphors (Section 8.3.3).
- Experts experienced problems with accidental triggering, and had developed techniques to mitigate its effects (Section 8.3.4).

**Effect of System Error on Performer Skill Acquisition** (Chapter 9).

- A 1% rate of system-related error had a significant effect on a performer's perception of system accuracy, responsiveness and sense of control (Section 9.3).
- A 5% rate of system-related error resulted in a significantly reduced improvement in movement smoothness over time, while a 10% rate had little effect (Section 9.5.1).
- A 5% rate of system error resulted in a reduced improvement in note selection accuracy, while a 10% rate had little effect (Section 9.5.2).
- A 5% rate of system error caused timing accuracy to initially worsen before improving, while a 10% rate had little effect on timing accuracy initially, becoming significantly worse after time (Section 9.5.3).

Table 10.1.: Summary of empirical findings.

the mapping from gesture features to sound parameters to provide musicians with expressive control over sound.

### Conceptual Metaphors

This research as found that glove musicians make use of conceptual metaphors in their mapping designs (Sections 6.3.3, 7.4.3 and 8.3.3). Metaphors are used by glove musicians to create transparent (Fels, Gadd and Mulder, 2002) mapping strategies, with the glove musicians are concerned with the audience's ability to perceive their control relationships (Section 7.4.3). Novel metaphors were also used to express aesthetic intentions (Section 8.3.3). This finding supports established literature that advocates for the use of conceptual metaphors in DMI design (Wilkie, Holland and Mulholland, 2010; Fels, Gadd and Mulder, 2002).

> "I wanted the controls the gloves had to be obvious to the audience, and I wanted that to be the strongest feature of the performance." – Novice participant (Section 7.4.3).

### Maintaining Control

The glove musicians were concerned with maintaining control over their mappings and minimising performance mistakes (Section 6.3.3, 7.5.4 and 8.3.1). This was prevalent due to the problems the glove musicians had with accidental triggering, which influenced the musicians' mapping decisions and effected the musicians' ability to implement mappings that expressed their aesthetic intentions. The experienced musicians have developed techniques to minimise this, designing mappings that provide additional levels of control while being intentionally visibly imperceptible to audiences (Section 8.3.4).

> "[The problems are] mainly the postures interacting with each other, you can accidentally hit a posture on your way to another one." – Novice participant (Section 7.4.3).

> "I do this when I don't want people to pay attention because it's almost no posture, and I usually do this one when I kind of want to do it with an open hand but I just need the control." –- Expert musician (Section 8.3.4).

## How does mapping practice develop as glove musicians gain experience?

### Use of metaphor

Novice glove musicians used metaphor to show that their actions are controlling the auditory aspects of their performance, and made use of established embodied metaphors of music, such as UP AND DOWN IN SPACE IS LOUDER AND SOFTER DYNAMICS (Wilkie, Holland and Mulholland, 2010), or metaphors based on the interactions of established instruments, such as "air-guitars" or "air-violins", as well as typical interactions for audio effects, such as control knobs and sliders (Sections 6.3.3 and 7.4.3).

> "Maybe because I'm used to seeing it with knobs, but it made sense to have the filter cut off being roll." – Novice participant (Section 7.5.4).

Meanwhile, the experienced glove musicians used metaphor to express a set of personal aesthetics, and had developed a more nuanced approach to using metaphors in mapping design (Section 8.3.3). The experienced musicians use novel metaphors that add a visual aesthetic element to their performances, such as releasing notes like feathers (Section 8.3.3), and their mapping design aesthetics has become a key element of their performance practice and personal creative outputs (Section 8.3.6).

A major contrast between the novice and experienced glove performers was how the experienced musicians' incorporated their use of metaphor into expressive ancillary movements, that did not have any action–sound controlling aspect, but formed a major part of their performances.

> "I knew that it was just forward, up, left, right but it looked like 'I'm going to put it into that wind bit over there, and I'm going to let the feather go over there'." – Expert musician (Section 8.3.3).

#### Establishing personal technique

Novice glove musicians quickly established preferred mapping choices for specific mapping tasks, and the time allocated to sub tasks while performing mapping tasks changed very little over the month-long study period (Section 7.5.2).

For experienced glove musicians, mapping design has become a significant part of their creative expression, with sound-related mapping choices forming an integral part of their performance practice (Section 8.3.6). More technical aspects of mapping are shared between musicians (Section 6.3.3 and 8.3.6).

### How do these factors impact a musician's ability to acquire skill with glove instruments?

#### Strategies to mitigate accidental triggering

Issues with accidental triggering and system-related errors were the main detrimental factor in end-user mapping design (Sections 7.4.3, 7.5.4 and 8.3.4). These errors were caused through poorly trained posture recognition and musicians' lack of understanding of where those errors were originating from, as well as poorly designed mappings that leads to “pass through postures”, where a third posture is recognised as a musician moves between postures (Section 8.3.4). Issues with pass through postures and accidental triggering had an effect on the musician’s mapping choices, particularly with postures, with the experienced musicians developing a series of subtle postures that are not supposed to be noticed by the audience but provide additional levels of control for the musician (Section 8.3.4).

A more robust posture recognition algorithm, or clearer feedback and instruction within the Glover software on how to provide good training data, could have addressed these accidental triggering issues and impacted the Glover performers’ mapping choices, allowing them to focus less on ergonomics and instead make more aesthetically lead decisions. This is reflected in Section 7.5, where in an expression-based mapping task, with no need to perform the “correct” notes, novice Glove musicians focused primarily on aesthetics and musical metaphor in their designs.

#### Effects of system-related error

System-related errors caused disruption to motor skill acquisition when the system errors were present but relatively uncommon, and there was no significant effect on skill acquisition when errors were present and more common (Sections 9.5.1 and 9.5.2). This suggests that system error had a detrimental effect on skill acquisition when participants were less able to perceive its presence.

## 10.4. Methodological Review

### 10.4.1. Musical Tasks and Open Exploration

Task-based studies drawn from traditional Human-Computer Interaction (HCI) research has been a major influence on DMI research (Barbosa *et al.*, 2015; Wanderley and Orio, 2002), while more recent DMI work has looked to use qualitative, observational methods used in the field of UX research (Gelineck and Serafin, 2012; Johnston, 2011). The research undertaken in this dissertation took a mixed methods approach, using both task-based experiments (Chapters 6, 7 and 9) and observational qualitative approaches (Chapters 7 and 8).

When using task-based experiments to reveal mapping design decisions and behaviour, the mapping design of musicians was highly dependent on the musical material used in the task. For note-based tasks, musicians were designing mappings to limit mistakes and maintain control, with little thought for aesthetics (Sections 6.3.3 and 7.5.4), while for expression-based tasks, the participants were considering aesthetics and metaphor more, but were influenced by the cultural context of the material (Section 7.5.4). Meanwhile, the observational, qualitative methods that allowed participants to use the gloves within their own creative practice allowed for more contextually relevant mapping behaviour to emerge, and enabled more effective investigation into why musicians made their mapping designs choices (Section 7.4.1, 7.4.3 and 8.3).

### 10.4.2. Usage Logging

In Chapter 7, the number of participants in the longitudinal study meant that statistically significant results could not be drawn from the usage logging used to indirectly observe participant behaviour. It was however effective in revealing observational insights and supporting the qualitative observations and interviews (Section 7.4.1 and 7.4.3). Usage logging would be more effective with either a larger sample (Nash and Blackwell, 2011), or for a more long-term investigation into the development of a glove musician's musical practice.

## 10.5. Limitations

There are a few limitations to the work presented in this dissertation. One such limitation is the lack of participant numbers in Chapters 6, and 7 and 8. A lack of practitioners for DMIs is common (McPherson and Kim, 2012), with only a few pairs of usable gloves available to this research. Efforts have been made throughout this work to make the most out of the resources available through focusing on detailed qualitative and observational insights into glove musicianship (Chapters 7 and 8), supplemented by a large scale, quantitative study (Chapter 9).

It would have been useful to provide more time to participants in the longitudinal study presented in Chapter 7. The time given was based on similar previous research (Gelineck and Serafin, 2012). More time, for example six months instead of the allotted one month, would have allowed for more observation of a musician's development of practice, particularly regarding the usage logging, as a longer user journey would have been captured.

While studying mapping within the context of a musician's personal musical practice allows for observations to be properly contextualised, this leads to a loss of quantitative precision (Siegel, 2012). Attempts were made to mitigate for this by the use of task-based methods alongside each participant's open exploration in Chapter 7. However, the lack of participant numbers and that how much time each participant spent using the gloves between tasks varied according to their personal practice reduced the validity of these quantitative observations.

## 10.6. Recommendations

From the research conducted in this dissertation, the following recommendations are put forward for mapping design with mid-air DMIs.

## Use simple mappings between gesture features and sound parameters.

Glove musicians used mappings with simple, one-to-one relationships between gesture features and sound parameters, which they used to facilitate their expressive intentions (Sections 7.4.1 and 8.3.1). The use of simple mappings goes against traditional mapping literature which advocates for complex mapping strategies (Dobrian and Koppelman, 2006; Hunt, Wanderley and Paradis, 2003).

The musicians used simple mappings to increase control intimacy, system reliability and minimise their performance mistakes, and also to increase the visibility of their mapping choices to their audiences (Sections 6.3.3, 7.5.4 and 8.3.1). This focus on visibility reflects established thoughts around transparency (Fels, Gadd and Mulder, 2002) and visibility (Nash, 2015) in music interaction.

Referring back to Section 2.5.4, where the concepts of gesture features and sound parameters are discussed, it is important to address that simple mappings should be *perceptually* simple, and do not necessarily need to be a systems level one-to-one mapping. Many acoustic instruments have complex many-to-many mappings that are easily understood by performers and audiences.

## Make room for performative ancillary gesture

A significant aspect of artistic expression for glove musicians was how they utilised non-sound producing movements in their performances to express aesthetic intentions and visual conceptual metaphors (Section 8.3.2).

Recent literature has discussed the idea of bottlenecks in DMIs (Jack, Stockman and McPherson, 2017), where the multidimensional aspects of bodily gesture are reduced to a specific set of movements that the DMI is designed to detect. This research has found that mid-air DMI musicians utilise the space around the bottlenecks, where the musician's gestures would be ancillary instead of sound-producing (Jensenius *et al.*, 2009), to express visual aesthetics.

### Use metaphors in mappings to support intuitive interactions.

Both embodied metaphors and conceptual metaphors influenced and were used by Glove musicians in their mapping designs, and were used by both novice (Sections 6.3.3 and 7.4.3) and expert glove musicians (Section 8.3.3) to design engaging mappings for themselves and for their audiences. For example, the embodied metaphor of HIGHER IN SPACE IS HIGHER IN PITCH was used extensively by Glove musicians, which provided for intuitive mapping due to its ubiquity. Meanwhile, novel conceptual metaphors such as RELEASING FEATHERS IS RELEASING NOTES allowed musicians to express their own aesthetic ideas around movement and music.

Mid-air instrument designers should leverage both established embodied musical metaphors and novel musical conceptual metaphors in their mapping designs. This recommendation supports existing theories around the use of metaphor in interaction design (Wilkie, Holland and Mulholland, 2010; Fels, Gadd and Mulder, 2002; Marx, 1994)

### Balance visual aesthetics with ergonomic control.

How a mapping strategy was perceived by an audience was important to the glove musicians' mapping design practice, the glove musicians endeavoured to strike a balance between performing visually engaging movements while maintaining a sense of control over their mappings (Sections 6.3.3, 7.5.4, 8.3.1 and 8.3.2). As such, mid-air instrument designers should be conscious of providing ergonomic controls while also providing visually engaging mappings.

### Consider accidental triggering and minimise system error.

Related to ergonomic control, accidental triggering was a major problem for glove musicians, and significantly impacted a musician's ability to design mappings that expressed their aesthetic intentions (Section 8.3.4), develop understanding around the gloves system (Section 7.4.3) and gain proprioceptive motor skill with the gloves (Section 9.5.1). This would often be caused by controls that "passed through" each other, for example a musician moving from a fist posture to an open hand posture would "pass through" a one finger point posture (Section 8.3.5). Misclassification errors from poorly trained posture recognition models also caused issues for glove musicians, with the lack of reliability affecting the musician's confidence with the gloves (Sections 7.4.3 and 8.3.5). Accidental triggering has been discussed in previous literature as being particularly pertinent with gestural interaction, due to a lack of feedback (Norman and Nielsen, 2010) and gesture recognition models including ambiguous or similar gestures (Morris, Wobbrock and Wilson, 2010).

One method of reducing accidental triggering could be to take temporal information into account. For instance, speech recognition (Trentin and Gori, 2003), predictive text (Ikegami *et al.*, 2017), and gait analysis (Du, Wang and Wang, 2015) make use of techniques such as hidden Markov models and recurrent neural networks to use previous input data to more accurately determine a user's next or current input. Utilising previous gestural input to help in posture recognition could be effective at reducing accidental triggering, particularly in performances of predetermined musical material, but could become a problem during improvised performances, where the performer's next actions have not been determined in advance, especially in performances that reject predictability.

An example source of inspiration for temporally contextual input is the Dasher Project (Inference Group, 2016), where users build up words and sentences through a continuously zooming interface. A predictive text model provides the user with options for the next letter, with each letters hit-zone size being determined by its probability to be next in the user's sentence.

### Evaluate in the context of a musician's personal practice.

This work used both task-based and open practice research methods, and found that the nature of the task has an impact on the mapping behaviour of glove musicians (Section 7.5.4). Methods that focused on investigating mapping design in a musician's personal practice revealed more relevant insights into real-world mapping practice (Sections 7.4.3 and 8.3). This supports existing literature which calls for using longitudinal methods that examine users' behaviour in an "in the wild" context (Nash, 2011; Kaye, 2009; Collins, 2007).

## 10.7. Future Directions

This work has investigated a new area of musical practice unique to DMIs: the creative design of mappings between gesture and sound by the end-user musician. By examining the mapping design practice of musicians using a mid-air DMI, this research contributes towards the established thoughts and ideas around mapping design for DMIs, with the insights and findings presented also relevant to other fields within HCI, such as mid-air interaction, end-user design tools and gesture recognition.

The empirical work presented made use of qualitative, observational research methodologies designed to elicit insights into mapping design within the context of a musician's personal creative practice, supporting and being supported by quantitative measurements of aspects of mapping design. This approach was made possible thanks to the emerging and growing community of practitioners that has built up around the Mi.Mu Gloves, a rare phenomenon within the DMI field (McPherson and Kim, 2012).

Understanding the creative practice of the end-users of a DMI has been shown to be invaluable in providing nuanced, authoritative and contextually relevant insights about how these instruments effectively facilitate musical expression. Future directions within the field will hopefully see the continued expansion of glove musicians, as well as the fostering of communities around other DMIs.

This research has highlighted the need to follow the development of a musician's creative practice, and future work, building directly on the work presented in this dissertation, could include observing the initial development of glove musicianship for a much longer period than that used in Chapter 7. This would reveal insights into the change in mapping behaviour as glove musicians develop into highly experienced glove musicians with several years of experience (Chapter 8), and could be effectively facilitated by the recent commercialisation of the Mi.Mu Gloves[1], opening up the technology to a much wider group of practitioners.

## 10.8. Concluding Remarks

Digital technology has the potential to revolutionise the way in which we compose, perform and experience music. However, for widespread and mainstream music interaction interfaces, interaction design remains rooted in traditional musical instruments (e.g. piano keyboards) or simple, one dimensional user interface elements (knobs, sliders and buttons), while novel musical interactions often fail to become established within musicians' long-term, professional creative practice.

By examining the creative practice of a group of musicians using a novel interface to facilitate their musical expression, and revealing insights into how these

---

[1] https://shop.mimugloves.com/

musicians achieve expressive performances through end-user mapping design, this dissertation contributes findings to aid in bringing novel music interactions to a wider community of musicians, composers, and audiences.

# Bibliography

ABRSM. (2019) *Grade 8 Piano Syllabus 2019 and 2020*. Available From: `https://gb.abrsm.org/fileadmin/user_upload/PDFs/Grade_8_Piano_Syllabus_2019___2020.pdf` [Last Accessed: 9th August 2019].

Aggarwal, J. K. and Cai, Q. (1999) Human motion analysis: A review. *Computer Vision and Image Understanding*. 73 (3), pp. 428–440.

Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S. and Koh, J. T. K. V. (2012) Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI. *Microsoft Research TechReport MSR-TR-2012-111*.

Alty, J. (1995) Navigating through compositional space: The creativity corridor. *Leonardo*. 28 (3), pp. 215–219.

Antle, A. N., Corness, G. and Droumeva, M. (2009) Human Computer Intuition? Exploring the cognitive basis for intuition in embodied interaction. *International Journal of Arts and Technology*. 2 (3), pp. 235–254.

Arfib, D., Couturier, J.-M., Kessous, L. and Verfaille, V. (2002) Strategies of Mapping Between Gesture Data and Synthesis Model Parameters using Perceptual Spaces. *Organised Sound*. 7 (2), pp. 127–144.

Arif, A. S., Stuerzlinger, W., Filho, E. J. d. M. and Gordynski, A. (2014) Error Behaviours in an Unreliable In-air Gesture Recognizer. In: *CHI '14 Extended*

*Abstracts on Human Factors in Computing Systems*. Toronto, ON, Canada, May 2014. ACM. pp. 1603–1608.

Auslander, P. (2008) *Liveness: Performance in a Mediatized Culture*. Routledge.

Bailly, G., Müller, J., Rohs, M., Wigdor, D. and Kratz, S. (2012) ShoeSense: a new perspective on gestural interaction and wearable applications. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2012*. Austin, TX, USA, May 2012. pp. 1239–1248.

Bandura, A. (1997) *Self-efficacy: The exercise of control*. W.H. Freeman.

Barbosa, J., Malloch, J., Wanderley, M. M. and Huot, S. (2015) What does "Evaluation" mean for the NIME community? In: *Proceedings of New Interfaces for Musical Expression (NIME) 2015*. Baton Rouge, LA, USA, June 2015. pp. 156–161.

Bargas-Avila, J. A. and Hornbæk, K. (2011) Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2011*. May 2011. pp. 2689–2698.

Bellona, J. (2006) jpb.mod: a Max/MSP data modification library for rapid prototyping. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2006*. Paris, France, June 2006.

Benyon, D., Turner, P. and Turner, S. (2005) *Designing interactive systems: People, activities, contexts, technologies*. Harlow, UK: Pearson Education.

Bevilacqua, F., Müller, R. and Schnell, N. (2005) MnM: a Max/MSP mapping toolbox. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2005*. Vancouver, Canada, May 2005. pp. 85–88.

Bin, S. A. (2018) *The Show Must Go Wrong: Towards an understanding of audience perception of error in digital musical instrument performance*. PhD thesis, Queen Mary University of London.

Birnbaum, D., Fiebrink, R., Malloch, J. and Wanderley, M. M. (2005) Towards a dimension space for musical devices. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2005*. Vancouver, Canada, May 2005. pp. 192–195.

Blackwell, A. F. (2006) The Reification of Metaphor as a Design Tool. *ACM Transactions on Computer-Human Interaction (TOCHI)*. 13 (4), pp. 490–530.

Blythe, M. A., Overbeeke, K., Monk, A. F. and Wright, P. C. (2004) *Funology: from usability to enjoyment*. 3rd ed. Springer Science & Business Media.

Blythe, M., Reid, J., Wright, P. and Geelhoed, E. (2006) Interdisciplinary criticism: analysing the experience of riot! a location-sensitive digital narrative. *Behaviour & Information Technology*. 25 (2), pp. 127–139.

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative research in psychology*. 3 (2), pp. 77–101.

Brower, C. (2000) A Cognitive Theory of Musical Meaning. *Journal of music theory*. 44 (2), pp. 323–379.

Brown, D. (2016) The Application of Established Gestural Languages in the Control Mappings of Free-hand Gestural Musical Instruments. In: *Proceedings of the International Conference on Live Interfaces (ICLI) 2016*. Brighton, UK, July 2016. pp. 181–186.

Brown, D., Nash, C. and Mitchell, T. (2016) GestureChords: Transparency in Gesturally Controlled Digital Musical Instruments through Iconicity and Conceptual Metaphor. In: *Proceedings of Sound and Music Computing (SMC) 2016*. Hamburg, Germany, Sept. 2016. pp. 85–92.

Brown, D., Nash, C. and Mitchell, T. (2017) A User Experience Review of Music Interaction Evaluations. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 370–375.

Brown, D., Nash, C. and Mitchell, T. (2018) Simple Mappings, Expressive Movement: A qualitative investigation into the end-user mapping design of experienced mid-air musicians. *Digital Creativity*. 29 (2–3), pp. 129–148.

Brown, D., Nash, C. and Mitchell, T. (2018) Understanding User-Defined Mapping Design for Mid-Air Musical Performance. In: *Proceedings of the 5th International Conference on Movement and Computing (MOCO)*. Genoa, Italy, June 2018. ACM. 27:1–27:8.

Brown, D., Renney, N., Stark, A., Nash, C. and Mitchell, T. (2016) Leimu: Gloveless Music Interaction using a Wrist Mounted Leap Motion. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2016*. Brisbane, Australia, July 2016. pp. 300–304.

Cadoz, C. (1988) Instrumental Gesture and Musical Composition. In: *Proceedings of the International Computer Music Conference (ICMC) 1988*. Kologne, Germany, Feb. 1988. pp. 1–12.

Cadoz, C. and Wanderley, M. M. (2000) Gesture – Music. *Trends in Gestural Control of Music*. pp. 71–93.

Caramiaux, B., Bevilacqua, F., Wanderley, M. M. and Palmer, C. (2018) Dissociable effects of practice variability on learning motor and timing skills. *PloS one*. 13 (3).

Caramiaux, B., Françoise, J., Schnell, N. and Bevilacqua, F. (2014) Mapping Through Listening. *Computer Music Journal*. 38 (3), pp. 34–48.

Caramiaux, B., Montecchio, N., Tanaka, A. and Bevilacqua, F. (2014) Adaptive Gesture Recognition with Variation Estimation for Interactive Systems. *ACM Transactions on Interactive Intelligent Systems (TiiS)*. 4 (4), 18:1–18:34.

Caramiaux, B. and Tanaka, A. (2013) Machine Learning of Musical Gestures. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2013*. Seoul, Korea Republic, May 2013. pp. 513–518.

Casasanto, D. (2009) When is a linguistic metaphor a conceptual metaphor? In: Evans, V. and Pourcel, S. *New Directions in Cognitive Linguistics*. John Benjamins Publishing Company.

Chandler, D. (2007) *Semiotics: the basics*. London, UK: Routledge.

Chapman, P., Selvarajah, S. and Webster, J. (1999) Engagement in multimedia training systems. In: *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences*. Jan. 1999. pp. 9–17.

Choi, A., Joo, S.-B., Oh, E. and Mun, J. H. (2014) Kinematic evaluation of movement smoothness in golf: relationship between the normalized jerk cost of body joints and the clubhead. *Biomedical Engineering Online*. 13 (1), p. 20.

Collins, D. (2005) A synthesis process model of creative thinking in music composition. *Psychology of Music*. 33 (2), pp. 193–216.

Collins, D. (2007) Real-time tracking of the creative music composition process. *Digital Creativity*. 18 (4), pp. 239–256.

Cox, A. (2016) *Music and Embodied Cognition*. Bloomington, IN, USA: Indiana University Press.

Csikszentmihalyi, M. (1996) *Creativity: Flow and the Psychology of Discovery and Invention*. New York: Harper Collins.

Dahl, S., Bevilacqua, F., Bresin, R., Clayton, M., Leante, L., Poggi, I. and Rasamimanana, N. (2010) Gestures in Performance. In: Godøy, R. I. and Leman, M. *Musical gestures: Sound, movement, and meaning*. New York, NY, USA and London, UK: Routledge.

Danto, A. C. (1981) *The transfiguration of the commonplace: a philosophy of art*. Cambridge, MA, USA and London, UK: Harvard University Press.

Di Donato, B. (2017) *MyoMapper*. Available From: `http://www.balandinodidonato.com/myomapper/` [Last Accessed: 20th April 2018].

Di Donato, B., Bullock, J. and Tanaka, A. (2018) Myo Mapper: a Myo armband to OSC mapper. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2018*. Blacksburg, VA, USA, June 2018. pp. 138–143.

Dijk, J. van and Hummels, C. (2017) Designing for Embodied Being-in-the-World: Two Cases, Seven Principles and One Framework. In: *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. Yokohama, Japan, Mar. 2017. ACM. pp. 47–56.

Dillon, A. (2002) Beyond usability: process, outcome and affect in human-computer interactions. *Canadian Journal of Library and Information Science*. 26 (4), pp. 57–69.

Dobrian, C. and Koppelman, D. (2006) The 'E' in NIME: Musical Expression with New Computer Interfaces. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2006*. Paris, France, June 2006. pp. 277–282.

Dourish, P. (2004) *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA, USA: MIT Press.

Drake, C. and Palmer, C. (2000) Skill acquisition in music performance: relations between planning and temporal control. *Cognition*. 74 (1), pp. 1–32.

Du, Y., Wang, W. and Wang, L. (2015) Hierarchical recurrent neural network for skeleton based action recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Boston, MA, USA, June 2015. pp. 1110–1118.

Duke, R. A., Cash, C. D. and Allen, S. E. (2011) Focus of attention affects performance of motor skills in music. *Journal of Research in Music Education*. 59 (1), pp. 44–55.

Eldridge, A. and Kiefer, C. (2017) The self-resonating feedback cello: interfacing gestural and generative processes in improvised performance. In: *Proceedings* *of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 25–29.

Elliott, R., Fischer, C. T. and Rennie, D. L. (1999) Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British journal of clinical psychology*. 38 (3), pp. 215–229.

Evans, P. (1975) *Motivation*. 2015 ed. New York, NY, USA: Psychology Press.

Fels, S., Gadd, A. and Mulder, A. (2002) Mapping Transparency Through Metaphor: Towards More Expressive Musical Instruments. *Organised Sound*. 7 (2), pp. 109–126.

Fiebrink, R. and Cook, P. R. (2010) The Wekinator: a system for real-time, interactive machine learning in music. In: *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)*. Utrecht, The Netherlands, Aug. 2010.

Fiebrink, R., Trueman, D. and Cook, P. R. (2009) A Meta-Instrument for Interactive, On-the-Fly Machine Learning. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, June 2009. pp. 280–285.

Fiebrink, R., Trueman, D., Britt, N. C., Nagai, M., Kaczmarek, K., Early, M., Daniel, M., Hege, A. and Cook, P. R. (2010) Toward Understanding Human-Computer Interaction In Composing The Instrument. In: *Proceedings of the International Computer Music Conference (ICMC) 2010*. New York, NY, USA, June 2010.

Fischer, G. (1999) Domain-oriented design environments: Supporting individual and social creativity. *Computational Models of Creative Design IV*. pp. 83–111.

Fischer, G. (1994) Domain-oriented design environments. *Automated Software Engineering*. 1 (2), pp. 177–203.

Fischer, G. and Girgensohn, A. (1990) End-user modifiability in design environments. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 1990*. Seattle, WA, USA, Apr. 1990. pp. 183–192.

Fontana, F., Camponogara, I., Cesari, P., Vallicella, M. and Ruzzenente, M. (2016) An Exploration on Whole-Body and Foot-Based Vibrotactile Sensitivity to Melodic Consonance. In: *Proceedings of the Sound and Music Computing Conference (SMC) 2016*. Hamburg, Germany, Sept. 2016. pp. 143–150.

Françoise, J. (2013) Gesture–sound mapping by demonstration in interactive music systems. In: *Proceedings of the 21st ACM International Conference on Multimedia*. Barcelona, Spain, Oct. 2013. pp. 1051–1054.

Françoise, J. (2015) *Motion-sound Mapping By Demonstration*. PhD, McGill University.

Fyans, A. C., Gurevich, M. and Stapleton, P. (2009) Spectator Understanding of Error in Performance. In: *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. Boston, MA, USA, Apr. 2009. pp. 3955–3960.

Fyans, A. C., Gurevich, M. and Stapleton, P. (2009) Where Did It All Go Wrong? A Model of Error From the Spectator's Perspective. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, June 2009. pp. 171–172.

Gavrila, D. M. (1999) The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding*. 73 (1), pp. 82–98.

Gelineck, S. and Serafin, S. (2012) Longitudinal evaluation of the integration of digital musical instruments into existing compositional work processes. *Journal of New Music Research*. 41 (3), pp. 259–276.

Gertler, M. (2009) Taylorism. In: *The Dictionary of Human Geography*. Oxford, UK: Blackwell Publishers.

Glaser, B. G. and Anselm, L. S. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Hawthorne, NY, USA: Aldine Publishing Company.

Gniotek, K. and Krucinska, I. (2004) The Basic Problems of Textronics. *Fibres and Textiles in Eastern Europe*. 12 (1), pp. 13–16.

Godøy, R. I., Haga, E. and Jensenius, A. R. (2006) Exploring music-related gestures by sound tracing. A preliminary study. In: *Proceedings of the 2nd International Symposium on Gesture Interfaces for Multimedia Systems*. Leeds, UK, May 2006. pp. 27–33.

Godøy, R. I., Song, M., Nymoen, K., Haugen, M. R. and Jensenius, A. R. (2016) Exploring Sound-Motion Similarity in Musical Experience. *Journal of New Music Research*. 45 (3), pp. 210–222.

Goebl, W. and Palmer, C. (2013) Temporal control and hand movement efficiency in skilled music performance. *PloS one*. 8 (1).

González-Sánchez, V. E., Dahl, S., Hatfield, J. L. and Godøy, R. I. (2019) Characterizing movement fluency in musical performance: Toward a generic measure for technology enhanced learning. *Frontiers in Psychology*. 10 84:1–84:20.

Grande, A. (2015) *experimenting w the mimu gloves*. Available From: `https://www.youtube.com/watch?v=UXxrHKgKVWo` [Last Accessed: 14th March 2020].

Grani, F., Paisa, R., Banas, J. S., Vogiatzoglou, I. and Serafin, S. (2016) Design and evaluation of a gesture driven wavefield synthesis auditory game. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2016*. Brisbane, Australia, July 2016. pp. 188–193.

Green, O. (2011) Agility and playfulness: Technology and skill in the performance ecosystem. *Organised Sound*. 16 (2), pp. 134–144.

Guiard, Y. (1987) Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*. 19 (4), pp. 486–517.

Gurevich, M. and Fyans, A. C. (2011) Digital Musical Interactions: Performer–system relationships and their perception by spectators. *Organised Sound*. 16 (2), pp. 166–175.

Gurevich, M., Stapleton, P. and Bennet, P. (2009) Designing for Style in New Musical Interactions. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, June 2009. pp. 213–217.

Gurevich, M., Stapleton, P. and Marquez-Borbon, A. (2010) Style and Constraint in Electronic Musical Instruments. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2010*. Sydney, Australia, June 2010. pp. 106–111.

Gurevich, M. and Treviño, J. (2007) Expression and its discontents: toward an ecology of musical creation. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2007*. New York, NY, USA, June 2007. pp. 106–111.

Gustafson, S., Bierwirth, D. and Baudisch, P. (2010) Imaginary Interfaces: Spatial Interaction with Empty Hands and Without Visual Feedback. In: *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*. New York, New York, USA, 2010. ACM. pp. 3–12.

Hampe, B. (2005) Image schemas in cognitive linguistics: Introduction. *From perception to meaning: Image schemas in cognitive linguistics*. 29 p. 1.

Hassenzahl, M., Burmester, M. and Koller, F. (2003) AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: *Mensch & Computer 2003*. pp. 187–196. Springer.

Hassenzahl, M. and Tractinsky, N. (2006) User experience – a research agenda. *Behaviour & Information Technology*. 25 (2), pp. 91–97.

Hattwick, I. and Wanderley, M. M. (2012) A Dimension Space for Evaluating Collaborative Musical Performance Systems. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2012*. Ann Arbor, MI, USA, May 2012. pp. 21–23.

Hincapié-Ramos, J. D., Guo, X., Moghadasian, P. and Irani, P. (2014) Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-air Interactions. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Toronto, Ontario, Canada, 2014. ACM. pp. 1063–1072.

Hodges, D. A. and Nolker, D. B. (1992) The acquisition of music reading skills. *Handbook of research on music teaching and learning*. pp. 466–471.

Holstein, P. *Free Violin Studies, Etudes, and Solos*. Available From: `https://fiddlerman.com/studies-etudes-and-music/free-violin-music-studies-and-etudes/` [Last Accessed: 8th March 2019].

Howard, B. and Howard, S. (2001) Lightglove: Wrist-Worn Virtual Typing and Pointing. In: *Proceedings of the Fifth International Symposium on Wearable Computers*. Zurich, Switzerland, Oct. 2001. pp. 172–173.

Hsu, W. T. and Sosnick, M. H. (2009) Evaluating Interactive Music Systems: An HCI Approach. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, June 2009. pp. 25–28.

Hunt, A. and Kirk, R. (1999) Radical User Interfaces for Real-Time Control. In: *Proceedings of the 25th EUROMICRO Conference*. Sept. 1999. pp. 6–12.

Hunt, A. and Kirk, R. (2000) Mapping Strategies for Musical Performance. *Trends in Gestural Control of Music*. pp. 231–258.

Hunt, A. and Wanderley, M. M. (2002) Mapping Performer Parameters to Synthesis Engines. *Organised Sound*. 7 (2), pp. 97–108.

Hunt, A., Wanderley, M. M. and Paradis, M. (2003) The Importance of Parameter Mapping in Electronic Instrument Design. *Journal of New Music Research*. 32 (4), pp. 429–440.

Hunt, A., Wanderley, M. and Kirk, R. (2000) Towards a Model for Instrumental Mapping in Expert Musical Interaction. In: *Proceedings of the International Computer Music Conference (ICMC) 2000*. Berlin, Germany, Sept. 2000. pp. 209–212.

Ikegami, Y., Sakurai, Y., Damiani, E., Knauf, R. and Tsuruta, S. (2017) Flick: Japanese Input Method Editor Using N-Gram and Recurrent Neural Network Language Model Based Predictive Text Input. In: *13th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*. Jaipur, India, Dec. 2017. pp. 50–55.

Inference Group. Inference Group (2016) *Dasher Project*. Available From: `www.inference.org.uk/dasher` [Last Accessed: 28th March 2020].

Jack, R. H., Stockman, T. and McPherson, A. (2016) Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument. In: *Proceedings of Audio Mostly 2016*. Norrköping, Sweden, Oct. 2016. pp. 116–123.

Jack, R. H., Stockman, T. and McPherson, A. (2017) Rich gesture, reduced control: the influence of constrained mappings on performance technique. In: *Proceedings of the 4th International Conference on Movement Computing*. London, UK, June 2017. 15:1–15:8.

Jensenius, A. R. (2007) *Action-sound: developing methods and tools to study music-related body movement*. PhD thesis, University of Oslo.

Jensenius, A. R. (2014) To Gesture or Not? An Analysis of Terminology in NIME Proceedings 2001–2013. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, UK, July 2014. pp. 217–220.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I. and Leman, M. (2009) Musical Gestures: Concepts and Methods in Research. In: Godøy, R. I. and Leman, M. (2010) *Musical gestures: Sound, movement, and meaning*. New York, NY, USA and London, UK: Routledge.

Jessop, E. (2009) The Vocal Augmentation and Manipulation Prosthesis (VAMP): A Conducting-based Gestural Controller for Vocal Performance. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2009*. Pittsburgh, PA, USA, June 2009. pp. 256–259.

Jiang, S., Sakai, K., Yamada, M., Fujimoto, J., Hidaka, H., Okabayashi, K. and Murase, Y. (2014) Developing a Wearable Wrist Glove for Fieldwork Support: A User Activity-Driven Approach. In: *IEEE/SICE International Symposium on System Integration (SII) 2014*. Toko, Japan, Dec. 2014. pp. 22–27.

Johnson, M. (2005) The philosophical significance of image schemas. In: Hampe, B. *From perception to meaning: Image schemas in cognitive linguistics*. Berlin, Germany & New York, NY, USA: Mouton de Gruyter. pp. 15–34.

Johnson, M. (2008) *The meaning of the body: Aesthetics of human understanding*. Chicago, IL, USA: University of Chicago Press.

Johnson, M. and Larson, S. (2003) "Something in the Way She Moves" – Metaphors of Musical Motion. *Metaphor and Symbol*. 18 (2), pp. 63–84.

Johnston, A. (2009) *Interfaces for Musical Expression Based on Simulated Physical Models*. PhD thesis, University of Technology, Sydney.

Johnston, A. (2011) Beyond Evaluation: linking practice and theory in new musical interface design. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2011*. Oslo, Norway, June 2011. pp. 280–283.

Jordà, S. (2004a) Digital Instruments and Players: Part I–Efficiency and Apprenticeship. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2004*. Hamamatsu, Japan, June 2004a. pp. 59–63.

Jordà, S. (2004b) Digital Instruments and Players: Part II–Diversity, Freedom and Control. In: *Proceedings of the International Computer Music Conference (ICMC) 2004*. Miami, FL, USA, Nov. 2004b. pp. 706–710.

Jordà, S. and Mealla, S. (2014) A methodological framework for teaching, evaluating and informing NIME design with a focus on expressiveness and mapping. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, UK, July 2014. pp. 233–238.

Jørgensen, M. H., Knudsen, A. S., Wilmot, T. M., Lund, K. D., Serafin, S. and Purwins, H. (2015) A Mobile Music Museum Experience for Children. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2015*. Baton Rouge, LA, USA, June 2015. pp. 36–37.

Juslin, P. N. (2003) Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance. *Psychology of Music*. 31 (3), pp. 273–302.

Kaye, J. N. (2009) *The epistemology & evaluation of experience-focused hci*. PhD thesis, Cornell University.

Kiani Galoogahi, H., Fagg, A., Huang, C., Ramanan, D. and Lucey, S. (2017) Need for speed: A benchmark for higher frame rate object tracking. In: *Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy, Oct. 2017. IEEE. pp. 1125–1134.

Kiefer, C., Collins, N. and Fitzpatrick, G. (2008) HCI Methodology For Evaluating Musical Controllers: A Case Study. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2008*. Genova, Italy, June 2008. pp. 87–90.

Kim, D., Hilliges, O., Izadi, S., Butler, A. D., Chen, J., Oikonomidis, I. and Olivier, P. (2012) Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. In: *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. Cambridge, MA, USA, Oct. 2012. pp. 167–176.

Klipfel, K. (2017) MIDI Motion: Interactive Music Composition Gloves. In: *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. Yokohama, Japan, 2017. ACM. pp. 757–760.

Kray, C., Nesbitt, D., Dawson, J. and Rohs, M. (2010) User-defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. Lisbon, Portugal, 2010. ACM. pp. 239–248.

Kvifte, T. and Jensenius, A. R. (2006) Towards a coherent terminology and model of instrument description and design. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2006*. Paris, France, June 2006. pp. 220–225.

Lakoff, G. and Johnson, M. (1980) *Metaphors we live by*. Chicago, IL, USA: University of Chicago Press.

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P. and Kort, J. (2009) Understanding, Scoping and Defining User Experience: A Survey Approach. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2009*. Boston, MA, USA, 2009. ACM. pp. 719–728.

Leman, M. (2008) *Embodied music cognition and mediation technology*. Cambridge, MA, USA: MIT Press.

Leman, M. and Godøy, R. I. (2010) Why study musical gestures? In: Godøy, R. I. and Leman, M. (2010) *Musical gestures: Sound, movement, and meaning*. New York, NY, USA and London, UK: Routledge. pp. 3–11.

Lepecki, A. (2006) *Exhausting dance: Performance and the politics of movement*. New York, NY, USA and London, UK: Routledge.

Lerdahl, F. (1988) Tonal pitch space. *Music Perception: An Interdisciplinary Journal*. 5 (3), pp. 315–349.

Levitin, D. J., McAdams, S. and Adams, R. L. (2002) Control Parameters for Musical Instruments: A Foundation for New Mappings of Gesture to Sound. *Organised Sound*. 7 (2), pp. 171–189.

Luhtala, M., Niemeläinen, I., Plomp, J., Turunen, M. and Tuomisto, J. (2012) Studying Aesthetics in a Musical Interface Design Process Through 'Aesthetic

Experience Prism'. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2012*. Ann Arbor, Michigan, 2012. University of Michigan.

Macionis, M. J. and Kapur, A. (2018) Sansa: A Modified Sansula for Extended Compositional Techniques Using Machine Learning. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2018*. Blacksburg, Virginia, USA, June 2018. Virginia Tech. pp. 78–81.

Mackenzie, I. S. (2013) *Human Computer Interaction: An Empirical Research Perspective*. Waltham, MA, USA: Morgan Kaufmann.

Magnusson, T. (2010) Designing Constraints: Composing and Performing with Digital Musical Systems. *Computer Music Journal*. 34 (4), pp. 62–73.

Magnusson, T. (2017) Musical organics: a heterarchical approach to digital organology. *Journal of New Music Research*. 46 (3), pp. 286–303.

Makelberge, N., Barbosa, Á., Perrotta, A. V. and Ferreira, L. S. (2012) Perfect Take: Experience design and new interfaces for musical expression. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2012*. Ann Arbor, MI, USA, May 2012.

Malloch, J., Sinclair, S. and Wanderley, M. M. (2007) From Controller to Sound: Tools for Collaborative Development of Digital Musical Instruments. In: *Proceedings of the International Computer Music Conference (ICMC) 2007*. Copenhagen, Denmark, 2007. pp. 65–72.

Malloch, J., Sinclair, S. and Wanderley, M. M. (2013) Libmapper: (a library for connecting things). In: *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. Paris, France, May 2013. pp. 3087–3090.

Marquez-Borbon, A. and Stapleton, P. (2015) Fourteen years of NIME: the value and meaning of "community" in interactive music research. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2015*. Baton Rouge, LA, USA, June 2015. pp. 307–312.

Marx, A. (1994) Using Metaphor Effectively in User Interface Design. In: *Conference Companion on Human Factors in Computing Systems 1994*. New York, NY, USA, Apr. 1994. pp. 379–380.

Mathews, M. V. and Abbott, C. (1980) The Sequential Drum. *Computer Music Journal*. 4 (4), pp. 45–59.

McCarthy, J., Wright, P., Wallace, J. and Dearden, A. (2006) The experience of enchantment in human–computer interaction. *Personal and Ubiquitous Computing*. 10 (6), pp. 369–378.

McPherson, A. P., Chamberlain, A., Hazzard, A., McGrath, S. and Benford, S. (2016) Designing for Exploratory Play with a Hackable Digital Musical Instrument. In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. Brisbane, QLD, Australia, 2016. ACM. pp. 1233–1245.

McPherson, A. P., Gierakowski, A. and Stark, A. M. (2013) The space between the notes: adding expressive pitch control to the piano keyboard. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2013*. Paris, France, May 2013. pp. 2195–2204.

McPherson, A. P., Jack, R. H. and Moro, G. (2016) Action-sound latency: Are our tools fast enough? In: *Proceedings of New Interfaces for Musical Expression (NIME) 2016*. Brisbane, Australia, July 2016. Griffith University. pp. 20–25.

McPherson, A. P. and Kim, Y. E. (2012) The problem of the second performer: Building a community around an augmented piano. *Computer Music Journal*. 36 (4), pp. 10–27.

McPherson, A. and Zappi, V. (2015) Exposing the Scaffolding of Digital Instruments with Hardware-Software Feedback Loops. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2015*. Baton Rouge, LA, USA, June 2015. pp. 162–167.

McPherson, G. E. (2005) From child to musician: Skill development during the beginning stages of learning an instrument. *Psychology of Music*. 33 (1), pp. 5–35.

McPherson, G. E. and McCormick, J. (2006) Self-efficacy and music performance. *Psychology of music*. 34 (3), pp. 322–336.

Mekler, E. D. and Hornbæk, K. (2016) Momentary Pleasure or Lasting Meaning? Distinguishing Eudaimonic and Hedonic User Experiences. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Santa Clara, California, USA, 2016. ACM. pp. 4509–4520.

Mendoza, V. and Novick, D. G. (2005) Usability over time. In: *Proceedings of the 23rd Annual International Conference on Design of Communication: Documenting & Designing for Pervasive Information*. Coventry, UK, Sept. 2005. pp. 151–158.

Merrill, D. and Paradiso, J. A. (2005) Personalization, expressivity, and learnability of an implicit mapping strategy for physical interfaces. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems, Extended Abstracts*. 2005. pp. 2152–2161.

Mi.Mu Ltd. MiMu (2018) *Mi.Mu Gloves*. Available From: `http://www.mimugloves.com` [Last Accessed: 15th July 2019].

Miranda, E. R. and Wanderley, M. M. (2006) *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. Middleton, Wisconsin, USA: AR Editions, Inc.

Mitchell, T. J., Madgwick, S. and Heap, I. (2012) Musical Interaction with Hand Posture and Orientation: A Toolbox of Gestural Control Mechanisms. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2012*. Ann Arbor, MI, USA, May 2012.

Mitchell, T. and Heap, I. (2011) SoundGrasp: A Gestural Interface for the Performance of Live Music. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2011*. Oslo, Norway, June 2011. pp. 465–468.

Modler, P., Myatt, T. and Saup, M. (2003) An Experimental Set of Hand Gestures for Expressive Control of Musical Parameters in Realtime. In: *Proceedings of New interfaces for Musical Expression (NIME) 2003*. Montreal, Canada, May 2003. pp. 146–150.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D. F., Group, Q. *et al.* (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*. 354 (9193), pp. 1896–1900.

Morreale, F., De Angeli, A. and O'Modhrain, S. (2014) Musical Interface Design: An Experience-oriented Framework. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, UK, July 2014. pp. 467–472.

Morris, M. R., Wobbrock, J. O. and Wilson, A. D. (2010) Understanding Users' Preferences for Surface Gestures. In: *Proceedings of Graphics Interface 2010*. Toronto, ON, Canada, 2010. pp. 261–268.

Mudd, T. (2019) 9. Material-Oriented Musical Interactions. In: Holland, S., Mudd, T., Wilkie-McKenna, K., McPherson, A. and Wanderley, M. M. *New Directions In Music and Human-Computer Interaction*. Springer. pp. 123–134.

Mulder, A. (2000) Towards a Choice of Gestural Constraints for Instrumental Performers. *Trends in Gestural Control of Music*. pp. 315–335.

Müllensiefen, D., Gingras, B., Musil, J. and Stewart, L. (2014) The musicality of non-musicians: an index for assessing musical sophistication in the general population. *PloS one*. 9 (2), pp. 1–23.

Müller, J., Geier, M., Dicke, C. and Spors, S. (2014) The BoomRoom: mid-air direct interaction with virtual sound sources. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2014*. Toronto, ON, Canada, May 2014. pp. 247–256.

Müller, S. and Mazzola, G. (2003) The Extraction of Expressive Shaping in Performance. *Computer Music Journal*. 27 (1), pp. 47–58.

Myers, B., Hudson, S. E., Pausch, R. and Pausch, R. (2000) Past, Present, and Future of User Interface Software Tools. *ACM Transactions on Computer-Human Interaction (TOCHI)*. 7 (1), pp. 3–28.

Nakano, T., Goto, M. and Hiraga, Y. (2006) An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features. In: *Ninth International Conference on Spoken Language Processing*. Pittsburgh, PA, USA, Sept. 2006.

Nakra, T. M. (2002) Synthesizing Expressive Music Through the Language of Conducting. *Journal of New Music Research*. 31 (1), pp. 11–26.

Nash, C. (2011) *Supporting Virtuosity and Flow in Computer Music*. PhD, University of Cambridge.

Nash, C. (2015) The Cognitive Dimensions of Music Notations. In: *Proceedings of the First International Conference on Technologies for Music Notation and Representation (TENOR 2015)*. Paris, France, May 2015. pp. 190–202.

Nash, C. and Blackwell, A. (2011) Tracking virtuosity and flow in computer music. In: *Proceedings of the International Computer Music Conference (ICMC) 2011*. Huddersfield, UK, Aug. 2011. pp. 575–582.

Nash, C. and Blackwell, A. (2014) Flow of Creative Interaction with Digital Notations. In: Collins, K., Kapralos, B. and Tessler, H. (2014) *The Oxford Handbook of Interactive Audio*. New York, NY, USA: Oxford University Press. pp. 387–404.

Nielsen, J. (1993) *Usability Engineering*. Cambridge, MA, USA: Academic Press, Inc.

Nielsen, J. (1994) Heuristic Evaluation. *Usability Inspection Methods*. 17 (1), pp. 25–62.

Nielsen, M., Störring, M., Moeslund, T. B. and Granum, E. (2004) A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In: Camurri, A. and Volpe, G. *Gesture Based Communication in Human-Computer Interaction*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 409–420.

Norman, D. A. (1988) *The design of everyday things*. Basic books.

Norman, D. A. and Nielsen, J. (2010) Gestural Interfaces: a step backward in usability. *interactions*. 17 (5), pp. 46–49.

O'Modhrain, S. (2011) A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal*. 35 (1), pp. 28–42.

Paine, G. (2009) Towards Unified Design Guidelines for New Interfaces for Musical Expression. *Organised Sound*. 14 (2), pp. 142–155.

Paine, G., Stevenson, I. and Pearce, A. (2007) The Thummer Mapping Project (ThuMP). In: *Proceedings of New Interfaces for Musical Expression (NIME) 2007*. New York, NY, USA, June 2007. pp. 70–77.

Palacio-Quintin, C. (2003) The Hyper-Flute. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2003*. Montreal, Canada, May 2003. pp. 206–207.

Palmer, C., Carter, C., Koopmans, E. and Loehr, J. D. (2007) Movement, planning, and music: Motion coordinates of skilled performance. In: *Proceedings of the International Conference on Music Communication Science*. 2007. pp. 119–122.

Pardue, L., Nian, D., Harte, C. and McPherson, A. (2014) Low-Latency Audio Pitch Tracking: A Multi-Modal Sensor-Assisted Approach. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, United Kingdom, 2014. pp. 54–59.

Pavlovic, V., Sharma, R., Huang, T. S. *et al.* (1997) Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 19 (7), pp. 677–695.

Peirce, C. S. (1986) *The Philosophy of Peirce: Selected Writings*. Mineola, NY, USA: Dover Publications Inc.

Poepel, C. (2005) On Interface Expressivity: A Player-Based Study. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2005*. Vancouver, Canada, May 2005. pp. 228–231.

Reeves, S., Benford, S., O'Malley, C. and Fraser, M. (2005) Designing the Spectator Experience. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2005*. Montreal, Canada, Apr. 2005. pp. 741–750.

Rehg, J. and Kanade, T. (1994) DigitEyes: Vision-Based Hand Tracking for Human-Computer Interaction. In: *Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects*. Nov. 1994. pp. 16–22.

Resnick, M., Myers, B., Nakakoji, K., Shneiderman, B., Pausch, R., Selker, T. and Eisenberg, M. (2005) Design Principles for Tools to Support Creative Thinking. In: *NSF Workshop Report on Creativity Support Tools*. Washington DC, USA, June 2005.

Reybrouck, M. (2012) Musical sense-making and the concept of affordance: an ecosemiotic and experiential approach. *Biosemiotics*. 5 (3), pp. 391–409.

Reybrouck, M. (2017) Music and Semiotics: An Experiential Approach to Musical Sense-Making. In: *Interdisciplinary Approaches to Semiotics*. InTech.

Rohrer, B., Fasoli, S., Krebs, H. I., Hughes, R., Volpe, B., Frontera, W. R., Stein, J. and Hogan, N. (2002) Movement smoothness changes during stroke recovery. *Journal of Neuroscience*. 22 (18), pp. 8297–8304.

Ross, P. R., Overbeeke, C. J., Wensveen, S. A. G. and Hummels, C. M. (2008) A designerly critique on enchantment. *Personal and Ubiquitous Computing*. 12 (5), pp. 359–371.

Rovan, J. B., Wanderley, M. M., Dubnov, S. and Depalle, P. (1997) Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer

Music Performance. In: *Proceedings of Kansei - The Technology of Emotion Workshop 1997*. Genoa, Italy, Oct. 1997. pp. 3–4.

Saussure, F. d. (1983) *Course in General Linguistics*. London, UK: Duckworth.

Schneider, A. (2010) Music and Gestures: A Historical Introduction and Survey of Earlier Research. In: Godøy, R. I. and Leman, M. *Musical gestures: Sound, movement, and meaning*. New York, NY, USA and London, UK: Routledge. pp. 69–100.

Scurto, H., Bevilacqua, F. and Françoise, J. (2017) Shaping and Exploring Interactive Motion-Sound Mappings Using Online Clustering Techniques. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 410–415.

Sengers, P., Boehner, K., Mateas, M. and Gay, G. (2008) The disenchantment of affect. *Personal and Ubiquitous Computing*. 12 (5), pp. 347–358.

Serafin, S., Trento, S., Grani, F., Perner-Wilson, H., Madgwick, S. and Mitchell, T. J. (2014) Controlling Physically based Virtual Musical Instruments using the Gloves. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, UK, July 2014. pp. 521–524.

Shepperd, M., Guo, Y., Li, N., Arzoky, M., Capiluppi, A., Counsell, S., Destefanis, G., Swift, S., Tucker, A. and Yousefi, L. (2019) The Prevalence of Errors in Machine Learning Experiments. *arXiv*. 1909.04436.

El-Shimy, D. and Cooperstock, J. R. (2016) User-Driven Techniques for the Design and Evaluation of New Musical Interfaces. *Computer Music Journal*. 40 (2), pp. 35–46.

Shneiderman, B. (1996) The eyes have it: A task by data type taxonomy for information visualizations. In: *Proceedings of the IEEE Symposium on Visual Languages 1996*. 1996. pp. 336–343.

Shneiderman, B. (2007) Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM*. 50 (12), pp. 20–32.

Siegel, D. (2012) Usability for Engaged Users: The Naturalistic Approach to Evaluation. In: Jacko, J. A. *The Human-Computer Interaction Handbook*. Boca Raton, FL, USA: CRC Press, Taylor & Francis Group. pp. 1245–1258.

Silva, E. S., Abreu, J. A. O. de, Almeida, J. H. P. de, Teichrieb, V. and Ramalho, G. L. (2013) A Preliminary Evaluation of the Leap Motion Sensor as Controller of New Digital Musical Instruments.

Sonami, L. *Lady's glove*. Available From: `http://sonami.net/ladys-glove/` [Accessed 4th May 2018].

Spool, J. (2005) What makes design seem intuitive? *User Interface Engineering*. 10 (1).

Stowell, D., Plumbley, M. D. and Bryan-Kinns, N. (2008) Discourse Analysis Evaluation Method for Expressive Musical Interfaces. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2008*. Genova, Italy, June 2008. pp. 81–86.

Sturman, D. J. and Zeltzer, D. (1994) A Survey of Glove-based Input. *IEEE Computer Graphics and Applications*. 14 (1), pp. 30–39.

Tanaka, A. (1993) Musical technical issues in using interactive instrument technology with application to the BioMuse. In: *Proceedings of the International Computer Music Conference (ICMC) 1993*. Waseda University, Japan, 1993. pp. 124–126.

Tanaka, A. (2010) Mapping out instruments, affordances, and mobiles. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2010*. Sydney, Australia, June 2010. pp. 88–93.

Taruskin, R. (1995) *Text and Act: Essays on Music and Performance*. New York, NY, USA: Oxford University Press.

Torre, G., Andersen, K. and Baldé, F. (2016) The Hands: The making of a digital musical instrument. *Computer Music Journal*. 40 (2), pp. 22–34.

Trentin, E. and Gori, M. (2003) Robust combination of neural networks and hidden Markov models for speech recognition. *IEEE Transactions on Neural Networks*. 14 (6), pp. 1519–1531.

Van Nort, D. and Wanderley, M. M. (2006) The LoM Mapping Toolbox for Max/MSP/Jitter. In: *Proceedings of the International Computer Music Conference (ICMC) 2006*. New Orleans, LA, USA, 2006. pp. 397–400.

Vertegaal, R. and Eaglestone, B. (1996) Comparison of input devices in an ISEE direct timbre manipulation task. *Interacting with Computers*. 8 (1), pp. 13–30.

Visi, F., Caramiaux, B., Mcloughlin, M. and Miranda, E. (2017) A Knowledge-based, Data-driven Method for Action-sound Mapping. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 231–236.

Vogl, R. and Knees, P. (2017) An Intelligent Drum Machine for Electronic Dance Music Production and Performance. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2017*. Copenhagen, Denmark, May 2017. pp. 251–256.

Vogt, D. M. and Wood, R. J. (2014) Wrist Angle Measurements Using Soft Sensors. In: *IEEE SENSORS 2014*. Valencia, Spain, Nov. 2014. pp. 1631–1634.

Waisvisz, M. (1985) The hands: A set of remote midi controllers. In: *Proceedings of the International Computer Music Conference (ICMC) 1985*. 1985.

Wanderley, M. M. (2001) Gestural Control of Music. In: *International Workshop Human Supervision and Control in Engineering and Music 2001*. Kassel, Germany, Sept. 2001. pp. 632–644.

Wanderley, M. M. and Orio, N. (2002) Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*. 26 (3), pp. 62–76.

Wang, R. Y. and Popović, J. (2009) Real-time Hand-tracking with a Color Glove. *ACM Transactions on Graphics (TOG)*. 28 (3), 63:1–63:8.

Watts, C., Murphy, J. and Barnes-Burroughs, K. (2003) Pitch matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with nontalented singing voices in conditions of varying feedback. *Journal of Voice*. 17 (2), pp. 185–194.

Wessel, D. and Wright, M. (2002) Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal*. 26 (3), pp. 11–22.

Wilkie, K., Holland, S. and Mulholland, P. (2010) What can the language of musicians tell us about music interaction design? *Computer Music Journal*. 34 (4), pp. 34–48.

Winkler, T. (1995) Making Motion Musical: Gesture Mapping Strategies for Interactive Computer Music. In: *Proceedings of the International Computer Music Conference (ICMC) 1995*. Banff, AB, Canada, 1995. pp. 261–264.

Winograd, T. and Flores, F. (1986) *Understanding computers and cognition: A new foundation for design*. Bristol, UK: Intellect Books.

WIRED UK. (2013) *Imogen Heap Performance with Musical Gloves Demo | WIRED 2012 | WIRED*. Available From: `https://www.youtube.com/watch?v=6btFObRRD9k` [Last Accessed: 14th March 2020].

Wobbrock, J. O., Morris, M. R. and Wilson, A. D. (2009) User-defined Gestures for Surface Computing. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2009*. Boston, MA, USA, 2009. ACM. pp. 1083–1092.

Wong, P. H.-K. (2011) *Conceptual Metaphor in the Practice of Computer Music*. MSc, Mills College.

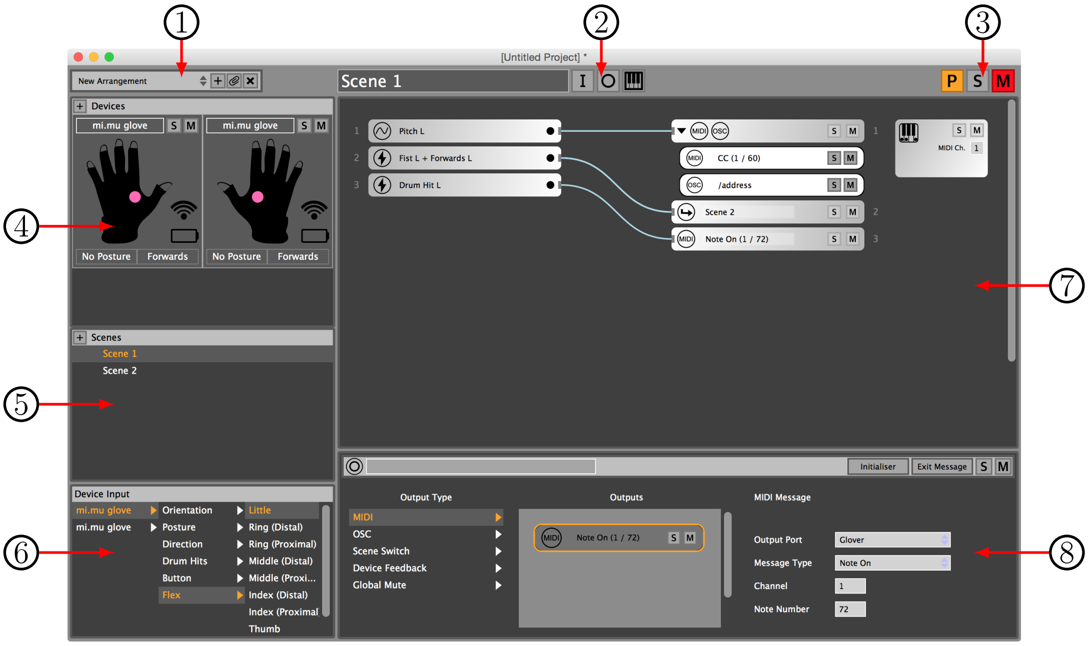
Zappi, V. and McPherson, A. (2014) Design and Use of a Hackable Digital Instrument. In: *Proceedings of the International Conference on Live Interfaces (ICLI) 2014*. Lisbon, Portugal, Nov. 2014.

Zappi, V. and McPherson, A. (2014) Dimensionality and Appropriation in Digital Musical Instrument Design. In: *Proceedings of New Interfaces for Musical Expression (NIME) 2014*. London, UK, July 2014. pp. 455–460.

Zappi, V. and McPherson, A. (2015) The D-Box: How to rethink a digital musical instrument. In: *Proceedings of the 21st International Symposium on Electronic Art*. Vancouver, BC, Canada, Aug. 2015.

# A. Glover Instructions

This document describes how to set up the Mi.Mu Gloves and their software application Glover. Glover is used to design relationships between the gestural data from the Gloves and MIDI or OSC output. This guide describes the process of setting up the Gloves and Glover and a brief overview of its features.



Output Port Glover
Message Type Note On
Channel 1
Note Number 72

Figure A.1.: An example Glover project.

1. Arrangement selection.
2. New mapping options.
3. Perform mode, solo and global mute.
4. Device overview panel.

5. Scene overview panel.
6. Device input panel.
7. Mapping panel.
8. Inspector panel.

## A.1. The Gloves

The Gloves are wireless gestural controllers, which use a variety of sensors to detect your hands orientation and posture. The Gloves send this data wirelessly through WiFi to Glover.

### A.1.1. Anatomy of a Glove

A pair of Mi.Mu Gloves should include:

- 2 Mi.Mu Gloves.
- 1 WiFi router.
- 2 USB-to-LiPo battery chargers.
- 2 USB power cables.

### A.1.2. Setting up the Gloves

1. The Gloves are powered wirelessly by LiPo batteries or wired by USB. To turn the gloves on, plug the power source into the the socket at the bottom of the computing board (Figure A.2.1). Some pairs may have an on/off switch, located at the bottom of the board. (Figure A.2.2). On powering-up, the glove will vibrate once and the battery indicator will light up (Figure A.2.3).
2. Plug both the Ethernet connection of the WiFi router into your laptop, and the power cable into the mains.

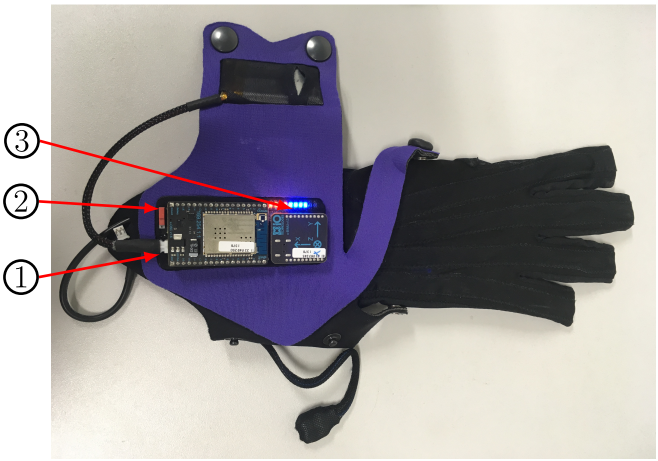


Figure A.2.: A Mi.Mu Glove.

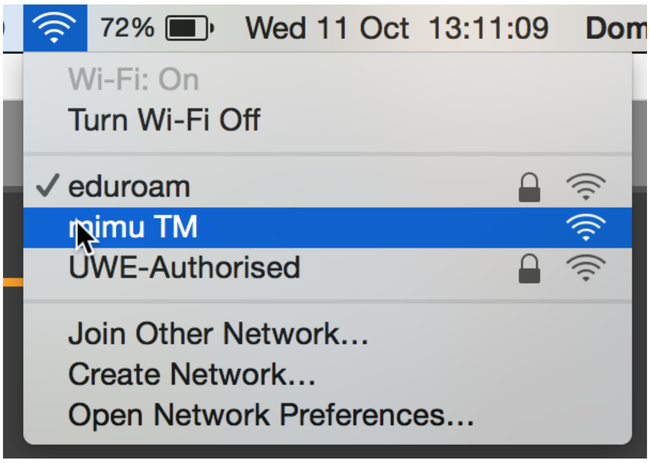


Figure A.3.: Connecting the WiFi.

3. Connect to this router (called “mimu <something>”) in your WiFi settings (Figure A.3).

## A.2. Adding the Gloves to Glover

Once a glove is on, you must set it up in Glover, which is done in the Devices section.

### A.2.1. Connection and calibration

The first thing to do is connect and calibrate each Glove. Follow these steps for each glove.

1. Click on the “+” button next in the Devices section in the Glover project. This brings up a menu of devices that Glover supports (Figure A.4).

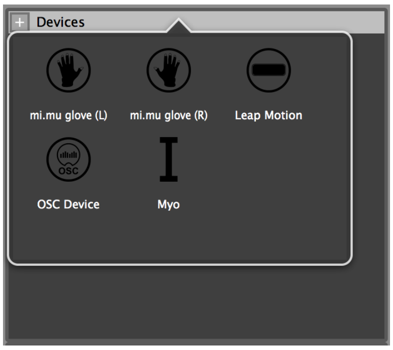
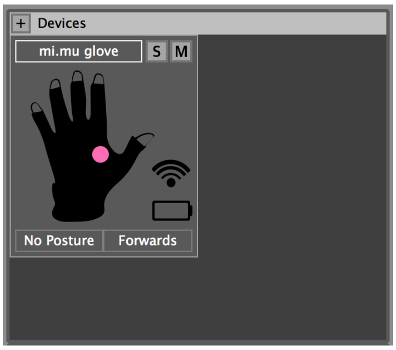
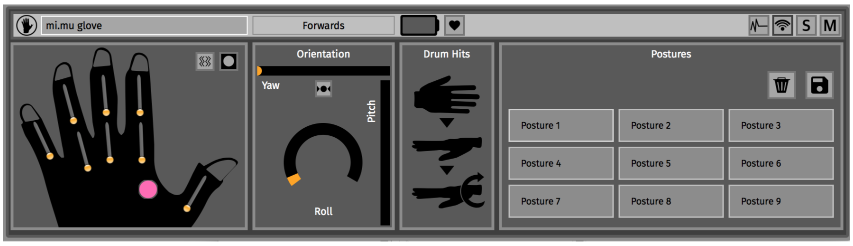


Figure A.4.: Device menu.

2. Select “Mi.Mu Glove (L)”. This will create a new Device object (Figure A.5a). Click on the new device object and its settings will be displayed in the inspector panel in the Glover project (Figure A.5b). Rename the glove from “mi.mu glove” to “left”.
3. Put the Gloves on and make sure they are turned on.



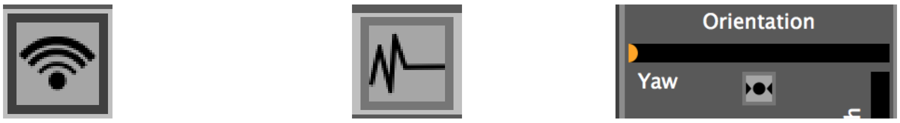
(a) New Device object.



(b) Device Inspector Panel(left hand).

Figure A.5.: Device Settings.

4. Click on the WiFi button to connect the Glove (Figure A.6a). It will flash red if no connection can be made, and will go orange once a connection is made. Connecting to a glove can take a few moments.
5. Glover needs to calibrate the sensors to your hand. Click on the Calibrate button (Figure A.6b), then clench and unclench your hand several times, before clicking on the Calibrate button again. This calibrates the gloves to your fingers, and tells Glover the minimum and maximum amounts that your fingers open and close.
6. Point your hand forwards. You may notice that Glover will think you are pointing left, right or even backwards. This is because you need to tell the gloves which direction you are facing. Point forwards, and click the “set forwards” button (Figure A.6c) underneath the Orientation slider in the device’s inspector.
7. Repeat steps 1–5 with the right-hand glove.



(a) WiFi button. (b) Calibrate button. (c) Set forwards button.

Figure A.6.: Device Setup Controls.

### A.2.2. Training the posture recogniser

Glover is able to recognise your hand's posture using the posture recogniser in the device inspector (Figure A.7). The posture recogniser uses the flex sensors along each finger, and so the hand's orientation or position does not effect the posture recogniser.

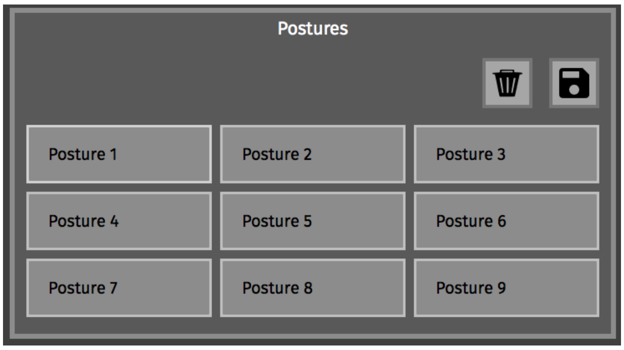


Figure A.7.: The Posture Recogniser.

1. Decide on the posture you want to make.
2. Making sure you have the appropriate Glove open in the inspector, right click (two-finger-click) on "Posture 1" and rename it to a more appropriate name.
3. Position your hand into the desired posture.
4. Click on the renamed "Posture 1" button several times. This gives the posture recogniser multiple examples. Around 10 examples should be sufficient.
5. Repeat steps 1–4 with any other desired postures.
6. Click the "Save" button.

7. Test the posture recogniser by making each posture and checking that the correct posture is being recognised.
8. You can reset the posture recogniser by clicking on the “trash” icon or redo a single posture by right-clicking (two-finger-click) on each posture button. You can also add new postures and training examples on top of your existing ones.

## A.2.3. Device Inputs

Device inputs are always available in the “Device Input” section of the Glover project (Figure A.8).

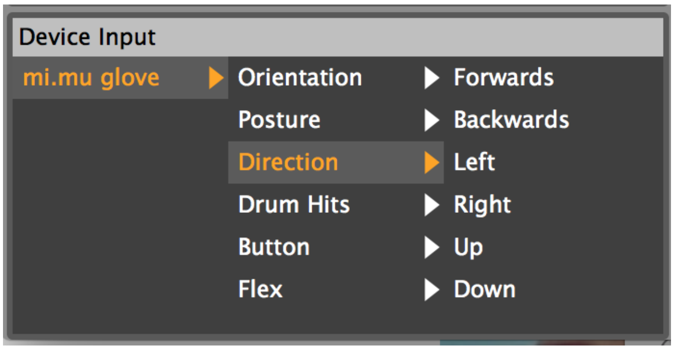


Figure A.8.: Device Inputs box.

Each mapping option: postures, orientations, directions etc., has one of three behaviours:

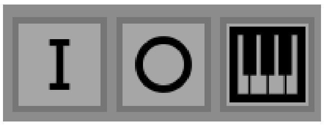
- **Movements:** Continuous controls from body movements, such as the pitch, yaw and roll of the wrist, and the amount of flex of each finger.
- **Events:** Trigger controls that notify that a specific action has occurred, such as drum hits.
- **Qualifiers:** Controls that can either be either occurring or not, such as specific hand postures or direction.

We will look at using these device inputs to map movements to MIDI data in the next section.

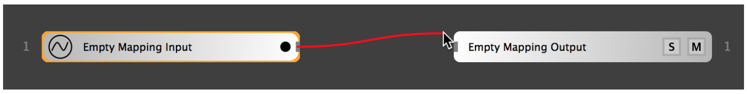
## A.3. Mappings

Mappings are how gestural inputs are connected to musical MIDI or OSC output, and multiple mappings can be contained within each Scene. These are found in the mapping panel of the Glover project (Figure A.1.7).

1. To create new input and output objects, click on the "I" and "O" buttons in the Glover project toolbar (Figure A.9a).
2. To connect an input to an output, click and drag from the grey tab on the right of an input object, and connect it to the corresponding tab on the output object (Figure A.9b).



(a) New mapping buttons

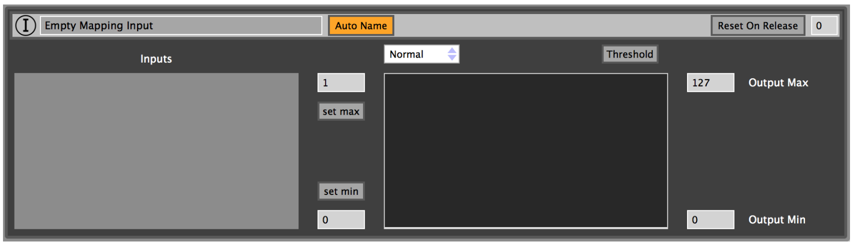


(b) Connecting mapping objects.

Figure A.9.: Mapping Controls.

### A.3.1. Inputs

1. Once you have created an input, click on the new input object. This brings up the new mapping input's inspector panel (Figure A.10).
2. To add a new input parameter, drag a device input from from the "Device Inputs" section into the "Inputs" box of the mapping input inspector (Figure A.11).



127 Output Max

0 Output Min

Figure A.10.: Mapping Input Inspector Panel.

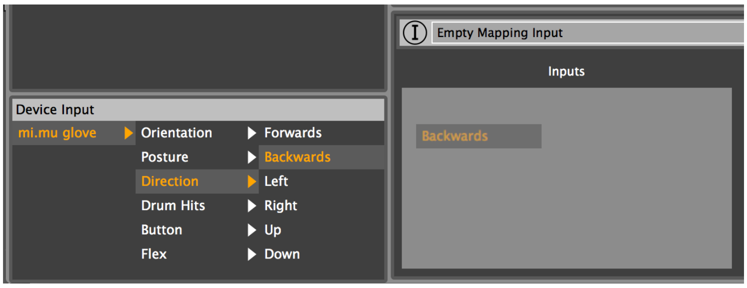
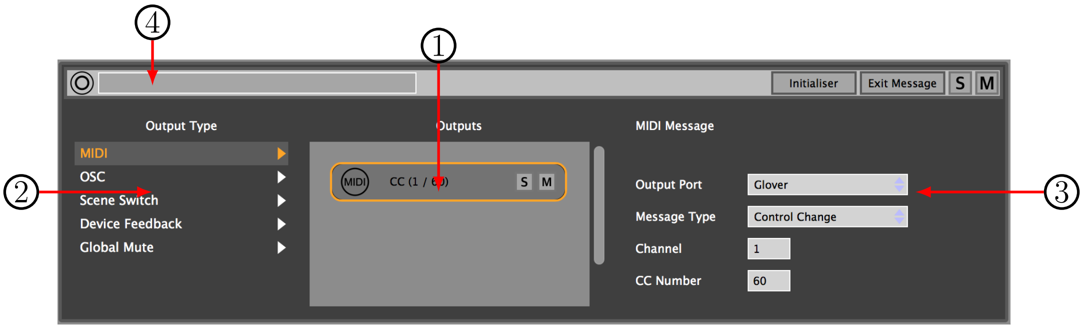


Figure A.11.: Dragging inputs.

## A.3.2. Outputs

The output objects are for passing the gestural data from the input objects to external MIDI outputs, or used to change settings in Glover or provide device feedback from Glove controls.



Output Port Glover

Message Type Control Change

Channel 1

CC Number 60

Figure A.12.: Output Inspector Panel.

1. Click on the output object you have created. This brings up the output inspector (Figure A.12).
2. To set the Output Type, drag the desired Output Type (Figure A.12.2) into the Outputs box (A.12.1).

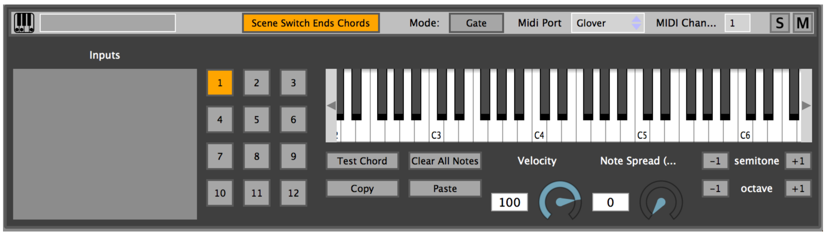
3. Change the settings of the output in the right hand section (Figure A.12.3).
4. Rename the output to something appropriate, e.g. "filter cutoff" (Figure A.12.4).

As well as MIDI you can send OSC messages, and you can control processes in Glover from your gestures, such as changing a scene, or triggering feedback on the gloves themselves (vibrations and LEDs).

### A.3.3. Instruments

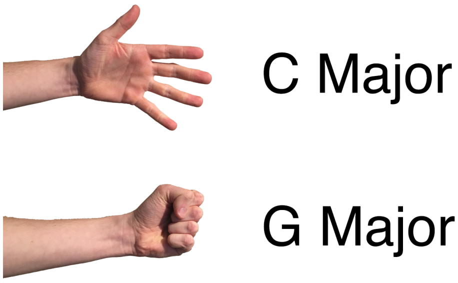
The third mapping option are the instruments. There are two instruments, the *chord machine* (Figure A.13a) and the *note matrix* (Figure A.14a).

The Chord Machine uses qualifiers to play multiple notes at once, while the Note Matrix splits a movement into a series of thresholds that triggers notes when they are crossed.



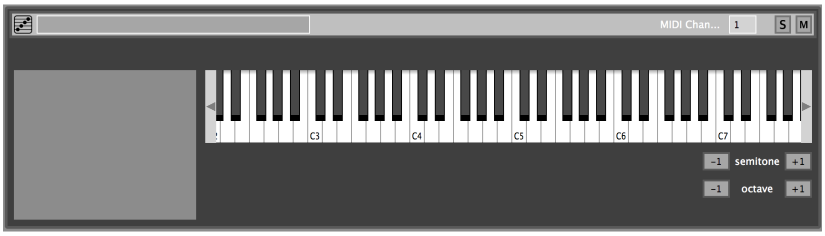
Mode: Gate
Midi Port Glover
MIDI Chan... 1
Velocity 100
Note Spread (... 0

(a) Chord Machine Inspector Panel.

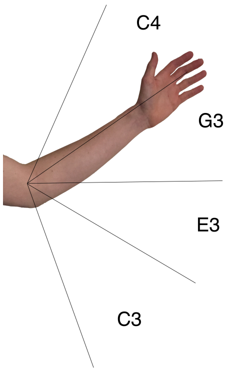


(b) An example of how the Chord Machine works. Here, an open hand posture is mapped to C Major chord, and a fist is mapped to a G Major.

Figure A.13.: The Chord Machine.



(a) Note matrix Inspector Panel.



(b) An example of how the Note Matrix works. Here, the pitch movement is split over the notes in a C Major chord.

Figure A.14.: The Note Matrix.

1. Click the the piano button (Figure A.9a) and choose the desired instrument. This will appear in the Instrument column on the right hand of the mapping panel (Figure A.15).

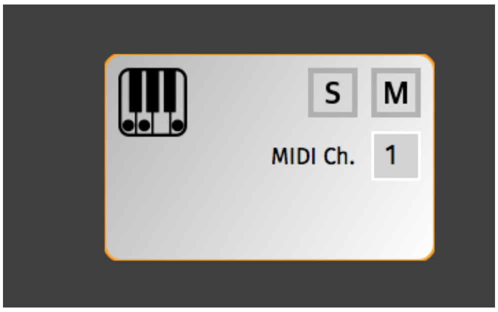


Figure A.15.: A new instrument in the mapping panel.

2. Click on the instrument to bring it up in the inspector.
3. Drag input parameters from the Device Inputs panel into the grey inputs box of the instrument.
4. Select the desired output notes on the piano roll.

## A.4. Organising Mappings

Mappings in Glover are organised into *scenes*, which are in turn organised into *arrangements* (Figure A.16). These help organise mapping strategies into suitable pieces, such as each song, or each section of a song (verse, chorus etc.).

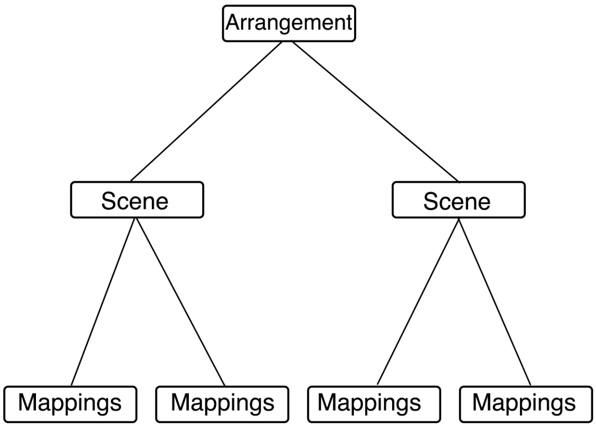


Figure A.16.: Mapping heirarchy.

### A.4.1. Scenes

1. To create new scene, click on the "+" button in the scene section of the Glover project (Figure A.17).

Figure A.17.: Scene overview panel.

2. Select the new scene, you will notice that the mapping panel is now empty, as it shows the mappings in this new empty scene.

3. To rename a scene, click on the scene's name in the Glover project toolbar.
4. To delete a scene, select it in the Scenes overview panel (Figure A.17) and press backspace.
5. Scenes can be switched between, and there is a mapping output option ("Scene Switch") which can be used to trigger scene switches from gestural actions.

### A.4.2. Arrangements

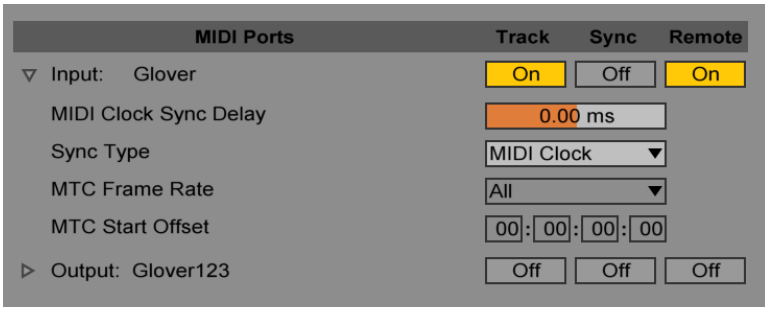
1. To create a new arrangement, click on the "+" button on the arrangement selector in the toolbar (Figure A.18).

Figure A.18.: Arrangement selector.

2. To rename an arrangement, click on the arrangement's name in the Arrangement selector.
3. To delete an arrangement, select the desired arrangement in the selector and click on the "x" button in the arrangement selector panel. Note: a Glover project must contain at least one arrangement, and so the last one cannot be deleted.

## A.5. Sending MIDI to Ableton Live

Occasionally Ableton Live will not seem to respond to MIDI information from Glover. In Live, go to Live -> Preferences and click on the Link MIDI tab. Make sure that Glover is available as a MIDI port, and make sure Remote is On (Figure A.19).



| MIDI Ports | Track | Sync | Remote |
|---|---|---|---|
| ▽ Input: Glover | On | Off | On |
| MIDI Clock Sync Delay | 0.00 ms | | |
| Sync Type | MIDI Clock | | |
| MTC Frame Rate | All | | |
| MTC Start Offset | 00:00:00:00 | | |
| ▷ Output: Glover123 | Off | Off | Off |

Figure A.19.: MIDI settings for Glover in Ableton.