Gesture-Sound Causality from the Audience's Perspective: An Investigation of the Influence of Mapping Perceptibility on the Reception of New Digital Musical Instruments

Gesten-Klang Kausalität aus der Zuschauerperspektive: Eine Untersuchung von dem Einfluss der Wahrnehmbarkeit des Mappings auf die Rezeption von Neuen Digitalen Musikinstrumenten

Gina Emerson

[redacted]

MA Musikwissenschaft
Erstgutachterin: Jun-Prof. Dr. Jin Hyun Kim
Zweitgutachter: Dr. Hauke Egermann
Institut für Musikwissenschaft
Humboldt-Universität zu Berlin

Contents

List of Figures

List of Tables

Introduction

From the *Radio Baton* to the *Reactable,* from MIDI keyboards to *The Hands*, the number, diversity and complexity of digital musical instruments (DMIs) has increased rapidly over the past four decades. Despite the range of contrasting artistic or commercial aims that stand behind their creation, these inventions all have one defining purpose; they are *systems for controlling digital sound synthesis* (after Wessel et al. 2002). In the achievement of this purpose, they typically operate as part of a multi-step process that begins with some form of input from a human performer and results in audio output, usually via algorithmic sound synthesis on a laptop computer (Fig. 1). This form of sound production, one that involves a separation of sound control and sound generation elements, brings about two ways in which DMIs starkly contrast their traditional, acoustic counterparts. Firstly, they are not sources of sound in themselves; it is only through the triggering of sound-producing algorithms that they can be used for sound generation. Secondly, DMIs can have any number of different input-output relationships, as determined by the mapping design (see Section 2.2. below). This possibility for endless reconfiguration grants DMIs a very transitory sense of identity in comparison to acoustic musical instruments and in turn often means that they have no single, intrinsic timbre.

For instrument designers and performers, DMIs afford limitless possibilities and flexibility for music-making. They also represent a way of returning a sense of bodily presence to the act of producing and performing electronic music, this for many being the aim behind the whole project of creating new digital musical instruments (see Miranda and Wanderley 2006, Leman 2008, Kim 2012, Bovermann et al. 2014). In contrast, what DMIs offer audience members is less immediately clear. The separation of control and sound generation elements can mean that the performer's physical interaction with the device frequently does not appear to correlate directly or as expected with the sonic output, thus making it potentially difficult for spectators to discern how gestures are translated into sounds.

The aim of this thesis is to explore how a potential lack of gesture-sound causality impacts spectators' reception of DMIs through empirical research carried out within the 3DMIN (*Design, Development and Dissemination of New Musical Instruments*) project at the Technical University, Berlin. Chapter One further introduces the term 'digital musical instrument' and provides a brief overview of the types of DMIs in existence. Chapter Two follows the DMI sound generation process step by step, defining additional relevant terminology. In Chapter Three, the focus shifts to the reception of DMIs, with a critical review of the existing literature in this area and a consideration of the useful insights from audiovisual perception research. Chapter Four presents the empirical study, which is followed by some concluding points and suggestions for further research.

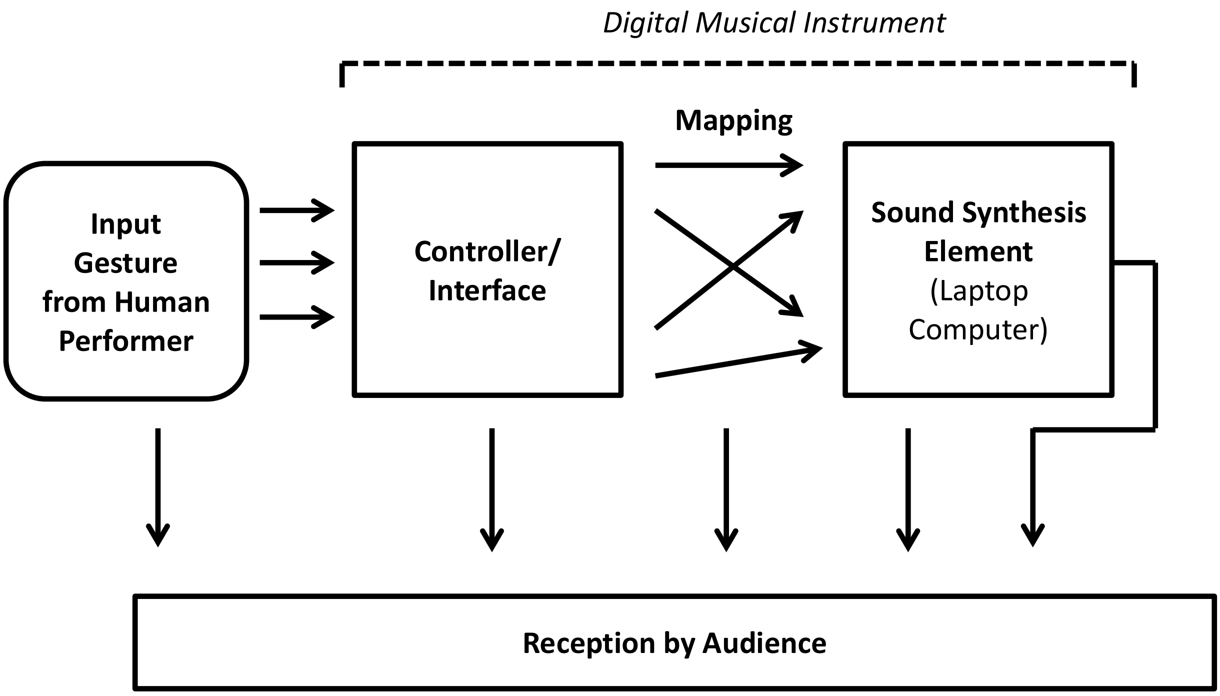

Fig. 1: The process of sound generation with a Digital Musical Instrument.

1. What is a Digital Musical Instrument (DMI)?

1.1 DMIs and Instrumentality

The differences between conventional instruments and digital musical instruments beg the question as to whether the term 'musical instrument' is even appropriate in reference to DMIs, leading to the issue of their status as instruments or their level of 'instrumentality' (Burrows 1987). Within the musicological discourse on the definition of 'musical instruments', there is a general consensus that the term is fluid but centred on the potential for an object to be used for sound generation; practically any object can be used for sound production and thereby become a musical instrument, however momentary this usage might be (Hornbostel and Sachs 1914; Burrows 1987; Alperson 2008; Bense 2013). Objects such as pianos, violins and flutes tend to be used for sound production more often than for other purposes and therefore have more permanently acquired musical instrument status. Further defining criteria for musical instruments that have been proposed include considerations of the intentions behind the production of an instrument (i.e. a musical instrument is an object *created for* sound production; Alperson 2007, 38) and the importance of human physical effort and skill in interaction with musical instruments (i.e. a musical instrument is an object that holds cultural value which defines how it should be used), which in turn leads to evaluation-related concepts such as virtuosity (Auslander 2009).

If the ability to produce sound with the object in question is a necessary feature of a musical instrument, then, strictly speaking, the majority of DMIs do not fulfil even this basic criterion of instrumentality. When their respective sound sources, which is most cases a laptop running sound synthesis software, are detached, DMIs are not necessarily any more suited to sound production than any other object. Their lack of a consistent, basic timbre further prevents them from being permanently distinguishable from other instruments, a criterion that I take to be intrinsic to instrumentality and that has been discussed within culturally-centred definitions of musical instruments (Alperson 2007; Cance et al. 2009).

In light of this, what would be alternatives to the term 'digital musical instrument'? Even though it typically refers to the control element of the device (see Fig. 1), the term 'interface' (also 'Tangible Acoustic Interface' or 'TAI') is frequently used interchangeably with 'DMI' and thus can be seen as an option, one which crucially avoids the baggage of instrumentality. It is the chosen noun for the title of the *New Interfaces for Musical Expression* (NIME) conference, the main annual event for those working in this field and well reflects the general purpose of DMIs, namely that of enabling human interaction with something external (in this case, digital sound synthesis on a computer). In a similar manner, the term 'controller' presents itself as a possibility, having already appeared eighty times in the titles of papers in NIME conference proceedings from 2001 to 2015.[1] Despite their popularity, these two options are potentially reductive as they do not acknowledge the sound-producing elements and accordingly do not take in the entirety of the device. 'Controller' also draws a parallel with the industry MIDI controllers to which it usually refers and therefore fails to encompass the variety of devices that have been invented. Another option with a number of proponents is 'virtual musical instrument' or 'VMI', which can be considered a precursor to 'DMI' and was developed precisely in response to this issue of sound production; if these devices do not produce sound, then they do not really *exist* as instruments (see Mulder 1994; Goto 2000; Bense 2013). This term does not, however, sufficiently account for the physicality that most DMIs do have; while there are many software instruments to which it is well-suited, DMIs tend to be physical objects that are tangible, not 'imagined' or 'virtual'.

A further potential direction which would successfully take into account the composite nature of these devices would be use to terms such as musical 'system' or 'network'. Wessel, Wright and Schott (2002, 1), for instance, refer to digital musical instruments as 'systems' comprising of an interface, mappings and an audio output component and the term has also been taken up in part by designers (e.g. Hsu and Sosnick 2009; Van Nort et al. 2012). A similar notion of DMIs as systems can be found in Gurevich and Fyans' (2011b) focus on 'digital musical interactions', which expands the concept of the overall system to include the human performer's actions, a concept that is paralleled by Marc Leman's idea of the musical instrument as a 'mediator' in musical communication (Leman 2008, 160-183).

Yet, despite its shortcomings and the number of available alternatives, the term 'digital musical instrument' has become accepted as the standard; 212 NIME papers make reference to 'instruments' in their titles and it is furthermore the term employed in Miranda and Wanderley (2006), the first monograph devoted entirely to the topic of DMIs. In spite of

[1] Figure calculated by running keyword searches through the NIME proceedings archive at http://www.nime.org/archives/, last accessed 12.10.2015

its potential incompatibility with the criteria of instrumentality, 'DMI' satisfactorily sees these devices as wholes, fully inclusive of sound control *and* generation elements. Appropriating the term 'instrument' also helps to position DMIs within the mainstream discourse on acoustic instruments, encouraging their wider acceptance. It is for these reasons that I adopt the term here and use 'controller' and 'interface' when referring specifically to the sound control elements of the instrument.

1.2 An Overview of Existing Digital Musical Instruments

While there are various important electronic instruments that can be considered predecessors to digital musical instruments, such as the theremin, the Trautonium and the Ondes Martenot, the historical development of DMIs truly begins with the invention of digital sound synthesis. The first computer program designed for digital sound generation, *MUSIC 1*, was created by Max Mathews at the Bells Laboratory at Stanford University in 1957, spawning five later versions and a variety of related software, such as Csound and CMusic (Roads and Mathews 1980; Mathews 1997). MUSIC 1 was capable of generating a single waveform with equal attack and decay values, for which only the pitch, amplitude and duration could be specified (Roads and Mathews 1980, 15). Whilst developing the later versions of MUSIC, Mathews also experimented with the creation of input devices for his new digital sound synthesis program. This resulted initially in the graphical input interface, *Graphic 1*, which was developed in 1965 and then followed by the *GROOVE* (Generated Realtime Operations On Voltage-controlled Equipment) system in 1970. This featured a number of novel input devices, such as joysticks and a wand for remote control (a predecessor of the *Radio Baton*, see below; Park and Mathews 2009, 13).

With the establishment of the MIDI (Musical Instrument Digital Interface) protocol in 1982 and the development of microcomputers shortly thereafter, it became far easier to transmit information between musical hardware and software (Manning 2013, 263-279). Despite this advancement, the keyboard interface that had dominated analogue synthesis largely prevailed, as demonstrated by the commercial success of such digital keyboard synthesisers as the Yahama DX7 from 1983 and the Korg M1 from 1988 (Miranda and Wanderley 2006, xix). Indeed, the keyboard synthesiser/MIDI controller still dominates the industry landscape, presumably due to the fact that the keyboard interface is simply familiar to many, given its status as a music education staple. Outside of the commercial realm, however, it was with the founding of more institutions and university departments for computer music research that the field of DMI development began to grow in the late 1980s and 90s,[2] leading up to the first NIME conference in 2001.

In the brief overview of types of DMIs below, I adopt Miranda and Wanderley's categorisation according to similarity with existing acoustic instruments (see Miranda and Wanderley 2006, 19-21). This leads to three overarching categories of instruments: augmented instruments, DMIs with controllers inspired by acoustic instruments (this combines Miranda and Wanderley's second and third categories into one for simplification) and DMIs with novel, alternative interfaces.

1.2.1 Augmented Instruments

The term 'augmented instruments' refers to acoustic or electric instruments that have been 'extended' through the addition of sensors and microphones that pick up aspects of the sound production that would usually not be audible (Miranda and Wanderley 2006, 21-22). The signals from the sensors can also be used to control digital sound synthesis directly. One of the earliest and most widely used examples of an augmented instrument is Yamaha's *Disklavier*, which first came on the market in 1987 and is in essence an acoustic piano equipped with a variety of sensors, enabling much extended functionality, above all the ability to self-record.[3] With the development of the *PianoBar*, an attachable device for producing MIDI recordings on an acoustic piano created by Robert Moog and Don Buchla in 2003, any acoustic piano could be spontaneously augmented in a manner similar to the *Disklavier*.[4] The most extensive work done on the production of augmented string instruments comes from Tod Machover's group at the Massachusetts Institute for Technology (MIT). The group's *Hyperinstruments* project began in 1986[5] and led to the development of a *hypercello*, *hyperviola* and *hyperviolin*, a series of sensor-augmented string instruments that each allow the performer to play self-accompanied by sounds synthesised out of the basic acoustic sound (Machover 1992). The *hypercello*, in particular, received considerable attention when performed on by Yo-Yo Ma, for whom Machover's 1991 composition, *Begin*

[2] Not to overlook the contributions by institutions already in existence by this point such as STEIM (S*tudio for Electo-Instrumental Music*) in Amsterdam and IRCAM (*Insitute de Recherche et Coordination Acoustique/Musique)* in Paris.
[3] http://usa.yamaha.com/products/musicalinstruments/keyboards/disklaviers/e3_series/?mode=series# , last accessed 10.10.15
[4] http://www.moogmusic.com/legacy/bob-moog-timeline, last accessed 10.10.15
[5] It is now part of the MIT Media Lab: https://www.media.mit.edu/research/groups/1450/hyperinstruments, last accessed 28.10.15.

Again Again... was written (see Fig. 2a).[6] There are also a number of notable wind and brass augmented instruments (see Impett 1994, Schiesser and Schacher 2012).

1.2.2 Instrument-inspired Controllers

This category includes DMIs with controllers that are in some way related to an existing acoustic instrument, usually in that the playing action of an acoustic instrument is taken as the basis for the interface design. The majority of industry digital synthesisers and MIDI controllers fall into this group, ranging from keyboard to guitar (e.g. the Yahama EZ -G) to wind controllers (e.g. the Yahama WX series or the Akai EWI/EVI devices) or even to controllers that combine the interfaces of multiple instruments in one (e.g. the Artiphon *Instrument 1*, see Fig. 2d).[7] These devices often enjoy wide appeal because they allow players of the respective acoustic instrument to apply their learned playing technique when synthesising sound digitally, making sound production more intuitive. Notable examples of novel DMIs that draw inspiration from but do not explicitly model an acoustic counterpart include the *Sequential Drum* (Mathews and Abbott 1980), the *Accordiatron* (Gurevich and von Muehlen 2001) or the *GuitarBot* by Eric Singer.[8]

1.2.3 Alternative DMIs

Naturally, the number of DMIs that bear no resemblance to an existing instrument and therefore fall into this final category is so large as to be beyond documentation. There are, however, a number of discernible trends in the design of alternative sound controllers, the three most prominent being: 1) devices which are operated via directly touching a surface or screen; 2) devices that are hand-operated through some form of remote control; and 3) DMIs with wearable controllers that are attached in some way to the body of the performer. Prominent examples from the first category include the *Lemur*, a popular touch screen instrument from the French company, JazzMutants, which was first unveiled at IRCAM in 2004.[9] A more recent development is the *Reactable*, a circular touch screen interface that is manipulated through a combination of direct touch and the movement of differently patterned cubes on the surface and which seeks to grant electronic music-making an interactive, ludic quality (Fig. 2b). The *Reactable* was created in 2005 by a team from the

---

[6] http://www.nytimes.com/1991/08/17/arts/review-music-yo-yo-ma-and-his-new-hyper-cello.html, last accessed 10.10.15
[7] http://artiphon.com, last accessed 10.10.15
[8] http://www.singerbots.com, last accessed 10.10.15
[9] http://www.jazzmutant.com/behindthelemur.php, last accessed 11.10.2015

Pompeu Fabra University in Barcelona and counts the experimental pop artist Björk among its most famous proponents.[10]

In the second category of hand-held or hand-controlled devices belong two of the best-known examples of DMIs, Mathews and Bole's *Radio Baton* from c. 1987 (Mathews 2000; Fig. 2c) and Don Buchla's *Buchla Lightning* from 1991.[11] Both devices operate via the movement of hand-held batons or wands, which emit a signal of a specific frequency that allows a receiver unit to track their movements (see Miranda and Wanderley 2006, 38-41 for a more detailed discussion of the differences between the two). A more recent instance of a long-range hand-controlled DMI is Naonext's *Crystal Ball* from 2012, the operating mode of which strongly mirrors that of the theremin.[12]

The earliest example of a wearable-like controller is Michael Waisvisz's *The Hands* from 1984 (Fig. 2e). Arguably a hand-held and not strictly a worn device, *The Hands* are a pair of aluminium plates shaped to fit naturally into the hands with embedded sensors measuring the distance between the hands and the degree of hand-tilt (Waisvisz 1985). *The Hands* paved the way for other, more clearly wearable DMIs such as Laetitia Sonami's *Lady's Glove*, the first version of which dates from 1991,[13] and Tomomi Adachi's *The Right Hand* from 2005/2006.[14] The wearable concept was extended by such full torso or full body sensor instruments as Suguru Goto's *BodySuit* from around 1997,[15] the *Conductor's Jacket,* also by Machover's team at MIT (Nakra 2000), and Adachi's *Infrared Sensor Shirt* from 2004.[16]

Having established what is and what is not meant by 'digital musical instrument' and the range of devices to which this term applies, the steps involved in the DMI sound production process can now be looked at in more detail.

[10] http://www.reactable.com/history/, last accessed 11.10.2015
[11] http://www.buchla.com/history/, last accessed 11.10.2015
[12] http://www.crystall-ball.com/UK/index.html, last accessed 11.10.2015
[13] http://sonami.net/ladys-glove/, last accessed 11.10.2015
[14] http://www.adachitomomi.com/n/activities.html, last accessed 11.10.2015
[15] http://suguru.goto.free.fr/Contents/Works/BodySuit/BodySuit-e.html, last accessed 12.10.15
[16] http://www.adachitomomi.com/n/activities.html, last accessed 11.10.2015

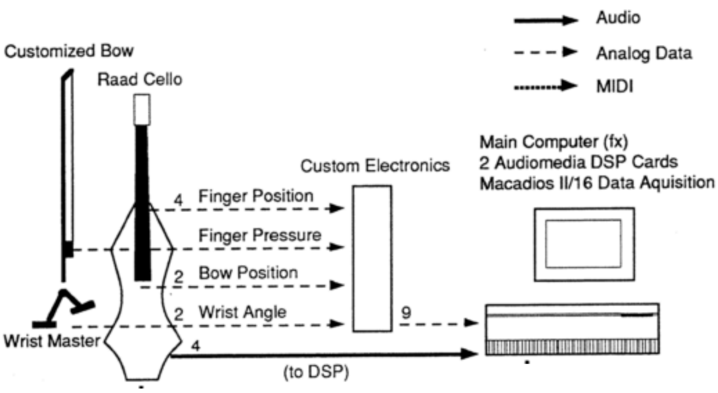

Fig. 2a): Excerpt from the *Hypercello* system diagram
(Machover 1992, 51)

b): The *Reactable Live!* S6
(http://reactable.com/live/, last accessed 28.10.15)

c): Max Mathews with the *Radio Baton, c. 1984*
(http://www.computerhistory.org/revolution/computer-graphics-music-and-art/15/222, last accessed 28.10.15)

d): The various playing configurations of the *Instrument 1* by Artiphon.
(http://techcrunch.com/2015/03/07/the-crowdfunded-artiphon-instrument-1-mixes-high-and-low-musical-tech/, last accessed 28.10.15)

e): Michael Waisvisz with *The Hands*, c. 1995
(*http://www.perfomap.de/map3/kapitel1/steimgeschichten, last accessed 28.10.15)*

2. From Gesture to Sound: The Process of DMI Sound Production

2.1 Gesture

The process by which sound is produced with a DMI begins with the performer's input. This most commonly takes the form of a 'gesture', a term that features ubiquitously in the literature on DMIs, appearing in 62% of all NIME conference papers from 2001 to 2013 (Jensenius 2014). A common starting point for defining 'gesture' is to contrast it with 'posture'; it denotes movement, not the static holding of a position (Miranda and Wanderley 2006, 4). However, gestures are viewed as distinguishable from mere movements in that they carry meaning; they are a means of communication and must therefore be interpreted by a perceiver as such. This definition of gesture is prevalent in the fields of linguistics, communication studies and psychology (e.g. Kendon 2004) and has been applied to music in a number of texts from the past ten years. Anthony Gritten and Elaine King, in their collected edition on music and gesture, define the term as 'a movement or change in state that becomes marked as significant by an agent' (Gritten and King 2006, xx), a similar interpretation to that of Marc Leman and Rolf Godøy, who suggest that a movement can only be considered a gesture if it is 'in some way a carrier of expression and meaning' (Leman and Godøy 2010, 7). This suggests that any one music-related movement could be construed alternately as meaningful or meaningless on different occasions. Therefore, gesture cannot truly be a fixed concept with a clear definition. Considering a movement to be a 'gesture' is simply the prerogative of the perceiver, it is not an intrinsic property of the movement itself.

It is perhaps due to the undefinable nature of gesture that the term has been quite generally employed by musical gesture researchers when establishing categorisations of different of musical gestures. Cadoz and Wanderley (2000), drawing on the work of Delalande (1988) propose a three-level musical gesture classification system, consisting of *effective*, *accompanist* and *figurative* gestures (ibid., 77-8). The first category includes gestures that are necessary to create sound, such as the bowing of a violin or the depression of a key on a piano. Accompanist or accompanying gestures tend to co-occur with effective gestures, but are not required for sound production; expressive body movements would be included here. Finally, figurative gestures encompass movement *within* the music, such as a change in melodic contour or harmony or as driven by a rhythmic figure. This latter notion of musical gesture essentially refers to how music can appear to reflect movement in physical space and has its roots in Baroque discourses of music and rhetoric (Johnson and Larson 2003; Hatten 2004; Schneider 2010). It has also generated theories relating to the gestural affordances of sounds that have in turn been applied in the establishment of mapping strategies (Godøy 2010, see the discussion of Caramiaux et al. 2014, pg. 14). Jensenius et al. (2010) have expanded Cadoz and Wanderley's typology by introducing further categories of functional, non-figurative gestures. They add *communicative* gestures, replace accompanist gestures with *sound-facilitating* gestures and introduce the term *sound-accompanying* gestures to refer to gestures that trace or imitate sound-producing movements (ibid., 24), establishing overall a finer-grained picture of musical gestures.

Considering the definitions of gesture presented above and their insistence on the transmission of meaning, I would argue that it is only the communicative gestures (and possibly also some examples of sound-accompanying gestures) in Jensenius' typology that qualify as gestures, as they have the purpose of carrying meaning. I would propose that the term 'action' is in fact more appropriate than 'gesture' where sound-producing movements are concerned. The action of bowing a stringed instrument carries no semantic meaning; it is executed purely for the production of sound. 'Action' does also appear more appropriate when referring to the movements carried out by performers with DMIs. While some wearable controllers do use communicative gestures for sound production, for the majority of interfaces the primary mode of physical interaction are actions comparable to those made with acoustic instruments (striking, plucking, pressing, blowing, bowing etc.), movements which are not carriers of meaning. Although I do consider 'action' a better terminological fit, I shall continue to use gesture here so as to maintain a clear connection to existing research.

2.1.1 Alternatives to Gestural Control

There exists a number of other possible alternatives to the use of physical movement in conjunction with a device for sound synthesis control. One such trend is *expressive gesture mapping,* the use of facial expressions and other semantically loaded movements as pure gestural input to control sound synthesis processes without a controller or intermediate device. The most developed system in this field is *EyesWeb*, a hardware platform featuring a range of different video camera and sensor systems combined with software for gesture recognition, extraction and sound synthesis (Camurri et al. 2000; 2007). Its most recent version, *EyesWeb XMI* (eXtended Multimodal Interaction), has been applied in a number of interactive installations, including a system that enables users to control a virtual orchestra with their movements (Camurri et al. 2007). Similar expressive gesture capture systems include the *Interactive Music Head* (Ng 2002), Rokeby's *Very Nervous System*[17] and STEIM's *BigEye*.[18] There are also a number of further alternatives to gestural control, such as Stefano Fasciani's system for the voice control of sound synthesis (Fasciani and Wyse 2012) and explorations into the use of biosignals, such as brain activity (via electroencephalogram) or heartbeat (via electrocardiogram) as input (see Miranda and Wanderley 2006, 173-218 for an overview of these approaches).

2.2 Mapping

Once an input gesture or other form of input has been decided upon and the appropriate capture method found,[19] it has to be linked to the synthesis parameters it is to control. The process of establishing these links between input and output is called *mapping* and is fundamental to the creation of the instrument's identity. As Hunt et al. (2002) maintain, mapping is far from a trivial aspect in DMI design as it impacts, among other things, 'the psychological and emotional response elicited from the performer' (ibid., 1), the aspect that forms the focus of the empirical work presented later in this thesis.

The majority of DMIs employ mapping strategies that fall into the category of *explicit mappings*, mapping schema that do not change of themselves once implemented and do not involve any form of machine learning (this is in contrast to *implicit mappings*, see below). The three main configurations for explicit mapping strategies are generally held to be: 1) *one-to-one* mappings, in which one input parameter controls one synthesis parameter; 2) *one-to-many* or *divergent* mappings, in which one input parameter relates to several synthesis parameters; and 3) *many-to-one* or *convergent* mappings, in which several input parameters map onto one synthesis parameter (Rovan et al. 1997; Kim 2012; Miranda and Wanderley 2006, Fig. 3).

Hunt et al. (2000; 2003) conducted a series of qualitative experiments to test the playability and learnability of differently-mapped devices. Participants were given the task of accurately reproducing a set of audio excerpts of increasing complexity played by a computer on three different interfaces: 1) a collection of four on-screen sliders, controlled by a mouse; 2) a set of four hardware sliders to be controlled physically by the users; and 3) a multiparametric device consisting of a mouse and a set of two physical sliders (Hunt and Kirk

[17] http://www.davidrokeby.com/vns.html, last accessed 28.10.15
[18] http://steim.org/2012/01/bigeye-1-1-4/, last accessed 12.10.15
[19] See Miranda and Wanderley 2006, Ch. 3 for an overview of the different sensor methods in usage.

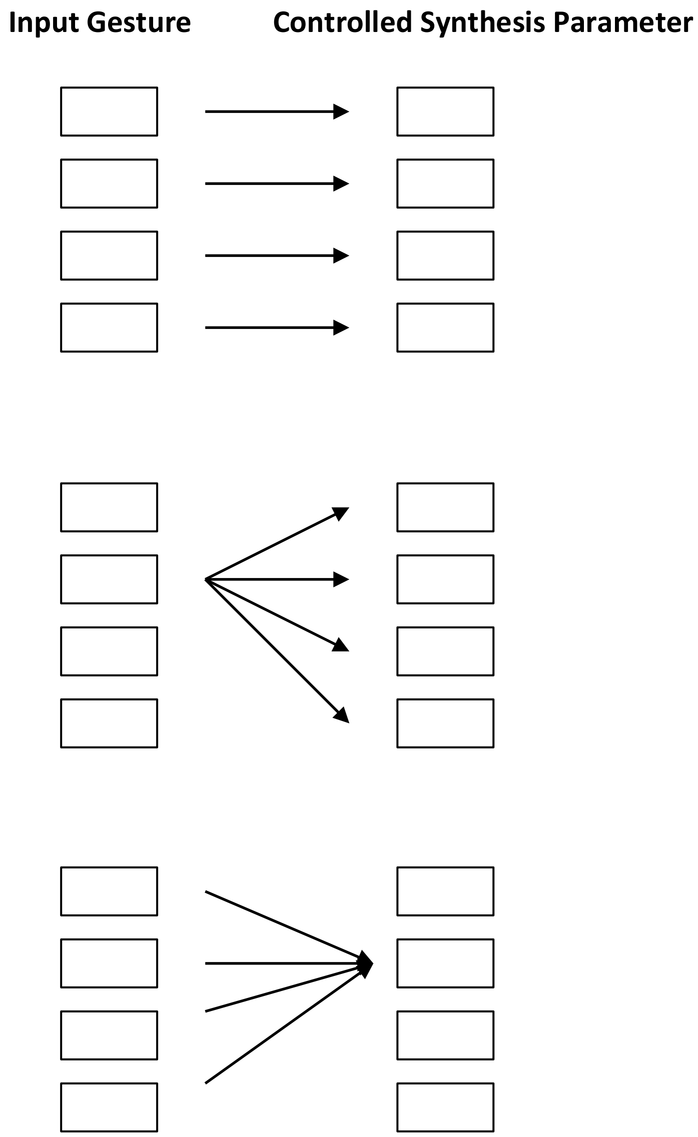

Fig. 3: Explicit mapping strategies. *(Top: one-to-one, Middle: one-to-many, Bottom: many-to-one)*.

2000, 236-9). The first two setups were one-to-one mappings with each of the four sliders controlling either the pitch, timbre, volume or the panning. In contrast, the multiparametric device featured several many-to-one configurations with differing combinations of mouse and slider control. This interface furthermore required continuous energy input from the user; sound was only produced when the mouse was moved. The accuracy of the participants' reproductions was recorded on multiple occasions over several weeks. The test scores for the multiparametric device were consistently better than the other two, except for the very simplest tasks (Hunt et al. 2003, 3) and the scores for this device improved as the tasks increased in complexity. The authors suggest that that a more complex interface that logically couples parameter control can indeed make harder tasks simpler to perform (ibid.). Qualitative evaluations collected from the participants also suggested that the many-to-one device was more intuitive and satisfying, largely due to the input of continuous energy that it required, although it involved a steeper learning curve. The simpler set-ups were met with greater frustration, a larger cognitive load and were overall described as being more limiting to expression (ibid.). The results confirms earlier support of convergent mappings in terms of the expressive potential they offer (as in Rovan et al. 1997, 69, see also Doornbusch 2002 and Dobrian and Koppelman 2006). A series of experiments conducted by Luciani et al. (2009) with a virtual bowed string have further revealed that continuous energy input and haptic resistance in the form of greater friction allow for the more successful completion of a task involving changes in bowing direction.

Aside from this, there have been few systematic empirical investigations comparatively exploring the advantages and disadvantages of different mapping strategies. There are, however, numerous models for approaching mapping as a designer that have arisen out of a mixture of empirical research and artistic exploration. Levitin et al. (2002), for example, propose that mapping strategies should 'exploit some intrinsic property of the musicians's cognitive map' (ibid., 183) and ground themselves in connections between action and sound that exist in the real-world, such as more force when breathing and striking resulting in a louder sound (see also Wessel et al. 2002 and Fels et al. 2002).

A similar approach has been developed by Caramiaux et al. (2014b), which also incorporates an *implicit* mapping strategy; it uses adaptive learning in the form of real-time neural network processing (see also Lee et al. 1991; Arfib et al. 2002; Merrill and Paradiso 2005). Their concept of 'mapping through listening' draws on a series of experiments carried out to ascertain how individuals describe sounds gesturally through asking participants to spontaneously perform gestures to audio clips (Caramiaux et al. 2010; Caramiaux et al. 2014a). The results suggest that when the source of sound is perceptible, participants perform actions that mimic those required to produce that sound. When this is not the case, gestures that trace a parameter of the sound (e.g. melody, dynamics etc.) tend to be enacted (Caramiaux et al. 2014b, 38). Based on these results, the authors have developed a system that plays back certain sounds from a database according to the gestures executed by the performer. These sounds can then, depending on which mode the system is in, be modulated in real-time via a shape-matching algorithm that continuously searches for the acoustic pattern that best matches the gestural pattern, before syncing the two temporally (ibid., 40). The mapping is therefore a reflection of the gestures that a performer would naturally make to the sound, a flexibility that is supported by the machine learning aspect as it adapts to form new mappings between gesture and sound as the performer conceives of them. Although the authors do not consider this, such a mapping of gestural affordances would potentially be more satisfying for and more readily understood by an audience (see Leman 2008, 164-168 for a similar argument).

Another noteworthy approach to the choice of mapping design comes from Alberto de Campo, who has a developed a patch (a defined collection of settings) for use in the *SuperCollider* programming language that allows for changes in mapping configuration during performance (de Campo 2014). The *Influx Patch* stems out of de Campo's 'lose control, gain influence' (LCGI) concept, which advocates the implementation of design elements that actively reduce the amount of direct control the performer has over the instrument, an endeavour that has the potential to invigorate the creative process, especially when applied to mapping:

> [...] In the chain human > physical device > gestural data > mapping > output process, it is precisely the mapping process that is the easiest to expose to change during performance... As a performer, one can choose to put oneself into situations where a performance system will surprise not only the audience, but also oneself. In other words, we argue that in concert situations it may be preferable to lose control if this loss can be made meaningful by gaining influence.
>
> (ibid., 218)

While it remains to be seen whether the randomisation would surprise an audience in a positive manner, the LCGI principle is certainly an interesting proposal for performers, challenging them to listen carefully and react accordingly to new configurations.

In light of these diverse examples of mapping strategies, it seems that certain categories, especially that of 'explicit' vs. 'implicit', may not be as distinct as earlier accounts have implied. The *Influx Patch*, for instance, blurs the boundaries between an explicit and an implicit system; the mappings employed remain in themselves linear but the performer is naive to their new configuration. Since no adaptive learning is involved, however, LCGI-principled systems fall more on the explicit side of the mapping strategy spectrum, perhaps in fact representing an 'any-to-any' configuration. It has been suggested elsewhere that the term 'mapping' is in fact too simplistic and serves ultimately to limit artistic exploration. Joel

Chadabe (2002) argues that as DMIs grow more varied, 'the concept of mapping becomes more abstract and does not describe the more complex realities of electronic instruments' (ibid., 4) and suggests that the concentration on describing mapping strategies in the literature leads to a failure to describe the instrument as a whole (ibid.). While this may be the case, mapping still represents what is perhaps the most important feature of DMI design. There can, however, be no definitive description of mapping designs that covers all possibilities. The act of categorisation is also perhaps best left to instrument designer/performer, who will have his or her individual model, which may or may not be viewed by them as 'implicit' or 'explicit', linear or non-linear.

2.3 Sound Synthesis: A Brief Discussion of Musical Practice

I think the question which is going to dominate the future is now understanding what kind of sounds we want to produce rather than the means of usefully generating these sounds musically.

Max Mathews[20]

The final step in the DMI performance process around which I have structured this chapter is the sound generation itself. A complete overview of the various programming environments currently used for the digital sound synthesis in conjunction with DMIs is beyond the scope of this thesis, this can be found elsewhere (Kim 2012, 100-113; Collins 2009, 31-33). Instead, I would like to devote this section to a brief consideration of the relationship between instrument design, sound design and musical practice. There has been very little discussion of the kinds of sounds or music that DMIs could be used to create and little to no justification or explanation of the decisions made regarding sound design in the wider DMI design literature. This facet is also notably omitted from Miranda and Wanderley's 2006 monograph, a book that covers practically every other aspect of DMI practice. This could be due to the fact that it is not always clear who has the responsibility for making and justifying such decisions and who feels qualified to do this; DMIs encourage the blending of the roles of instrument designer, composer and performer, roles that have been kept much more distinct outside this field. For those individuals that consider themselves as taking on all three of these positions, the chosen approach to music-making is often improvisation. While this is most certainly a legitimate form of musical expression, it should still be preceded by clear decisions concerning the choice of musical structure, timbre and other parameters and aim to realise a specific artistic intention, just as a more fixed composition would. As Sergio Jordá (2004)

[20] Park and Mathews 2009, 19

similarly notes, creating a new musical instrument can be conceived of as being equivalent to creating new music, or at least of having the intention to do so, and thus, a greater importance should be placed on sound design to realise this aim.

Standing in contrast to the DMI improvisers are a number of DMIs that have had pieces specifically composed for them or that have been designed to play certain compositions. This obviously comes with a series of other challenges: how is it possible to compose a piece of music exploiting the idiosyncracies of a particular instrument when the instrument has no fixed timbre? Such works have tended to be produced by composers with a practical knowledge of instrument design and who therefore set out with the intention of developing a system that allows them to realise their musical ideas. This is, for example, the approach employed by Tod Machover, who has created various works for DMIs, such as *Jeux Deux* for Disklavier, orchestra and live visuals (2005) and the groundbreaking opera *Death and The Powers* (2010), which features a robotic chorus and a system that uses the live input from the singers to control other aspects of the setup, such as the visual effects (Jessop et al. 2011). Other examples of recent compositions for DMIs include Marije Baalman's *Gewording* ('Becoming') from 2013, which combines live coding and performance with a glove controller, Truus de Groot's duet for Crackle Box and Triomne Synth from 2014 and the piece *Balgerei* ('Tussle' or 'Quarrel'), also composed in 2014 as part of the 3DMIN project for an accordian-like instrument called *PushPull* (see Chapter 4 below). Encouraging the creation of compositions for DMIs would be a way of giving them a firmer position in contemporary musical practice, turning them into instruments with repertoires. This greater focus on the music to be produced with DMIs could in turn aid audiences in their reception, a matter that I shall return to in the concluding section.

3. Perceived Causality and its Importance in DMI Reception

3.1 Audiovisual Perception Research on Causality

Before introducing the existing literature on audience responses to DMIs, it is necessary to define more clearly what is meant here by causality. When experiencing a musical performance, spectators receive auditory and visual information and, on the basis of this, determine which kinds of actions by the performer result in which kinds of sounds; in short, *they perceive the live performance of cause-and-effect relationships*. In DMI performances, this relationship is determined by the mapping design as described above, which may or may not allow for the perceptibility of causal relationships. Several studies in the field of audiovisual perception research have explored the basis for the perception of causality between the visual and auditory streams, suggesting that it is largely the spatiotemporal correspondence of two events that enables humans to perceive one event as being caused by another (see Spence 2007 for an overview of the work in this area). If a visible action seems to precede an appropriate sound, humans combine or 'integrate' the two streams of information and perceive the action as having caused the sound. The successive event has to occur less than 200 msecs after the initial event in order for this to be the case (Guski and Troje 2003, 791). In addition, percussive sounds are more readily perceived as causal than sustained sounds (Schutz and Kubovy 2009, 1793). The question under investigation here, however, is whether causal stimuli are somehow aesthetically more pleasing to perceive than acausal stimuli. While there is little to no existing work directly on this topic, there are two strands of research that offer related insights into this possibility: 1) research on the general impact of visual information on the reception of musical performances; and 2) investigations into the reception of congruent or synchronous audiovisual stimuli.

3.1.1 Visual Information and the Perception of Musical Performance

Various studies from this first line of thought have demonstrated the fundamental principle that visual information accompanying the act of listening to music can influence what we hear and, at times, heighten our enjoyment of it. Schutz and Lipscomb (2007), for example, report an illusion in which the duration of short-length tones played by a marimba player were perceived by participants as longer when accompanied by a video of the performer making a longer performance gesture (see McGurk and Macdonald 1976, Schutz 2008 and Behne and Wollner 2011 for more audiovisual illusions). Concerning the communication of emotion and musical tension through performance, Vines et al. (2008) found that visual information can both increase and reduce perceived tension at different points in a piece, as revealed through an audio-only vs. visual-only vs. audiovisual paradigm involving ratings of performances of a Stravinsky clarinet piece. This was then extended to emotional responses; the intended emotion was shown to be most intensely perceived in the audiovisual condition (Vines et al. 2011; see also Dahl and Friberg 2007; Silveira 2013; Vuoskoski 2013). These results indicate that visual input contributes significantly to the perception of musical performance, from which it can be inferred that DMI performances that present conflicting or confusing audiovisual information might struggle to be effective at communicating musical intentions to an audience.

3.1.2 Insights from Multimedia Perception: Congruency Theory

A particularly relevant concept that has arisen from research on the audiovisual perception of multimedia is congruency theory, which relates to the human ability to judge the extent to which information from the auditory and visual streams occurs in synchrony. This most commonly takes the form of some kind of co-occurrent accent in both streams, such as the coincidence of the downbeat and a change of scene or other significant visual event. As part of her Congruence-Association Model, Annabel Cohen has proposed that congruent moments between the video and soundtrack in film attract greater attention from the viewer (Cohen 2013, 28). Furthermore, a number of related studies have demonstrated that temporally congruent audiovisual stimuli are given a more positive aesthetic rating by participants. Iwamiya et al. (2000) investigated this by presenting participants with short animated videos of a ball on grid, which were accompanied by a simple audio track featuring drums and a bassline. The downbeats of the audio track were temporally congruent with changes of shot in the video track for some clips and incongruent for others. The authors found that higher ratings of perceived congruency correlated with higher impressiveness ratings of the clips.

How do these insights relate back to the perception of causality? It is possible to suggest that congruent stimuli are aesthetically preferable because they are simpler to perceive; there is less conflicting information to be dealt with. This in turn implies that higher degrees of perceptible causality might similarly be aesthetically preferable. Clearer information can be expected to result in a higher level of understanding, which can be considered as the necessary basis for further judgements, a point I shall expand upon in the following section.

3.2 DMI Reception Research

The majority of existing research on DMIs has focused largely on the technical aspects of such devices, considering their attributes from the performer or instrument designer's perspective and thus evaluating them in terms of their potential expressivity and playability, as the literature cited in Chapter 2 does. More recently, papers regarding audience evaluations of DMIs have emerged, many of which touch upon the issue of gesture-sound causality. These fall into two groups: 1) research that draws on the issue of *liveness;* and 2) investigations into audiences' understanding of DMIs.

3.2.1 Liveness

The term *liveness* comes from the work of Philip Auslander (1999; 2009), who uses it to refer to a certain cultural value that live performance is deemed to have. Live musical performance has *authenticity*; it involves effort and skill and is thereby often considered to be experientially different, and in part more valuable, than listening to a recording (Auslander 1999, 73-128). The notion of liveness invokes many concepts that are relevant when considering audience perceptions of DMIs and has been taken up in DMI-focused research most prominently by John Croft in his 2007 essay, *Theses on Liveness*. Here, Croft breaks down the concept of liveness into two different kinds, *procedural* liveness and *aesthetic* liveness. The former is applicable wherever live sound is produced and manipulated in real time, it essentially refers to the most basic criterion for a live performance. Aesthetic liveness, in contrast, is a level above procedural liveness, in which some kind of input makes 'aesthetically meaningful' differences to the audio output in a real-time context, by which a significant, perceptible change or development is meant (Croft 2007, 61). This aspect of an identifiable relationship between input and output from the spectator's perspective is central to Croft's overall concept of what can legitimately be considered a live performance:

> Thus the onus of justification of liveness is shifted to the causal link between the performer's action and the computer's response. It is a question of the specificity of the relation: if many perceptibly different inputs generate outputs with no pertinent differences [...], then the liveness is merely procedural and not aesthetic – pre-recorded sounds would do the job as well or better.
>
> (ibid.)

On the basis of this, Croft offers eight conditions for the relationship between input and output in DMIs and similar technologies that, if fulfilled, would allow for greater aesthetic liveness in performances and also grant such devices a stronger claim to instrumentality (ibid., 64-5). These can be summarised into two underlying principles: that the performer's gestural input should always be met with an fitting response from the instrument (Conditions 1-3, 6 and 8) and that the instrument should itself have some degree of internal consistency, which in turn makes it learnable by a performer (Conditions 4, 5 and 7). By drawing on the spectator's point of view in order to better inform artistic practice, Croft's essay represents a very important step in the theoretical work on DMI reception and allows for the development of more concrete audience-centred approaches to evaluation (see also Schloss 2003).

A small collection of studies have taken up the issues of liveness and causality in empirical investigations. Bown et al. (2014) conducted an online survey which asked respondents to listen to eight audio recordings of live performances by laptop musicians and then rate various aspects, including the degree of perceived liveness (how live they felt the performance sounded), the perceived level of performer activity and overall enjoyment. Perceived liveness correlated as expected with perceived performer activity (ibid., 16), with judgements of performer activity being based on how accurate the timing of changes in the music were (e.g. rhythms that sounded quantised were judged as 'non-live') and the respondents' technical knowledge of electronic music production, which was generally very high (ibid., 15). Tracks that were rated as non-live received lower enjoyment ratings, which goes some way to affirming Auslander and Croft's ideas about the persistance of the value of live performance. It seems though that the connection between liveness and performer activity could be more convincingly confirmed through the presentation of audiovisual stimuli; in the audio-only paradigm, the performer's activity becomes highly abstracted and it is really the amount of musical change that is being rated. Furthermore, the respondents were told that they were listening to live performances, which might have biased them away from giving lower ratings of perceived liveness (the authors acknowledge this, ibid. 16).

Another attempt to investigate liveness empirically can be found in Berthaut et al. (2015), who explicitly take up Croft's link between liveness and causality. They showed participants video recordings of performances with three DMIs that had been created for the experiment and designed specifically to make gesture-sound causalities difficult to perceive: one featured a temporal delay between input gesture and sounded result, one was mapped so that discrete or short gestures resulted in sustained sounds (and vice versa) and the final instrument was only partially under the performer's control, with some audio features at times produced automatically by the computer (ibid., 385). The videos were presented both with and without animated pointers that were colour-coded to represent the musical parameter being controlled at any one time (pitch, timbre, loudness and musical pattern). These 'visual augmentations' were designed to clarify the mapping for the participants and were thus tested in experiment for their ability to achieve this. Participants were simply asked to rate the extent to which they thought the performer's gestures influenced the music (described here as a rating of performer *agency*, i.e. their level of activity and the extent to which this is effective or has an impact) and how confident they were about this. As predicted, the visual augmentations by and large increased ratings of agency and confidence ratings, which indeed suggests that the sense of liveness, or the perceived agency of the performer as it was measured as here, does improve with a clearer understanding of gesture-sound causality (ibid., 385-6).

3.2.2 Audience Understanding of DMIs

The most extensive empirical work on audience understanding of DMIs comes from A.C. Fyans, Michael Gurevich and Paul Stapleton, who through a series of studies have explored the extent to which the skill displayed by the performer is a factor in audience evaluations (Gurevich and Fyans 2011a), as well as in how far audience members are able to notice errors made in DMI performances (Fyans, Gurevich and Stapleton 2010). In Gurevich and Fyans' later article (2011b), these investigations are combined into a single paradigm. Participants were asked to watch recordings of performances with a theremin and a *Tilt Synth*, a DMI developed specifically for the experiment, and were then interviewed on their perceptions of instrumentality, skill and error. The authors shaped the interview questionnaire around five questions, adapted from Bellotti et al. (2002):

1. Address: How does the spectator know that the performer is directing communication to the system?
2. Attention: How does the spectator know that the system is responding to the performer?
3. Action: How does the spectator think the user controls the system?
4. Alignment: How does the spectator know that the system is doing the right thing?

5. Accident: How does the spectator know when the performer or the system has made a mistake?

(Gurevich and Fyans 2011b, 170)

The performers were not experts in their instruments, which thus ensured that several errors occurred (eight for the theremin player, five for the *Tilt Synth*). Overall, only two out of twenty-seven participants correctly identified errors in the *Tilt Synth* performance; it was generally held that such an instrument was 'error-free', as it did not appear to allow for any fine-grained control of pitch, in contrast to the theremin (ibid., 172). The participants' descriptions of how they thought the instruments functioned, especially for the *Tilt Synth*, were largely vague and inaccurate. The language used by participants furthermore suggested that the two instruments were quite differently perceived: when talking about the theremin, they tended to focus more on the instrumentalist's gestural control and its relationship to the sonic output and made positive comparisons to acoustic instruments, whereas for the *Tilt Synth*, their general lack of understanding of its functioning prevented this (ibid., 173). This also implies that perceived skill, which arises from greater understanding, is still an important category in the evaluation of a performance, as has been suggested elsewhere (e.g. Auslander 2009, 603-4).

Whereas Gurevich et al.'s work involved the evaluation of a DMI created for an investigation into the reception of DMIs, there have been other studies that have integrated audience evaluations into the design process of new DMIs. Barbosa et al. (2012) drew on evaluation models from O'Modhrain (2011) and Gurevich and Fyans (2011b) to produce a spectator questionnaire for the purpose of evaluating *Illusio,* a touch-screen live looping device that enables users to group together numbers of recorded loops and display them hierarchically, thereby allowing for greater variation than with conventional looping machines (Barbosa et al. 2013). They adopted the five aspects of Address, Attention, Action, Alignment and Accident from Gurevich and Fyans as above, translating these into a questionnaire with Likert-scale ratings relating to how much the spectator felt he or she understood certain aspects of the system, namely how comprehensible the performer's actions were (cause comprehension), if the spectator had enough information to understand the output (effect comprehension), comprehension of the gesture-sound relationship (mapping comprehension), the understanding of the performer's intention (intention comprehension) and finally, how perceptible errors were (error comprehension) (ibid., 405). The results suggest that *Illusio* was in general not very well understood by the survey respondents, who in some cases did give high understanding ratings on the scales, especially for the effect and mapping comprehension items, but who then contradicted this by giving confused or incorrect answers to the open-ended questions (ibid., 407). The designers intend to implement this feedback by enhancing *Illusio'*s visualisations, thereby offering more information for the spectators (see also Lai and Bovermann 2013 for another audience survey model).

3.3 Summary: Causality and Reception

This existing literature on DMI reception provides useful insights into the various concepts that could be involved in an audience's reception of a DMI performance, namely the perception of skill, overall understanding, perceived agency and liveness. It is necessary though to consider how these constructs might organise themselves in the process of perception and evaluation and how gesture-sound causality fits into this chain. As illustrated in Fig. 4 below, I propose that a clear gesture-sound causality is the foundation for understanding the basic functioning of the instrument, which then underlies such higher-level evaluative concepts as perceived skill and liveness. It still remains to be shown, however, to what extent causality plays a role in aesthetic and emotional responses to DMI performances (dashed arrow in Fig. 4),[21] an aspect that has been neglected in previous research and that the study presented in the following section seeks to address.

[21] 'Aesthetic response' is used here to refer to general liking and enjoyment, rather than judgements of 'beauty' or value.

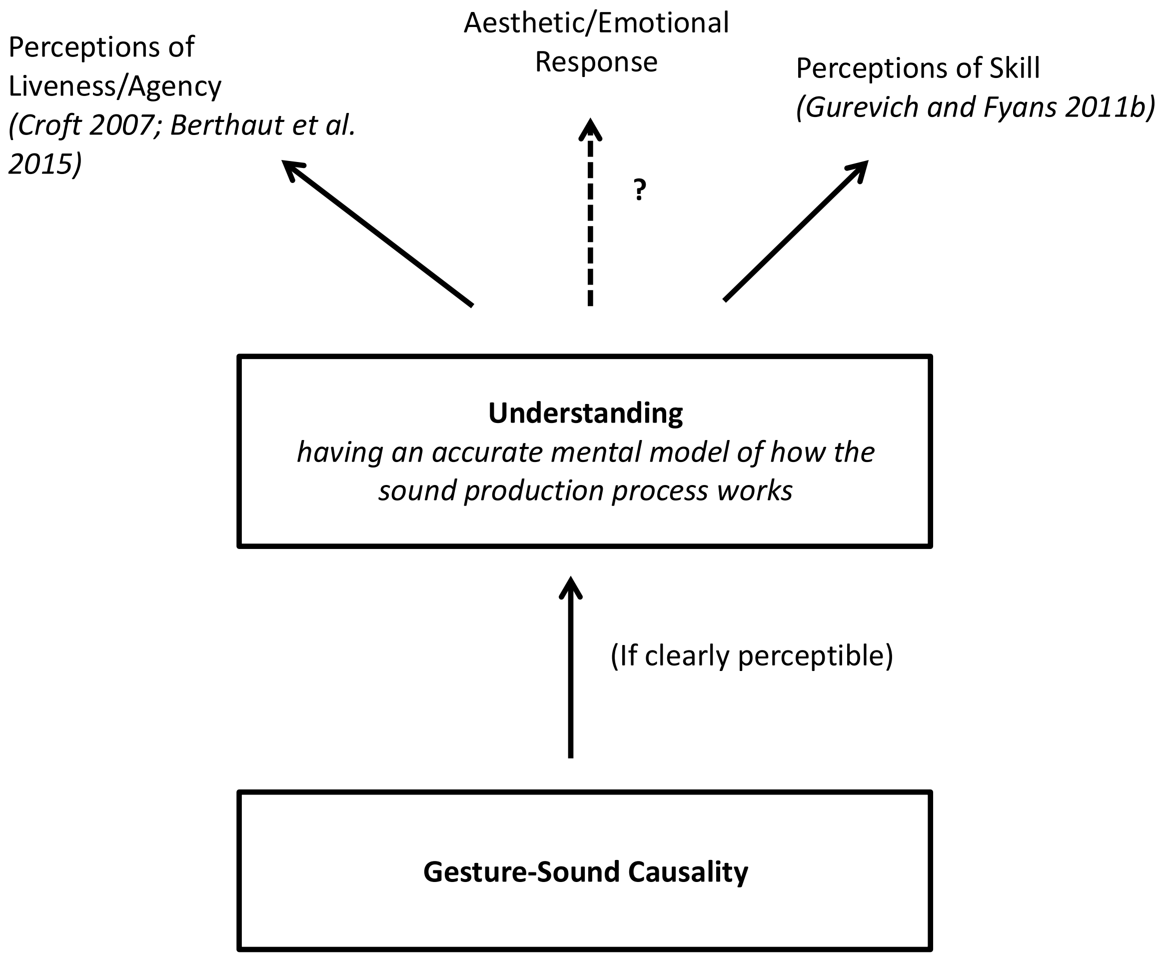

Fig. 4: A model of the DMI reception process.

4. Experimental Study: Do spectators rate DMI performances with mappings that have a higher degree of perceptible gesture-sound causality more positively?

4.1 Aims and Hypotheses

In order to investigate the effect of different levels of gesture-sound causality on the reception of performances with new DMIs, a two-part study was conducted, which comprised of a preliminary survey (see Section 4.2) and an experiment. The preliminary study was conducted at a live event to assess audience reactions to performances with new DMIs in an explorative manner and to collect common words and expressions that spectators use to describe the experience of perceiving such a performance, with the aim of using these terms in the creation of the experimental questionnaire. The experimental task itself then involved the rating of video clips of performances for two groups of DMIs: one with mapping designs that were more likely to be perceived as *causal* and one with mapping designs that were more likely to be perceived as *acausal*. These clips were presented in their original version and in a manipulated version, a mismatch of the audio and video tracks that created a unified acausal comparison condition (see Section 4.3.2). In light of the existing research presented in the preceding chapter, which posits connections between causality, understanding and liveness as well as between congruent stimuli, attention and effectiveness, the following hypotheses were made:

1. a) The DMIs in the causal group will be perceived as more causal than the acausal group.
   b) The original recordings will be rated as more causal than the manipulations.
   c) There will be a significant interaction between the factors Instrument Category and Manipulation, such that the effect of the manipulation on perceived causality ratings will be stronger for causal DMIs than for acausal DMIs.

2. a) The DMIs in the causal group will be rated more positively (i.e. as more interesting/ enjoyable, see Section 4.3.3) than the acausal group.
   b) The original recordings will be rated more positively than the manipulations.
   c) There will be a significant interaction between the factors Instrument Category and Manipulation, such that the effect of the manipulation on aesthetic and emotional ratings will be stronger for causal DMIs than for acausal DMIs.

4.2 Preliminary Study

A preliminary questionnaire study was carried out at a 3DMIN concert,[22] featuring four performances with instruments developed in the context of a 3DMIN-run seminar at the University of the Arts, Berlin. Alongside *S/A/S/A,* an instrument that was also used in the experiment (see Table 1 for details, pg. 33), these included a modular analogue synthesiser partly controlled via the movement of conductive concrete cubes on a plexiglass surface, an electronically-controlled percussion ensemble and a DMI with an organic controller made from a sheet of bacterial cellulose, a thermally conductive material stimulated here through contact with human touch and breath.[23] The questionnaire featured two questions per concert item, the first of which asked the respondents to describe in note form what they had paid most attention to during the performance, while the latter requested a more general description of the respondents' overall impression of the performance, encouraging them to touch upon any emotional responses and offer their thoughts on the instrument design and the music. Responses could be given in either English or German.[24]

4.2.1 Respondents

49 respondents completed the questionnaire (20 females, average age: 29.8 years). 30 respondents (61%) described themselves as being musically trained and 29 (59%) described themselves as being familiar with electroacoustic music. However, there appeared to be no notable difference in the reactions to the performances by respondents according to their musical expertise and prior exposure to such music.

4.2.2 Question 1: Attention

The majority of answers relating to attention made some mention of having focused on the gestures or movements of the performers and their relationship with the resultant sounds, which justifies the undertaking of further research into spectators' perceptions of this aspect of DMI performance. Answers frequently included such terms as 'gesture' (n=7), 'movement' (n=21), 'relation' (n=8) and 'interaction' (n=16), with some reporting having focused on interpreting the movements of specific parts of the body (e.g. the hands, n=25). Most

[22] The concert took place on 12.02.2015 at the Lab for Emerging Arts and Performance, Berlin.
[23] Video clips of the instruments can be found here: http://www.3dmin.org/activity/3dmin-concert-series/ws1415-student-concert/, last accessed 27.10.2015
[24] I will use English equivalents for German words when discussing the results. Equivalent German terms are included in the scores given for word usage frequency.

notably, fourteen respondents specifically mentioned trying to understand the gesture-sound causality:

> I was trying to figure out how the moving of the cubes correspond to the changes in sound.
>
> (Respondent 16)
>
> [I was] trying to understand the logic of relation between [the] players, relating the blocks as causes to effects.
>
> (Respondent 9)
>
> [I focused on the] hands/hand movements and [their] connection with the sounds.[25]
>
> (Respondent 21)
>
> How are the sounds produced?[26]
>
> (Respondent 20)
>
> I wanted to understand how the sounds are produced, so I paid attention to the movements.[27]
>
> (Respondent 11)
>
> It was a shame that there was a large gap between movement and interaction – no direct feedback [was] detectable.[28]
>
> (Respondent 49)
>
> [It was] unclear for me, what was made by [by the] laptop and what not.
>
> (Respondent 30)

Such comments suggest that mapping configuration can be the source of some confusion for audience members. Indeed, two respondents mentioned that a brief explanation of how the presented instrument works would have been helpful to them (Nos. 4 and 11). However, in

[25] Original German: ‚[...] Hände/Handbewegung und Zusammenhang mit Sound.'
[26] Original German: ‚Wie entstehen die Klänge?'
[27] Original German: ‚Ich wollte verstehen, wie die Klänge entstehen, also habe ich auf die Bewegungen geachtet.'
[28] Original German: ‚Ich fand es schade, dass zwischen Bewegung und Interaktion eine starke Lücke klaffte - keine direkte Rückkoppelung spürbar.'

contrast to this, one respondent (No. 10) did state that having to try to understand how the sound production worked made the performance appear as 'magic'.

4.2.3 Question 2: Overall Impressions

A summative content analysis was carried out on the answers to Question 2 (Hsieh and Shannon 2005; Mayring 2010). The words used to describe the performances were organised into three dimensions of spectator response that emerged from the data: terms pertaining to attention, emotional responses and words indicating preference or liking (Fig. 5). The most frequently used word overall was 'interesting' (n=27), followed by a number of other general, positive adjectives ('good', 'great' etc.) but answers also included more specific, emotionally-loaded terms such as 'tense'/'exciting' (n=13), 'threatening' (n=3) and 'theatrical'/'dramatic' (n=4). Given that much of the music produced (perhaps with the exception of the percussion setup) was freely structured, atonal and often involved only abstract, isolated sounds, it is notable that such a wide range of adjectives was employed and that emotional responses were reported; prior research demonstrates that atonal music provokes weaker emotional responses in listeners than tonal music (Daynes 2011). This may well be due to the respondents' high overall level of familiarity with electroacoustic/experimental music.

Further to this, there were eight responses that questioned whether 'instrument' was the correct word to use for the devices displayed, which shows how audiences can also become engaged with issues of instrumentality when faced with the peculiarities of DMIs:

> Is it even an instrument if you can't control it and the sounds are not predictable?[29]
>
> (Respondent 6)
>
> The instrument and the computer aren't separated enough to understand properly the way the instrument works.
>
> (Respondent 31)
>
> [It is] more of a sound-instrument than a musical instrument.[30]
>
> (Respondent 40)

[29] Original German: ‚Ist es überhaupt ein Instrument, wenn man es nicht kontrollieren kann und die Sounds sind nicht vorsehbar?'
[30] Original German: ‚Eher ein Klang-Instrument als ein Musikinstrument.'

Overall, the preliminary study confirmed that the perception of gesture-sound causality makes up an important part of the audience experience of a concert with new DMIs. This insight and the collected terms from Question 2 served as a useful basis for the design of the experiment.

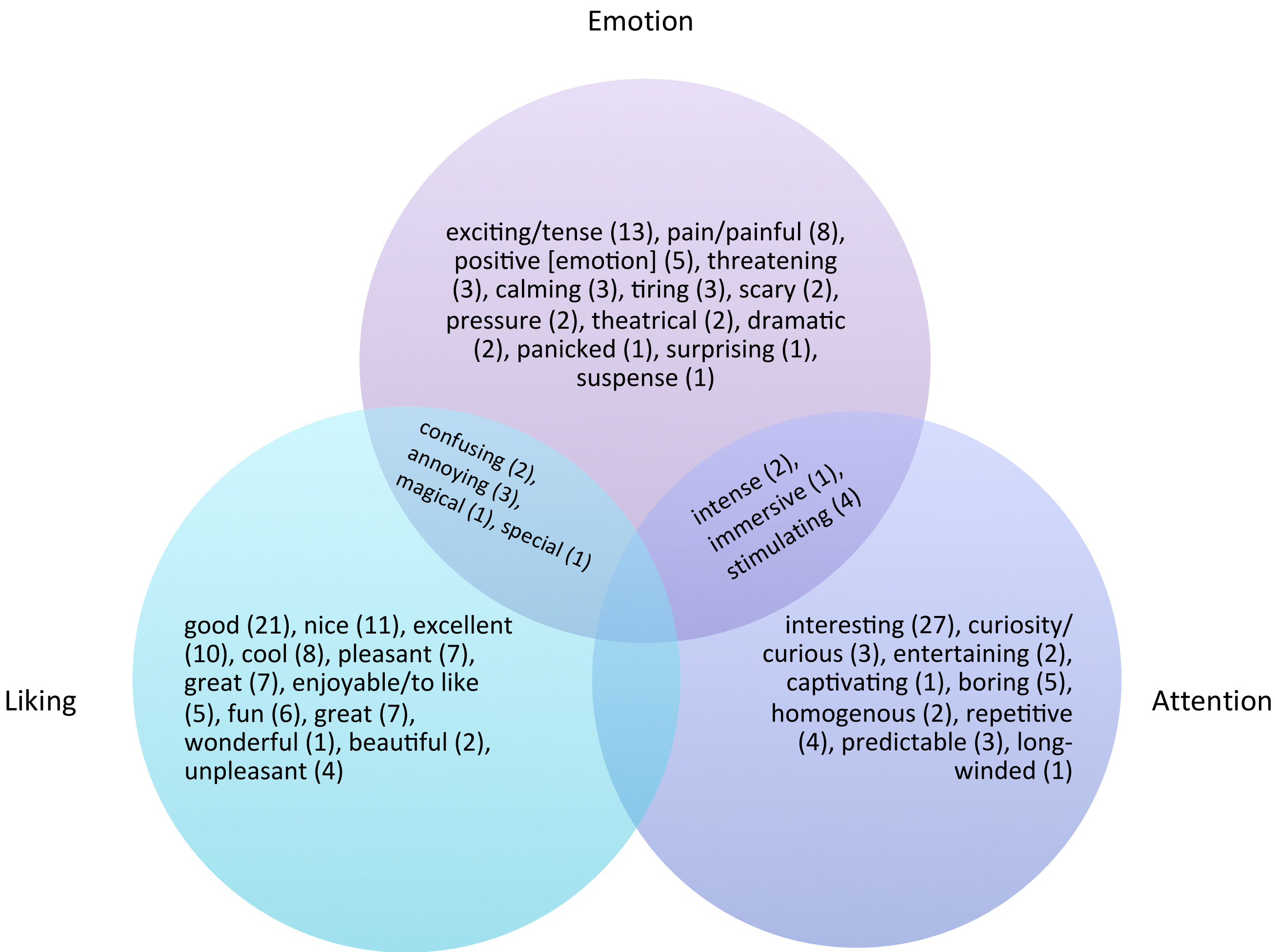

Fig. 5: Venn diagram showing the distribution of words used in answers to Question 2, organised according to dimension of reception (Attention, Emotion, Liking).

4.3 Main Study: Methods

4.3.1 Participants

Thirty-one participants (17 females, mean age: 26.8 years, range: 20-51, SD: 5.53) completed the experiment, after having been recruited via university mailing lists. Four were graduate students at the Audio Communication Group of the Technical University, Berlin participating for credit points. The remainder were other students from a variety of programmes, who received 10 Euros for participating. None of the participants had attended the concert at which preliminary study was conducted. The participants had received an average of 4.9 years of musical training (range: 0-27 years, SD: 5.66), with eighteen self-reporting as amateur musicians, twelve as non-musicians and one as a professional. Participants were asked to rate on scales of 1 to 7 how often they listen to experimental music and their liking for it. It appears that the sample generally had quite a low exposure to experimental music but that preference for it was quite high.[31]

4.3.2 Stimuli and Apparatus

Five DMIs, which had all except one been developed in connection with the 3DMIN project, were recorded to make the video stimuli for the experiment, using a JVC GY-DV301 Camcorder with direct audio input from the instruments. When selecting the instruments for the recording, attention was paid to gathering a wide range of controllers with differing degrees of perceptible causality. Three of the five controllers were then categorised as being more causal than the other three, resulting in two groups of instruments. This classification centred around the perceptibility of the mapping design and was based on a total of four factors that draw on John Croft's conditions for liveness and instrumentality in DMIs (Croft 2007, 64-5) and on the five spectator questions used by Gurevich and Fyans (2011b; see pg. 22 above):

1. **Type of controller:** DMIs that make use of interfaces resembling existing acoustic instruments are more likely to facilitate the understanding of gesture-sound relationships via the familiarity they offer; interfaces that allow for greater gestural variation from the performer (e.g. a wearable controller vs. a more laptop-based setup) could have a higher degree of perceptible causality.

[31] Mean rating of 2.77 (SD: 1.765) in response to the question: 'How often do you listen to experimental music'; Mean rating of 4.19 (SD: 1.6) in response to the question: 'How much do you like experimental music?'

2. **Latency/Temporal Delay:** while this can be difficult to avoid, a significant amount of latency is likely to impact the perceptibility of the mapping design and overall believability of the instrument, this could also be the case for other forms of temporal delay (e.g. delay/reverb effects or looping in the sonic output).
3. **Consistency of mapping design:** the relationship between the performer's actions and the resulting sounds should be consistent to aid perceptibility (similar to Croft 2007, Condition 6).
4. **Resulting output is appropriate for the given input:** larger gestures should map onto increased volume; short or percussive sounds should result from short, discrete actions and the equivalent applies for sustained sounds; no movement should mean no sound, continuous sound should involve some kind of continuous effort (similar to ibid., Condition 1).

In the categorisation process, an emphasis was furthermore placed on how the instrument appeared in the recording (see Table 1 for the instrument details).; since DMIs are endlessly re-programmable, it is difficult to make a decisive classification. *PushPull*, for instance, uses the *Influx Patch* (see pg. 15) to randomise the mapping strategy at will but this mode was not employed during the recording.

The performers each played a short performance of approximately 5-10 minutes in length and were asked to make this as varied as possible by including contrasting sections; the outline of a section of abstract sounds followed by a more tonal, melodic sounds was suggested. Ideally, the music played by each instrument would have been kept the same but given the vast differences in timbre and playing style between the instruments, it was decided that this would not be possible. The request to play in different styles resulted in one to three different performance styles or 'modes' per recording. These changes in mode also in some instances corresponded to a change in performance patch, hence the use of the word 'mode' rather than 'style'. In order to produce a controlled acausal condition for all instruments, the videos were manipulated in *VideoPad* Version 4.07 as follows. A 1 minute clip was selected from each of the 10 modes available. In each instance, the first 30 seconds were not selected so that the recording showed the performance at a more developed stage. For all of the instruments with two available modes, the audio from the second 1 minute clip was placed over placed over the video of the first and vice versa, resulting in two original and two manipulated versions.

Name	Creator(s)	Description	Causal or Acausal?	No. of Perfor-mance Modes	Manipulation Method
The Bass *(Fig. 7)*	Pierre-Alexandre Tremblay	**Augmented instrument** **Hardware:** 6-string fretless bass guitar played with distortion and volume pedals **Software:** Feeds via RME 400 sound card to laptop running Max/MSP performance patches, further control via iPad running Lemur	**Acausal** *Despite familiar interface, effects and looping create unclear gesture-sound relationship via temporal delay*	2	Mode 2 Audio --> Mode 1 Video, Mode 1 Audio --> Mode 2 Video
The Finger *(Fig.6)*	Dominik Hildebrand Marques-Lopez	**Wearable controller** **Hardware:** x-OSC I/O board combined with JeeNode wireless board and GravityPlug accelerometer, 2 potentiometers and 3 push buttons for control, plexiglass for outer shell of the glove **Software:** SuperCollider	**Causal** *Height and speed of gestures map logically onto changes in pitch/volume, no movement = no sound*	1	2 min clip selected, halved, 2nd Half Audio --> 1st Half Video, 1st Half Audio --> 2nd Half Video
Jerry *(Fig. 7)*	Christoph Schultz, Marten Seedorf	**Laptop instrument** **Hardware:** LogiLink USB Number Keypad and optical mouse **Software:** Pure Data	**Acausal** *Laptop-based setup provides little visual information, continuous sound produced without continuous visible effort*	2	Mode 2 Audio --> Mode 1 Video, Mode 1 Audio --> Mode 2 Video
PushPull *(Fig. 6)*	Till Bovermann, Dominik Hildebrand Marques-Lopez, Amelie Hinrichsen	**Instrument-inspired controller** **Hardware:** Complex combination of sensors within the latex bellow - inertial accelerometer measures acceleration of the hand, further light sensors measure the degree of contraction and two microphones pick up airflow in and out of the valves (see Hinrichsen et al. 2014), plus Arduino microcontroller **Software:** SuperCollider, uses the *Influx Patch* for variable mapping configuration (de Campo 2014, see pg. 15 above)	**Causal** *Design and playing action inspired by the accordian, maps as expected onto output parameters, no movement = no sound*	2	Mode 2 Audio --> Mode 1 Video, Mode 1 Audio --> Mode 2 Video
S/A/S/A *(Fig. 6)*	Julius Fischötter	**Alternative controller** **Hardware:** 10 triangular fields of different materials, made conductive either through copper under-plates or through being mixed with iron casting powder and connected to laptop via a Cypress CY8CKIT-049 USB device, fields are pressed, struck, scratched etc. **Software:** SuperCollider, with links to Ableton effect racks	**Causal** *Discrete/sustained actions mapped accordingly, no movement = no sound*	3	Mode 2 Audio --> Mode 1 Video, Mode 3 Audio --> Mode 2 Video, Mode 1 Audio --> Mode 3 Video

Table 1: Details of the five DMIs recorded for the experiment.

Fig. 6: The DMIs in the Causal Group.

Clockwise from top left: The Finger, PushPull and S/A/S/A

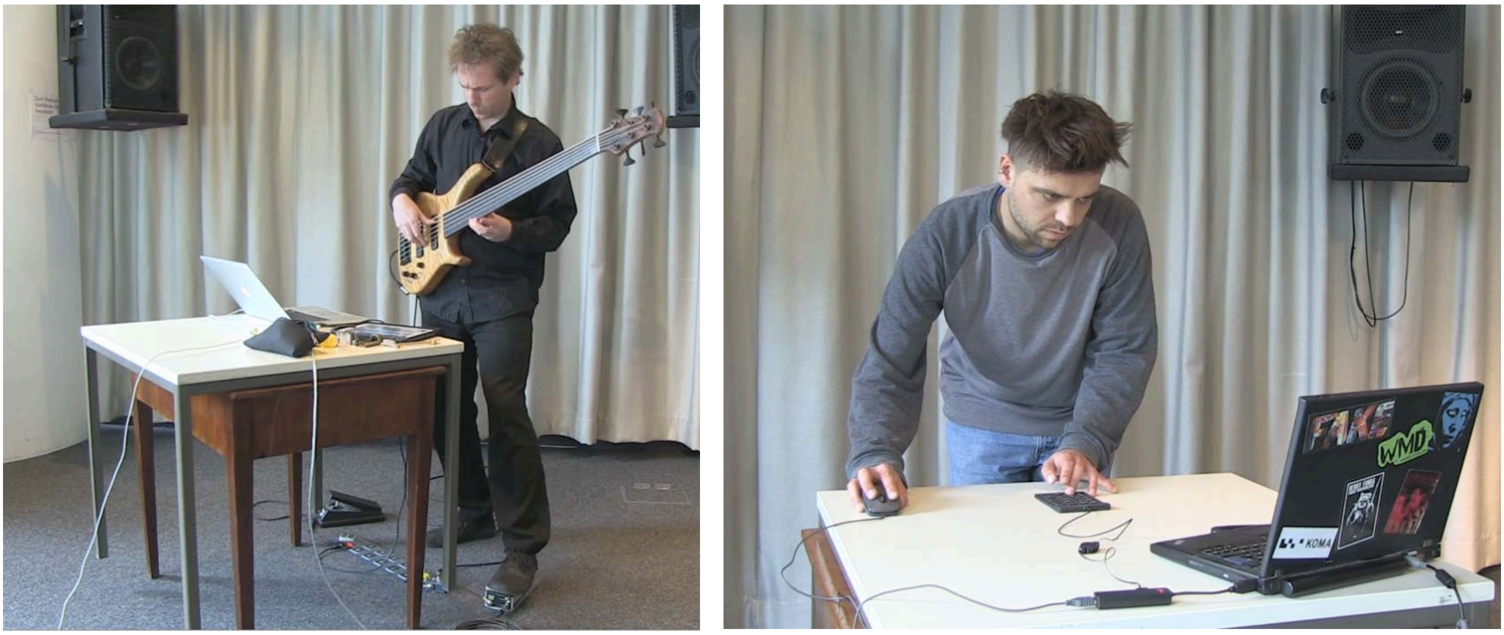

Fig. 7: The DMIs in the Acausal Group. *Left: The Bass, Right: Jerry*

For *S/A/S/A* (three modes), the audio from Mode 2 was placed over the video from Mode 1, the audio from Mode 3 over Mode 2's video and the audio from Mode 1 over Mode 3's video, creating a total of six clips. Finally, for *The Finger* (two modes), a 2 minute clip was selected, cut in half and the audio from the latter half placed over the first half and vice versa. This resulted in a total of 22 1 minute clips, eleven originals and eleven manipulations (summarised in Table 1 above). The clips (35 Mbits/s, 1920 x 1080 pixels, .avi format) were presented on and the data collected with *PsychoPy* Version 1.82.01,[32] running on a Lenovo Z500 series laptop, which was connected to an external 24" monitor and Philips SPA 5300 2.1 desktop loudspeakers. The participants could adjust the volume as desired.

4.3.3 Procedure

The participants each watched all 22 clips once in a randomised order. After each clip, they were asked to rate their level of agreement with the following adjectives on a scale from 1 to 7: *interesting, good, exciting, pleasant, boring, repetitive/predictable, stimulating, calming, virtuosic*.[33] These were the terms that occurred most frequently in the preliminary study data (see Fig. 5).[34] The rating of virtuosity was added separately and was included as it has been a common topic in DMI research and assessed as a possible factor in the evaluation of such

[32] http://www.psychopy.org/PsychoPyManual.pdf, last accessed 28.10.15

[33] Original German terms: *interessant, gut, spannend, angenehm, langweilig, repetitiv/vorhersagbar, anregend, beruhigend, virtuos*

[34] Terms with very similar meanings were excluded e.g. 'good' was used to cover 'nice', 'excellent' and 'cool', even though these were frequently used in their own right.

devices by spectators (Fyans et al. 2010, Gurevich and Fyans 2011b). Two further questions followed the adjective ratings, one asking how much attention the participant paid to the performance (i.e. how fixating was the performance)[35] and in how far the participant thought that the music was influenced by the gestures of the performer.[36] At the very end, a demographic questionnaire was filled out, which also asked participants whether more information about the instruments would have helped their evaluation. The participants were kept naive to the manipulation and were not given any information about any of the instruments or the 3DMIN project.

4.4 Results

A principal component analysis (PCA) was carried out on the response variables to simplify the data for further analysis and to see if the questionnaire items were received as expected by the participants, in terms of the emotional and attentional response dimensions used. The items 'Perceived Causality' and 'Virtuosity' were excluded from the PCA as they are distinct from the other ratings and do not fit theoretically into the emotion/attention response model. A correlation matrix for the response variables was produced and it was found that the item 'Repetitive/Predictable' correlated very poorly with almost all other response variables (only one coefficient > 0.3). It was thus excluded from the analysis.

The PCA was conducted on 8 items with orthogonal rotation (varimax). The Kaiser-Meyer-Olkin statistic confirmed the sampling adequacy for the analysis (KMO = 0.903; all KMO values for individual items were > 0.6) and Bartlett's test of sphericity identified correlations of sufficient size between items to justify the PCA (X^2 (28) = 4682.68, $p < 0.01$). The initial analysis revealed two components with eigenvalues over 1, which together explained 81.9% of the variance. These components were further confirmed by the scree plot and were therefore retained for the rotation. Table 2 shows the factor loadings after rotation. The loadings suggest that Component 1 can be interpreted in terms of the arousal-valence model as 'High Arousal, Positive Valence' and Component 2 as 'Low Arousal, Positive Valence' (see Russell 1980), which suggests that the questionnaire accurately reflected the underlying emotion model as intended. Factor scores were calculated using the Anderson-Rubin method and used for the calculation of further analyses. The data was reduced further by averaging

[35] Original German: ‚Wie fesselnd fanden Sie die Aufführung?'
[36] Original German: ‚Inwieweit wurde die Musik von der Gestik und den Bewegungen des/der Performer(s) Ihrer Meinung nach beeinflusst?'

the scores for the two manipulated and two original conditions for each instrument and then calculating mean scores for the causal and acausal groups and their manipulations.

	Rotated Factor Loadings	
Item	**High Arousal, +**	**Low Arousal, +**
Exciting	**.92**	.12
Interesting	**.90**	.10
Attention Paid	**.89**	.17
Boring	**-.82**	-.12
Stimulating	**.82**	.25
Good	**.81**	.41
Calming	.01	**.95**
Pleasant	.44	**.82**
Eigenvalues	5.23	1.32
% of variance	65.4	16.5

Table 2: Summary of PCA results (n=31), loadings > 0.4 are displayed in bold

4.4.1 Perceived Causality (Stimulus Treatment Check)

A two-way repeated measures ANOVA (Manipulation and Instrument Category) was carried out to test the effect of the manipulation and of instrument group on perceived causality (see Hypothesis 1, pg. 26; Fig. 8) and thereby confirm whether or not the stimuli were received as expected. The results show that the manipulation had a significant effect on perceived causality ratings, $F(1, 30) = 33.11$, $p < 0.01$ and that there was furthermore a significant difference between the instrument groups in the results for this rating, $F(1, 30) = 19.32$, $p < 0.01$. The interaction effect was significant, implying that the effect of manipulation on perceived causality varied significantly according to instrument category, $F(1, 30) = 4.51$, $p < 0.05$, $r = 0.36$.

A set of four post hoc paired sample t-tests were conducted to explore the interaction effect in more depth (Table 3). The participants' ratings of perceived causality confirmed the categorisation of the instruments (Hypothesis 1a) and the causality ratings were significantly lower in the manipulated condition, indicating that the experimental manipulation was accurately produced (Hypothesis 1b). The manipulation also had a stronger effect on ratings of perceived causality for the causal DMIs than for the acausal DMIs (Hypothesis 1c).

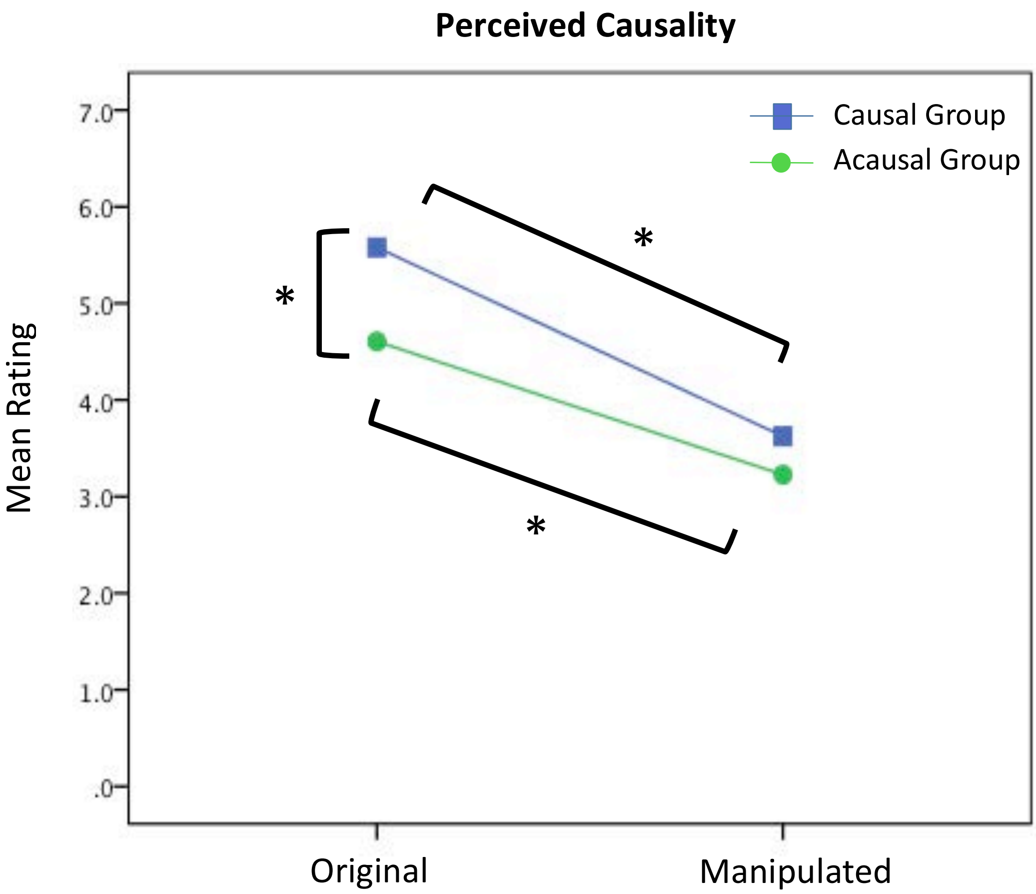

Fig. 8: Interaction between Manipulation and Instrument Category for Perceived Causality ratings. *Brackets mark significant differences from post hoc t-tests, $p < 0.0125$ (Bonferroni correction).

	Condition Means	*SDs*	*df*	*t*	*r*
Causal Original vs. Acausal Original	5.58, 4.60	0.79, 1.39	30	5.67*	.72
Causal Manip. vs. Acausal Manip.	3.62, 3.23	1.11, 1.28	30	1.68	.29
Causal Original vs. Causal Manip.	5.58, 3.62	1.24, 1.14	30	5.89*	.73
Acausal Original vs. Acausal Manip.	4.60, 3.23	1.35, 1.19	30	4.49*	.63

Table 3: Results of paired sample t-tests for Perceived Causality ratings, *$p < 0.0125$ (Bonferroni corrected).

4.4.2 Factor 1: High Arousal, Positive Valence

A two-way repeated measures ANOVA (Manipulation and Instrument Category) was conducted to test the effect of the manipulation and of instrument group on the positivity of ratings for Factor 1 (Hypothesis 2, pg. 26; Fig. 9). This was calculated using the factor scores from the PCA. The results show that the manipulation had a significant effect on the rating items that comprise Factor 1 (Exciting, Interesting, Attention Paid, Boring, Stimulating and Good), $F(1, 30) = 15.93$, $p < 0.01$ and that there was furthermore a significant difference between the instrument categories in the results for this factor, $F(1, 30) = 59.10$, $p < 0.01$. The interaction effect was also significant, suggesting that the effect of the manipulation on the ratings of the Factor 1 items varied significantly between the two groups, $F(1, 30) = 18.75$, $p < 0.01$, $r = 0.62$.

The results of the post hoc t-tests (Table 4) show that the causal group was indeed rated significantly more positively than the acausal instruments (Hypothesis 2a) and that the manipulations were in general rated less positively than the originals (Hypothesis 2b). The ratings for the causal group showed a significant difference between original and manipulated versions but this was not the case for the acausal group, indicating that they were far less strongly affected by the manipulation (Hypothesis 2c).

4.4.3 Factor 2: Low Arousal, Positive Valence

A further two-way repeated measures ANOVA (Manipulation and Instrument Category) was conducted to test the effect of the manipulation and instrument group on the positivity of ratings for Factor 2 (Hypothesis 2, pg. 26; Fig. 10). Once again, this was calculated using the factor scores from the PCA. The manipulation had a significant effect on the rating items that comprise Factor 2 (Calming, Pleasant), $F(1, 30) = 8.89$, $p < 0.01$ and there was furthermore a significant difference between the two groups in the results for this factor, $F(1, 30) = 45.94$, $p < 0.01$. The interaction effect was not significant, $F(1, 30) = 2.74$, $r = 0.28$, but post hoc t-tests (Table 5) still revealed a significant difference in Factor 2 ratings between the original and manipulated conditions for the causal group and not for the acausal group (Hypothesis 2c; the differences between the versions, independent of the instrument category factor, also support Hypothesis 2b). The acausal group, however, received significantly higher ratings than the causal group on Factor 2, in contrast to Hypothesis 2a; this result is interpreted in Section 4.5 below.

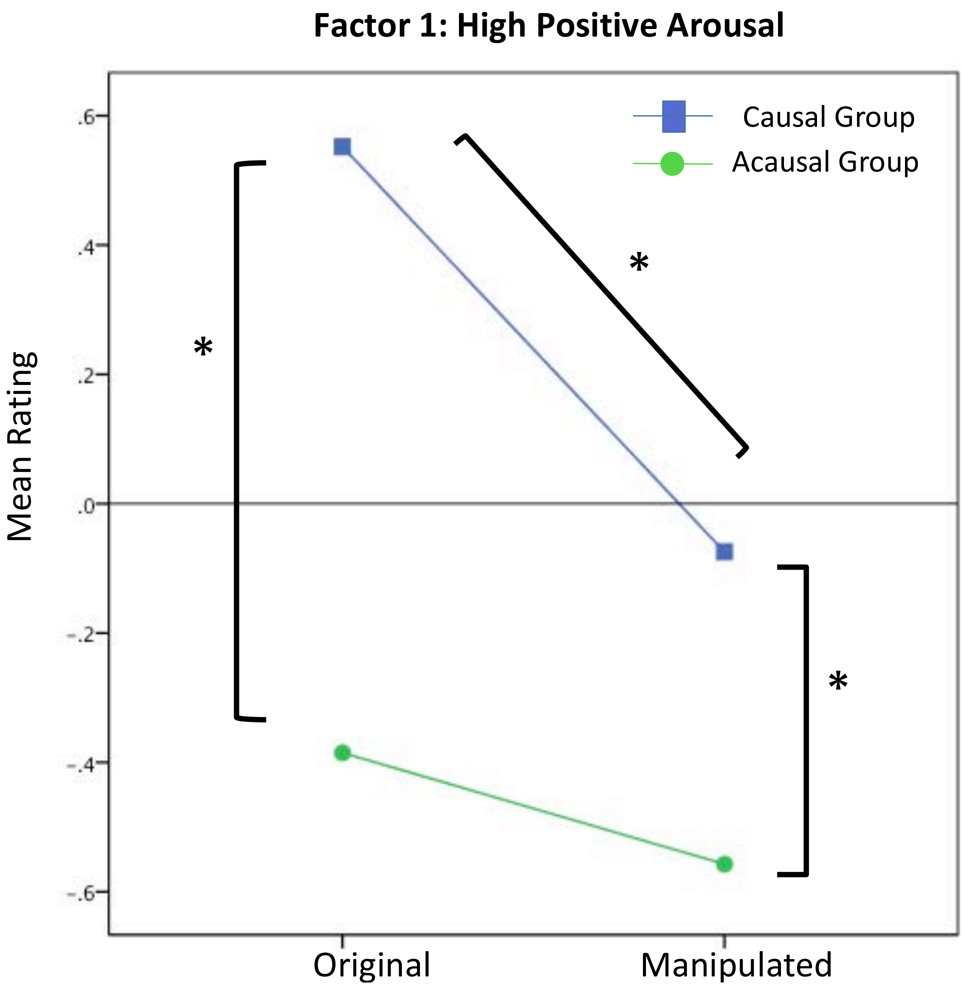

Fig. 9: Interaction between Manipulation and Instrument Category for Factor 1 ratings. *Brackets mark significant differences from post hoc t-tests, p < 0.0125 (Bonferroni correction).

	Condition Means	***SDs***	***df***	***t***	***r***
Causal Original vs. Acausal Original	0.55, -0.39	0.55, 0.68	30	8.22*	.83
Causal Manip. vs. Acausal Manip.	-0.07, -0.56	0.54, 0.65	30	4.93*	.67
Causal Original vs. Causal Manip.	0.55, -0.07	0.55, 0.54	30	4.99*	.67
Acausal Original vs. Acausal Manip.	-0.39, -0.56	0.68, 0.65	30	1.74	.30

Table 4: Results of paired sample t-tests for Factor 1 ratings, *p < 0.0125 (Bonferroni corrected).

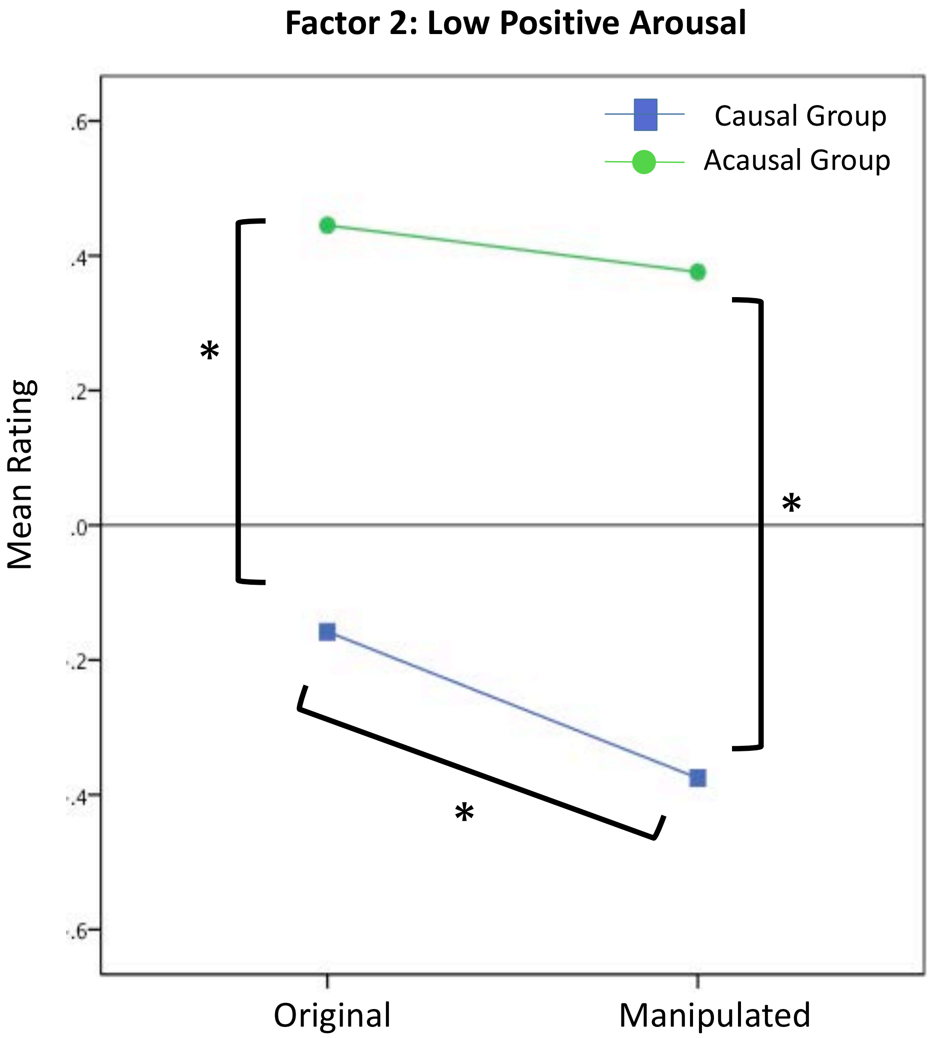

Fig. 10: Interaction between Manipulation and Instrument Category for Factor 2 ratings. *Brackets mark significant differences from post hoc t-tests, p < 0.0125 (Bonferroni correction).

	Condition Means	*SDs*	*df*	*t*	*r*
Causal Original vs. Acausal Original	-0.16, 0.45	0.49, 0.70	30	-5.33*	.69
Causal Manip. vs. Acausal Manip.	-0.38, 0.38	0.46, 0.73	30	-7.11*	.79
Causal Original vs. Causal Manip.	-0.16, -0.38	0.49, 0.46	30	4.67*	.64
Acausal Original vs. Acausal Manip.	0.45, 0.38	0.70, 0.73	30	0.86	.16

Table 5: Results of paired sample t-tests for Factor 2 ratings, *p < 0.0125 (Bonferroni corrected).

4.4.4 Virtuosity

A final two-way repeated measures ANOVA (Manipulation and Instrument Category) was conducted to test the effect of the manipulation on ratings of virtuosity (Hypothesis 2, pg. 26; Fig. 11). The results indicate that the manipulation had a significant effect on ratings of virtuosity, $F(1, 30) = 21.24$, $p < 0.01$ and that there was a significant difference between the instrument groups in the results for this response variable, $F(1, 30) = 13.39$, $p < 0.01$. The interaction effect was also significant, implying that the effect of manipulation on the virtuosity ratings varied significantly between the causal and acausal groups, $F(1, 30) = 15.26$, $p < 0.01$, $r = 0.58$.

The results of the post hoc t-tests (Table 6) revealed significantly lower ratings of virtuosity for the acausal group in comparison to the causal instruments (Hypothesis 2a) and lower ratings of virtuosity for the manipulations than for the originals (Hypothesis 2b). This difference was only just significant for the Acausal Original vs. Acausal Manipulated pair, $p = 0.012$, in contrast to the large significant difference for the causal group, which shows once more that the manipulation had a stronger effect on ratings for the causal DMIs (Hypothesis 2c).

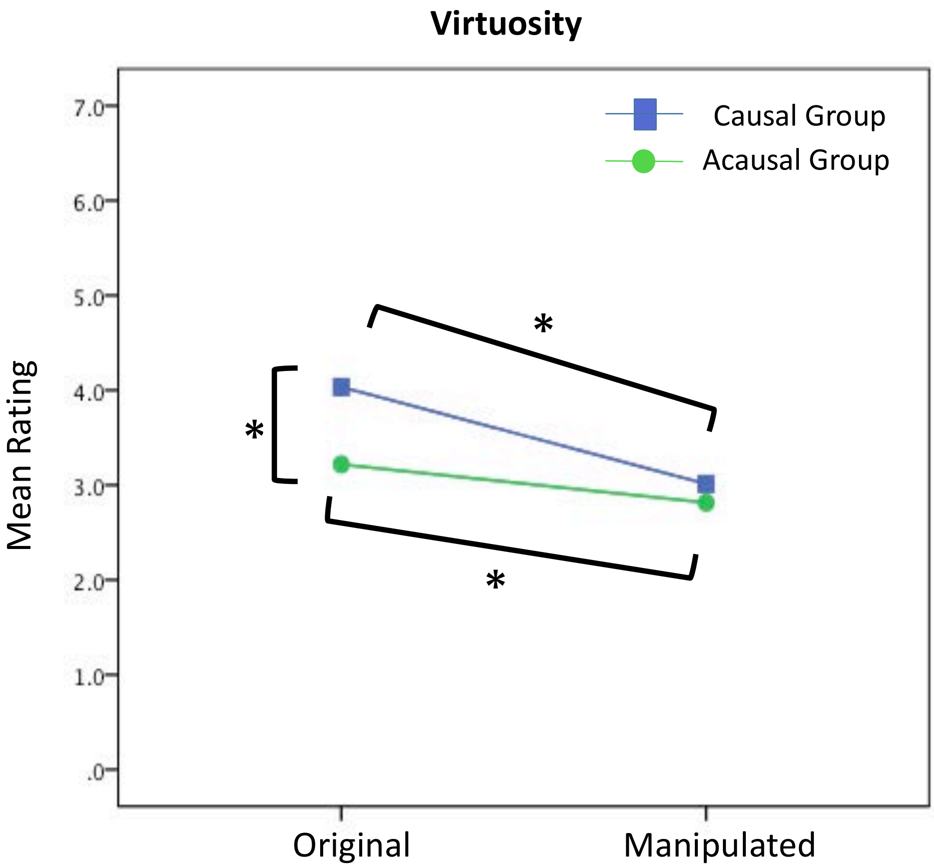

Fig. 11: Interaction between Manipulation and Instrument Category for Virtuosity ratings. *Brackets mark significant differences from post hoc t-tests, p < 0.0125 (Bonferroni correction).

	Condition Means	*SD*	*df*	*t*	*r*
Causal Original vs. Acausal Original	4.03, 3.22	1.24, 1.35	30	4.52*	.63
Causal Manip. vs. Acausal Manip.	3.01, 2.81	1.15, 1.19	30	1.46	.26
Causal Original vs. Causal Manip.	4.03, 3.01	1.24, 1.15	30	5.28*	.69
Acausal Original vs. Acausal Manip.	3.22, 2.81	1.35, 1.19	30	2.66*	.43

Table 6: Results of paired sample T tests for Virtuosity ratings, *p < 0.0125 (Bonferroni corrected).

4.5 Discussion

The reported results indicate that a lack of perceptible causality does overall have a negative impact on ratings of performances with DMIs. The instruments in the causal group, *PushPull*, *The Finger* and *S/A/S/A*, were viewed as considerably more interesting, more successful at holding the participants' attention and as demonstrating more skill than the acausal instruments and in comparison to their manipulated versions. Following on from the model of the DMI reception process presented above (see Section 3.2.3), I would posit that this result arises from the greater *understanding* that clearer gesture-sound causality offers spectators. The ability to perceive reliable relationships between gesture and sound establishes a necessary foundation for further judgements of a performance. This notion is supported further by the finding that the manipulations had, as predicted, a much stronger impact on the ratings for the causal DMIs than on those for the acausal DMIs. The acausal group received no significant difference in ratings between original and manipulated versions for Factors 1 and 2 (the Virtuosity ratings are also only just significantly different). This heavily suggests that participants struggled to perceive a difference between the original and manipulated versions of the acausal DMIs, which in turn implies that they were less able to figure out what the gesture-sound relationship was supposed to be and had thereby a poorer understanding. If a manipulation creates almost the same impression upon an audience as the original performance, this also shows how a lack of perceptible gesture-sound causality could thwart performers' attempts at communicating an artistic or expressive goal to spectators.

The results for the Low Arousal Factor are, at first glance, more difficult to interpret; given that the acausal instruments were expected to evoke a negative response, their high rating on the Low Arousal, Positive Valence factor (which includes the items 'Calming' and 'Pleasant') may seem to contradict Hypothesis 2a (see pg. 26). However, the results require interpretation in combination with those from Factor 1. The causal group were received as interesting, stimulating (high ratings on Factor 1) and, accordingly, *not calming* (low ratings on Factor 2) whereas the acausal group were received as boring and offering little stimulation (low ratings on Factor 1) but as calming and therefore, also pleasant (high ratings on Factor 2). It is certainly possible that participants either 1) perceived the lack of stimulation provided by the acausal instruments to be calming and a positive feature; or that 2) 'calming' was interpreted by the participants as equating 'boring'. The ratings for this factor may have also been confounded by the music played on the instruments in the acausal group, which could well have been perceived as calming and pleasant independent of the instrument design (this is discussed further below). In any case, high ratings on Factor 1 and Factor 2 for the causal group would have been a highly contradictory result; the participants were certainly consistent in rating the acausal group as less stimulating and the causal group as more so.

How could the experimental paradigm used here be improved upon? Firstly, some aspects of the experimental questionnaire could perhaps have benefited from further explanation. This applies in particular to the ratings for virtuosity, the averages for which lie around middle point of the scale, suggesting that participants were hesitant about giving more extreme ratings. Regarding the possible confounding factors in the study, an alternative design could have involved creating DMIs specifically for the experiment. This would have potentially allowed for further control of a number of factors, namely the music played on the instruments, the appearance of different performers in the video clips and greater control over the differing degrees of gesture-sound causality displayed. Gurevich and Fyans (2011b) and Berthaut et al. (2015), for instance, both employ instruments designed specifically for the experiment, with the former providing a discussion of the advantages of doing so in Marquez-Borbon et al. (2011). While it is certainly appealing to attempt to control these further aspects, such paradigms ultimately lose a lot in terms of ecological validity. By using recordings of a range of DMIs that were already in existence, the study presented here can be considered to well reflect the reality of DMI practice but this claim to ecological validity could have been strengthened through a more thorough simulation of live performance; the stimuli could have been presented on a larger screen and with a more immersive loudspeaker system. For future work on this topic, a paradigm involving ratings either during, with a continuous rating method, or immediately after live concert performances with DMIs could be considered for maximal validity of audience experience.

Conclusion: Designing the DMI Experience

The study presented here indicates that the relationship between gesture and sound is indeed an important factor for audiences, with higher causality, as created by the mapping and the type of controller, presumably providing them with more information, more reference points for evaluating the performance. Being able to perceive and then mentally model the sound generation process generates greater interest and also appears to provide a basis for the assessing the amount of skill involved in the performance. This is a result that carries potentially controversial implications for DMI practice. Since a lack of perceptible gesture-sound causality does lead to more negative ratings, how should this be countered? To what extent do DMI artists have a responsibility towards their audiences? To respond to these questions and to consider how they might form future directions of research, it is necessary to return to discussing the intention behind the production of DMIs and how this relates to the topics of musical communication, understanding and aesthetic appreciation.

Over the course of this thesis, I have proposed three possible intentions for the production of DMIs, as supported by the literature: 1) the intention of finding alternative ways of controlling digital sound synthesis, so as to make the composition and performance of digital music more reliant on the bodily effort of the performer and therefore more intuitive and satisfying; 2) the intention of creating new music (see Section 2.3); and 3) the intention of realising and communicating a particular artistic goal (see Section 2.3). For the latter two intentions, I would suggest that the audience's experience and level of comprehension is central to their successful realisation: if the audience cannot decode the performer's basic interaction with the instrument, the performer's intentions will remain opaque to them.

How, then, can DMI designers and performers go about increasing audience understanding and thereby improve the effectiveness of their musical expression? One approach would be to draw directly on the findings presented here and design DMIs that allow for a more clearly perceptible gesture-sound causality. For the instruments in the causal experimental group, this perceptibility manifested itself primarily in the amount of visual information the device allowed for and mappings that were logical. Indeed, the four factors impacting mapping perceptibility (type of controller, latency or other temporal delay, consistency of mapping design and having an appropriate input-output response, see Section 4.3.2.) as well as the principle of greater consideration of sound design, could be taken up as guidelines for future DMI designs (see similar propositions in Croft 2007, O'Modhrain 2011 and Barbosa et al. 2012). For many, however, delineating a framework for DMI design such as this could represent too much of an imposition on creative practice.

An alternative approach would be to consider how to make mappings more transparent to audiences in terms of performance presentation. Directly offering information to audiences on the mapping design and the general functioning of the instrument would be an easily implementable solution and one that audiences would seem to actively welcome; in response to the final question of whether more information about the instruments would have helped them evaluate the performances, 58% (n=18) of participants responded affirmatively. This information could take the form of programme notes or pre-performance demonstrations but there have been some suggestions of ways of providing information during performance. Berthaut et al. (2013) have, for example, designed a display system that can be placed under a tabletop DMI, which illustrates, via 3D visualisations, the ongoing sound generation process (a version of this was tested in Berthaut et al. 2015, see pgs. 21-22 above). There are also several live coding performers who include projections of their screen display in their setup, allowing spectators to see that changes in the code affect the sounds generated, from which a basic sense of the gesture-sound causality can be extracted (Brown and Sorensen 2007; 2009; McLean et al. 2010).

These various strategies could lead into directions for further research. One possible initial line of investigation could seek to establish the extent to which audiences understand or receive artistic intention accurately and could involve both the audience's and the designer/performer's perspective in this. Artists could be asked to note down their intentions and their means for realising them (e.g. the expression of particular emotion or the creation of an certain atmosphere) prior to performance and then the extent to which this matches the expression as perceived by the audience could be qualitatively evaluated for a number of different DMIs. A similar quantitative paradigm could investigate the communication of expressive intent via the comparison of audiovisual and video-only conditions. Through this, it could be established whether or not DMI performances provide enough visual information for the recognition of emotional expression without audio, which has shown to be possible for performances with acoustic instruments (Vines et al. 2006; 2011). The potential for supplementary information to modulate the accuracy of musical communication could be added to these paradigms and could also thereby compare the efficacy of different modes of providing information to audiences (e.g. programme notes vs. pre-performance demonstration vs. visual presentation accompanying performance).

However, the very act of providing information, through whichever means, is one that should be considered carefully. A balance needs to be struck between providing a basis for informed appreciation and altogether removing the unexpected and the provocative from musical performance. This is where research relating to DMIs has the opportunity feed into the wider discourse on the accessibility of new music. To draw on Fred Lerdahl's writings on composition and music cognition, contemporary art forms need to, within reason, challenge our perceptual and cognitive capacities and to push at natural constraints in order to remain interesting (Lerdahl 1992; 2001). Performances with DMIs *do* offer this challenge; they test audiences' perceptual limits both by presenting unclear gesture-sound causalities and by often working, in terms of sound design, at the border between noise and music. In some instances, this can add up to one challenge too many. By considering matters from the audience's perspective and keeping their potential constraints in mind, the project of creating new DMIs could result in richer musical and perceptual experiences that challenge and provoke audiences, but not to the point of alienating them. This is essential in order to move away from instrument design for its own sake and towards the production of new digital musical instruments that are not only interesting to play, but also to perceive.

Bibliography

Alperson, P. (2008). The Instrumentality of Music. *The Journal of Aesthetics and Art Criticism*, *66*(1), 37–51.

Arfib, D., Couturier, J. M., Kessous, L., & Verfaille, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, *7*(02), 127–144.

Arfib, D., Couturier, J.-M., & Kessous, L. (2005). Expressiveness and Digital Musical Instrument Design. *Journal of New Music Research*, *34*(1), 125–136.

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *Journal of Vision*, *6*(3), 260–268.

Auslander, P. (1999) *Liveness: Performance in a Mediatized Culture*. Routledge: London and New York.

– (2009) Lucille Meets GuitarBot: Instrumentality, Agency, and Technology in Musical Performance. *Theatre Journal*, *61*(4), 603–616.

Barbosa, J., Calegario, F., Teichrieb, V., Ramalho, G., & McGlynn, P. (2012). Considering Audiences' View Towards an Evaluation Methodology for Digital Musical Instruments. *NIME 2012: Proceedings of the International Conference on New Interfaces for Musical Expression*, 403–408.

Barbosa, J., Calegario, F., Teichrieb, V., Ramalho, G., & Cabral, G. (2013). Illusio : A Drawing-Based Digital Music Instrument. *NIME 2013: Proceedings of the International Conference on New Interfaces for Musical Expression*. Retrieved from: http://nime.org/proceedings/2013/nime2013_220.pdf, last accessed 01.11.2015

Behne, K.-E., & Wollner, C. (2011). Seeing or hearing the pianists? A synopsis of an early audiovisual perception experiment and a replication. *Musicae Scientiae*, *15*, 324–342.

Bellotti, V., Back, M., Edwards, W. K., Grinter, R. E., Henderson, A., & Lopes, C. (2002). Making sense of sensing systems: five questions for designers and researchers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *1*(1), 415–422.

Bense, A. (2013). *Musik und Virtualität: Digitale Virtualität im Kontext computerbasierter Musikproduktion*. epOs Music: Osnabrück.

Bergsland, A., & Åse, T. (2012). Using a seeing / blindfolded paradigm to study audience experiences of live-electronic performances with voice. *NIME 2012: Proceedings of the International Conference on New Interfaces for Musical Expression*, 433–436.

Berthaut, F., Marshall M., Subramanian S., Hachet M. (2013). *Rouages*: Revealing the Mechanisms of Digital Musical Instruments to the Audience. *NIME 2013: Proceedings of the International Conference on New Interfaces for Musical Expression*, 165–169.

Berthaut, F., Coyle D. Moore, J. W., Limerick H. (2015). Liveness Through the Lens of Agency and Causality Liveness Through the Lens of Agency and Causality. *NIME 2015: Proceedings of Conference on New Interfaces for Musical Expression.* 382-6.

Bongers, B. (2000). Physical Interfaces in the Electronic Arts Interaction Theory and Interfacing Techniques for Real-time Performance. In: *Trends in Gestural Control of Music*, Battier and Wanderley (eds.) 41–70. Paris: IRCAM.

Bovermann, T., Egermann, H., Foerstel, A., Hinrichsen, A., Hildebrand Marques Lopes, D., Pysiewicz, A., Hardjowirogo, S., Weinzierl, S., & de Campo, A. (2014). 3DMIN – Challenges and Interventions in Design, Development and Dissemination of New Musical Instruments. *Proceedings of ICMC/SMC*, Athens. Retrieved from: http://www.3dmin.org/wp-content/uploads/2014/10/2014-icmc_intro3dmin-final.pdf, last accessed 01.11.2015

Bown, O., Bell, R., & Parkinson, A. (2014). Examining the Perception of Liveness and Activity in Laptop Music: Listeners' Inference about what the Performer is Doing from the Audio Alone. *NIME 2014: Proceedings of the International Conference on New Interfaces for Musical Expression*, 13–18.

Brown, A. R. & Sorensen, A. C., (2007). aa-cell in Practice: An Approach to Musical Live Coding. *Proceedings of the International Computer Music Conference*, 292-299.

- Brown, A. R., & Sorensen, A. C. (2009). Interacting with Generative Music through Live Coding. *Contemporary Music Review*, *28*(1), 17–29.

Burrows, D. (1987). Instrumentalities. *Journal of Musicology*, *5*(1), 117–125.

Cadoz, C., & Wanderley, M. M. (2000). Gesture – Music. In: *Trends in Gestural Control of Music*, Battier and Wanderley (eds.) 71–94. Paris: IRCAM.

Camurri, A., Hashimoto, S., Ricchetti, M., Ricci, A., Suzuki, K., Trocca, R., & Volpe, G. (2000). Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal*, *24*(1), 57-69.

Camurri, A., Coletta, P., Varni, G., & Ghisio, S. (2007). Developing multimodal interactive systems with EyesWeb XMI. *NIME 2007: Proceedings of the International Conference on New Interfaces for Musical Expression,* 305-308.

Cance, C., Genevois, H., & Dubois, D. (2009). What is instrumentality in new digital msuical devices? A contribution from cognitive linguistics and psychology. *Proceedings of the 2009 JIM Conference*. Retrieved from: http://www.researchgate.net/publication/39064353_What_is_instrumentality_in_new_digital_musical_devices__A_contribution_from_cognitive_linguistics_and_psychology, last accessed 15.10.2015

Caramiaux, B., Bevilacqua, F., & Schnell, N. (2010). Towards a gesture-sound cross-modal analysis. In: Kopp and Wachsmuth (Eds.) *Gesture in Embodied Communication and Human-Computer Interaction*. Berlin, Heidelberg: Springer. 158-170.

Caramiaux, B. (2014). Mapping Through Listening. *Computer Music Journal*, *38*(3), 34–48.

Caramiaux, B., Bevilacqua, F., Bianco, T., Schnell, N., Houix, O., & Susini, P. (2014). The Role of Sound Source Perception in Gestural Sound Description. *ACM Transactions on Applied Perception*, *11*(1), 1–19.

Chadabe, J. (2002). The limitations of mapping as a structural descriptive in electronic instruments. *NIME 2002: Proceedings of the Conference on New interfaces in Musical Expression*, 1-5.

Cohen, A. (2013): Congruence-Association Model of Music and Multimedia: Origin and Evolution. In: Tan et al. (eds) *The Psychology of Music in Multimedia.* Oxford: Oxford University Press.

Collins, N. (2010). *Introduction to Computer Music*. Hoboken: Wiley.

Croft, J. (2007). Theses on Liveness. *Organised Sound*, *12*(01), 59.

Dahl, S. & Friberg, A. (2007): Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception: An Interdisciplinary Journal. 25*(5): 433–54.

Daynes, H. (2011). Listeners' perceptual and emotional responses to tonal and atonal music. *Psychology of Music*, *39*(4), 468–502.

De Campo, A. (2014). Lose control, Gain influence: Concepts for Metacontrol. *Proceedings of the International Computer Music Conference.* 217–222.

Delalande, F. (1988). La gestique de Gould: Elements pour une semiologie du geste musical. In G. Guertin (Ed.), *Glenn Gould Pluriel*. Quebec: Louise Courteau, 83-111.

Dobrian, C., Koppelman, D. (2006). The 'E' in NIME: Musical Expression with New Computer Interfaces. *NIME 2006: Proceedings of Conference on New Interfaces for Musical Expression*. 277–282.

Doornbusch, P. (2002). Composers' views on mapping in algorithmic composition. *Organised Sound*, *7*(02), 145–156.

Fasciani, S., & Wyse, L. (2012). A voice interface for sound generators: adaptive and automatic mapping of gestures to sound. *NIME 2012: Proceedings of the 12th International Conference on New Interfaces for Musical Expression.* Retrieved from: http://www.nime.org/proceedings/2012/nime2012_57.pdf, last accessed 01.11.2015

Fels, S., Gadd, A., & Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, *7*(02), 109–126.

Fyans, C., Gurevich, M., & Stapleton, P. (2010). Examining the Spectator Experience. *NIME 2010: Proceedings of the 2010 Conference on New Interfaces for Musical Expression.* 451–454.

Godøy, R. I., & Leman, M. (Eds.) (2010). *Musical Gestures: Sound, Movement, and Meaning*. New York and London: Routledge.

Green, O. (2014). NIME, Musicality and Practice-led Methods. *NIME 2014*: *Proceedings of the International Conference on New Interfaces for Musical Expression*. 1–6.

Gritten, A., & King, E. (Eds.). (2006). *Music and Gesture*. Ashgate.

Gurevich, M. & Fyans, A.C. (2011a). Perceptions of Skill in Performances with Acoustic and Electronic Instruments. *NIME 2011*: *Proceedings of the International Conference on New Interfaces for Musical Expression.* 495–498.

– Gurevich, M. & Fyans, A.C. (2011b). Digital Musical Interactions: Performer–system relationships and their perception by spectators. *Organised Sound*. 16(2): 166-75.

Gurevich, M., & von Muehlen, S. (2001). The Accordiatron: A MIDI controller for interactive music. *NIME 2001: Proceedings of the International Conference on New interfaces for Musical Expression*. 1-3.

Guski, R., & Troje, N. F. (2003). Audiovisual phenomenal causality. *Perception & Psychophysics*, *65*(5), 789–800.

Hatten, R. S. (2004). *Interpreting musical gestures, topics, and tropes: Mozart, Beethoven, Schubert*. Indiana University Press.

Hinrichsen, A., Hardjowirogo, S., Hildebrand Marques Lopes, D., Bovermann, T. (2014). *PushPull*. Reflections on Building a Musical Instrument Prototype. *Proceedings of the International Conference on Life Interfaces*. Retrieved from: http://www.3dmin.org/wp-content/uploads/2014/03/Hinrichsen_2014.pdf, last accessed 01.11.2015

von Hornbostel, E. M., & Sachs, C. (1914). Systematik der Musikinstrumente. Ein Versuch. *Zeitschrift für Ethnologie*, 553-590.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288.

Hsu, W., & Sosnick, M. (2009). Evaluating Interactive Music Systems : An HCI Approach. *NIME 2009: Proceedings of the International Conference on New Interfaces for Musical Expression*, 25–28.

Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance. *Trends in Gestural Control of Music*, In: *Trends in Gestural Control of Music*, Battier and Wanderley (eds.) 231-258. IRCAM: Paris.

Hunt, A., & Wanderley, M. (2002). Mapping performer parameters to synthesis engines, *Organised Sound. 7(2),* 97–108.

Hunt, A., Wanderley, M. & Paradis, M. (2003). The importance of parameter mapping in electronic instrument design. *Journal of New Music Research.* 32(4): 429-440.

Impett, J. (1994). A meta-trumpet(-er). *Proceedings of the International Computer Music Conference*.

Iwamiya, S., Sugano, Y., & Kouda, K. (2000): The Effects of Synchronization of Temporal Structure of Sound and Motion Picture on the Impression of Audiovisual Context. *IEEE SMC Conference Proceedings*. 1222-1225.

Iwamiya, S. (2013). Perceived congruence between auditory and visual elements in multimedia. In: Tan et al. (eds) *The Psychology of Music in Multimedia.* Oxford: Oxford University Press.

Jensenius, A. R., Wanderley, M.M., Godøy, R. I. & Leman, M. (2010). Musical Gestures: Concepts and Methods in Research. In: Godøy, R. I., & Leman, M. (Eds.) (2010). *Musical Gestures: Sound, Movement, and Meaning*. New York and London: Routledge, 12-36.

Jensensius, A.R. (2014). To Gesture or Not? An Analysis of Terminology in NIME Proceedings 2001-2013. *NIME 2014: Proceedings of the International Conference on New Interfaces for Musical Expression*. 217–220.

Jessop, E., Torpey, P. A., & Bloomberg, B. (2011). Music and Technology in *Death and the Powers*. *NIME 2011: Proceedings of the 2011 Conference on New Interfaces for Musical Expression*. 349–354.

Johnson, M., & Larson, S. (2003). Something in the Way She Moves: Metaphors of Musical Motion. *Metaphor and Symbol*, *18*(2), 63–84.

Jordà, S. (2004). Instruments and Players: Some Thoughts on Digital Lutherie. *Journal of New Music Research*, *33*(3), 321–341.

Jordà, S. and Maella, S. (2014). A Methodological Framework for Teaching, Evaluating and Informing NIME Design with a Focus on Expressiveness and Mapping. *NIME 2014: Proceedings of the International Conference on New Interfaces for Musical Expression*, 233–238.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kim, J. H. (2012). *Embodiment in interaktiven Musik- und Medienperformances – unter besonderer Berücksichtigung medientheoretischer und kognitionswissenschaftlicher Perspektiven*. Osnabrück: epOs.

Lai, C., & Bovermann, T. (2013). Audience Experience in Sound Performance. *Proceedings of the International Conference on New Interfaces for Musical Expression*, 170–173.

Lee, M., Freed, A., & Wessel, D. (1991). Real-time neural network processing of gestural and acoustic signals. *Proceedings of the International Computer Music Conference*. 277-280.

Lerdahl, F. (1992). Cognitive constraints on compositional systems. *Contemporary Music Review*, *6*(2), 97–121.

– (2001). *Tonal Pitch Space*. Oxford: Oxford University Press.

Leman, M. (2008) *Embodied Music Cognition and Mediation Technology*. Cambridge, London: MIT Press.

Levitin, D. J., McAdams, S., & Adams, R. L. (2002). Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, *7*(2), 171–189.

Lipscomb, S. D. (1998). Synchronization of musical sound and visual images: Issues of empirical and practical significance in multimedia development. *The Journal of the Acoustical Society of America*.

Luciani, A., Florens, J.-L., Couroussé, D., & Castet, J. (2009). Ergotic Sounds: A New Way to Improve Playability, Believability and Presence of Virtual Musical Instruments. *Journal of New Music Research*, *38*, 309–323.

Machover, T. (1992). *Hyperinstruments*: A Progress Report. MIT Media Lab.

Magnusson, T. (2009). Of Epistemic Tools: musical instruments as cognitive extensions. *Organised Sound*, *14*(02), 168.

Mayring, P. (2010). *Qualitative Inhaltsanalyse: Grundlagen und Techniken*. Weinheim and Basel: Beltz Verlag.

Malloch, J., Birnbaum, D., Sinyor, E., & Wanderley, M. M. (2006). Towards a new conceptual framework for digital musical instruments. *Proc. of the 9th International Conference on Digital Audio Effects*, 49–52.

Manning, P. (2013). *Electronic and Computer Music*. Oxford: Oxford University Press.

Marquez-Borbon, A., Gurevich, M., Fyans, A., & Stapleton, P. (2011). Designing digital musical interactions in experimental contexts. *Contexts*, 373–376.

Mathews, M. V., & Abbott, C. (1980). The Sequential Drum. *Computer Music Journal*, *4*(4), 45–59.

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*. 264: 746–748.

McLean, A., Griffiths, D., & Collins, N. (2010). Visualisation of live code. *Electronic Workshops in Computing: Electronic Visualisation and the Arts (EVA 2010)*, 26–30.

Merrill, D. J., & Paradiso, J. (2005). Personalization, expressivity, and learnability of an implicit mapping strategy for physical interfaces. *CHI 2005: Conference on Human Factors in Computing Systems*.

Miranda, E. R. & Wanderley M.M. (2006): *New digital musical instruments: Control and Interaction beyond the keyboard*. Computer Music and Digital Audio Series: Vol 21. Middleton: AR Editions.

Mulder, A. (1994). Virtual Musical Instruments : Accessing the Sound Synthesis Universe as a Performer. *Proceedings of the 1st Brazilian Symposium on Computer Music.* 243–250.

Nakra, T. M. (2000). Inside the Conductor's Jacket: Analysis, Interpretation and Musical Synthesis of Expressive Gesture. Unpublished PhD Thesis. MIT Media Lab.

Ng, K. (2002). Sensing and mapping for interactive performance. *Organised Sound*. *7(2),* 191–200.

O'Modhrain, S. (2011). A Framework for the Evaluation of Digital Musical Instruments. *Computer Music Journal*, *35*(1), 28–42.

Park, T. H., & Mathews, M. (2009). An Interview with Max Mathews. *Computer Music Journal*, *33*(3), 9–22.

Roads, C., & Mathews, M. (1980). An Interview with Max Mathews, *Computer Music Journal, 4(4),* 15-22.

Rovan, J. B., Wanderley, M. M., Dubnov, S., & Depalle, P. (1997). Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Kansei: The Technology of Emotion Workshop. Proceedings of the AIMI International Workshop*, 68–73.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-78.

Schiesser, S., & Schacher, J. C. (2012). SABRe : The Augmented Bass Clarinet. *NIME 2012: Proceedings of the International Conference on New Interfaces for Musical Expression*, 109–112.

Schloss, W. A. (2003). Using Contemporary Technology in Live Performance: The Dilemma of the Performer. *Journal of New Music Research.* 32(3): 239–242.

Schneider A. (2010). Music and gestures: a historical introduction and survey of research. In R.I. Godøy, M. Leman (eds.). *Musical Gestures. Sound, Movement, and Meaning.* New York and London: Routledge, 69-100.

Schutz, M. (2008). Seeing Music? What musicians need to know about vision. *Empirical Musicology Review*, *3*(3), 83–108.

Schutz, M., & Kubovy, M. (2009). Causality and cross-modal integration. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(6), 1791–1810.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*(2004), 888–897.

Schutz, M., & Manning, F. (2012). Looking beyond the score: The musical role of percussionists' ancillary gestures. *Music Theory Online*, *18*(1), 1–14.

Silveira, J. M. (2013). The effect of body movement on listeners' perceptions of musicality in trombone quartet performance. *International Journal of Music Education*.

Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science and Technology*, *28*(2), 61–70.

Van Nort, D., Braasch, J., & Oliveros, P. (2012). Mapping to musical actions in the FILTER system. *NIME 2012: Proceedings of the International Conference on New Interfaces for Musical Expression*. Retrieved from: http://www.nime.org/proceedings/2012/nime2012_235.pdf, last accessed 01.11.2015

Van Nort, D. (2014). Mapping Control Structures for Sound Synthesis: Functional and Topological Perspectives. *Computer Music Journal*, *38*(3), 6–22.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*(1), 80–113.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., Dalca, I. M., & Levitin, D. J. (2011). Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition*, *118*(2), 157–170

Vuoskoski, J. K., Thompson, M. R., Clarke, E. F., & Spence, C. (2013). Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception & Psychophysics*. 76: 591–604.

Waisvisz, M. (1985) *The Hands*: a set of remote MIDI controllers. *Proceedings of the International Computer Music Conference*. 86–89.

Wanderley, M. M.; Orio, N. (2002). Musical Expression: Borrowing Tools. *Computer Music Journal*, *26*(3), 62–76.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667.

Wessel, D., & Wright, M. (2002). Problems and Prospects for Intimate Musical Control of Computers. *Computer Music Journal*, *26*(3), 11–22.

Wessel, D., Wright, M., & Schott, J. (2002). Intimate Musical Control of Computers with a Variety of Controllers and Gesture Mapping Metaphors. *NIME 2002: Proceedings of the International Conference on New Interfaces for Musical Expression.* 1–3.

Humboldt-Universität zu Berlin
Kultur-, Sozial- und Bildungswissenschaftliche Fakultät

Name: ………………………………………………….. Vorname: …………………………………………….

Matrikelnummer:……………………………….

Eidesstattliche Erklärung zur

☐ **Hausarbeit** ☐ **Magisterarbeit**

☐ **Bachelorarbeit** ☐ **Masterarbeit**

Ich erkläre ausdrücklich, dass es sich bei der von mir eingereichten schriftlichen Arbeit mit dem Titel

…………………………………………………………………………………………………

…………………………………………………………………………………………………

um eine von mir erstmalig, selbstständig und ohne fremde Hilfe verfasste Arbeit handelt.

Ich erkläre ausdrücklich, dass ich *sämtliche* in der oben genannten Arbeit verwendeten fremden Quellen, auch aus dem Internet (einschließlich Tabellen, Grafiken u. Ä.) als solche kenntlich gemacht habe. Insbesondere bestätige ich, dass ich ausnahmslos sowohl bei wörtlich übernommenen Aussagen bzw. unverändert übernommenen Tabellen, Grafiken u. Ä. (Zitaten) als auch bei in eigenen Worten wiedergegebenen Aussagen bzw. von mir abgewandelten Tabellen, Grafiken u. Ä. anderer Autorinnen und Autoren (Paraphrasen) die Quelle angegeben habe.

Mir ist bewusst, dass Verstöße gegen die Grundsätze der Selbstständigkeit als Täuschung betrachtet und entsprechend der fachspezifischen Prüfungsordnung und/oder der Allgemeinen Satzung für Studien- und Prüfungsangelegenheiten der HU (ASSP) bzw. der Fächerübergreifenden Satzung zur Regelung von Zulassung, Studium und Prüfung der Humboldt-Universität (ZSP-HU) geahndet werden.

Datum ………………………………..   Unterschrift

Einverständniserklärung (gilt nur für Abschlussarbeiten)

Ich bin damit einverstanden, dass bei Empfehlung durch den Betreuer/die Betreuerin meiner Abschlussarbeit diese in die entsprechende Zweigbibliothek der Kultur-, Sozial- und Bildungswissenschaftlichen Fakultät der HU aufgenommen wird und als Publikation der Bibliothek zur Verfügung gestellt wird.

☐ **ich bin einverstanden** ☐ **ich bin <u>nicht</u> einverstanden**

Berlin, ...   (Unterschrift)