MUSIC COMPUTING GROUP

DEPARTMENT OF COMPUTING AND COMMUNICATIONS

FACULTY OF MATHEMATICS, COMPUTING AND TECHNOLOGY

THE OPEN UNIVERSITY

Tabletop Tangible Interfaces for Music Performance: Design and Evaluation

---

Author:

Anna Xambó Sedó

MSc (Universitat Pompeu Fabra)

BA, MA (Universitat de Barcelona)

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

Supervisors:

Robin Laney (The Open University, UK)

Chris Dobbyn (The Open University, UK)

Sergi Jordà Puig (Universitat Pompeu Fabra, Spain)

Examiners:

Eduardo Reck Miranda (Plymouth University, UK)

Janet van der Linden (The Open University, UK)

Submitted 30 September 2014

Abstract

This thesis investigates a new generation of collaborative systems: *tabletop tangible interfaces (TTIs) for music performance* or *musical tabletops*. Musical tabletops are designed for professional musical performance, as well as for casual interaction in public settings. These systems support co-located collaboration, offered by a shared interface. However, we still know little about their challenges and opportunities for collaborative musical practice: in particular, how to best support beginners or experts or both.

This thesis explores the nature of collaboration on TTIs for music performance between beginners, experts, or both. Empirical work was done in two stages: 1) an exploratory stage; and 2) an experimental stage. In the exploratory stage we studied the Reactable, a commercial musical tabletop designed for beginners and experts. In particular, we explored its use in two environments: a multi-session study with expert musicians in a casual lab setting; and a field study with casual visitors in a science centre. In the experimental stage we conducted a controlled experiment for mixed groups using a bespoke musical tabletop interface, SoundXY4. The design of this study was informed by the previous stage about a need to support better real-time awareness of the group activity (*workspace awareness*) in early interactions. For the three studies, groups musical improvisation was video-captured unobtrusively with the aim of understanding natural uses during group musical practice. Rich video data was carefully analysed focusing on the nature of social interaction and how workspace awareness was manifested.

The findings suggest that musical tabletops can support peer learning during multiple sessions; fluid between-group social interaction in public settings; and a democratic and ecological approach to music performance. The findings also point to how workspace awareness can be enhanced in early interactions with TTIs using auditory feedback with ambisonics spatialisation.

The thesis concludes with theoretical, methodological, and practical implications for future research in New Interfaces for Musical Expression (NIME), tabletop studies, and Human-Computer Interaction (HCI).

Acknowledgements

Preface

The work presented in this thesis has led to the following publications, in chronological order:

1. Xambó, A., Laney, R., Dobbyn, C., & Jordà, S. (2011, July). Collaborative music interaction on tabletops: An HCI approach. In *BCS HCI 2011 Workshop on When Words Fail: What Can Music Interaction tell us about HCI?* Newcastle Upon Tyne, UK.

2. Xambó, A., Laney, R., Dobbyn, C., & Jordà, S. (2012, September). Towards a taxonomy for video analysis on collaborative musical tabletops. In *BCS HCI 2012 Workshop on Video Analysis Techniques for HCI*. Birmingham, UK.

3. Xambó, A., Laney, R., Dobbyn, C., & Jordà, S. (2013). Video analysis for evaluating music interaction: Musical tabletops. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 241–258). London: Springer-Verlag.

4. Xambó, A., Hornecker, E., Marshall, P., Jordà, S., Dobbyn, C., & Laney, R. (2013). Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction*, *20(6)*, 34.

5. Xambó, A., Roma, G., Laney, R., Dobbyn, C., & Jordà, S. (2014, June–July). SoundXY4: Supporting tabletop collaboration and awareness with ambisonics spatialisation. In B. Caramiaux, K. Tahiroglu, R. Fiebrink, & A. Tanaka (Eds.), *NIME '14: Proceedings of the 14th International Conference on New Interfaces for Musical Expression, 30 June – 4 July 2014, London* (pp. 40–45). London: Goldsmiths, University of London.

The first two publications are position papers, which briefly describe the work of this thesis at early stages. Part of the methodology chapter (Chapter 3) has been published in the third publication, a book chapter. Part of the first study on the Reactable (Chapter 4) has been included in the fourth publication, a journal article. Finally, part of the last study on ambisonics spatialisation (Chapter 6) has been included in the last publication, a conference paper.

The study presented in Chapter 4 was held at the Music Technology Group - Universitat Pompeu Fabra in Barcelona (Spain), of which an ethics approval was not required. The research of the other two studies that are respectively presented in Chapters 5 and 6 was held in the UK, and required ethics approval. For each of the two studies, there are listed below the project title and the reference number that were assigned by the OU Human Resources Ethics Committee (HREC). Dr Duncan Banks, the chair of the HREC, signed both:

- Chapter 5: "Evaluating collaboration on interactive musical tabletops - Evaluating the Reactable in the INTECH Science Centre (Winchester, UK)", HREC/2011/#906/1.

- Chapter 6: "SoundXY4: Supporting real-time auditory feedback in musical tabletops", HREC/2013/1493/Xambó/1".

There have been the following posters, oral presentations, and dissemination activities associated to the work presented in this thesis, in chronological order:

- Xambó, A. (2011, March 8). *Designing and evaluating interactive systems: Musical tabletops for collective music performance.* Poster session presented at The Open University Poster Competition 2011, Milton Keynes, UK.

- Xambó, A. (2011, May 17). *Tabletop groupware for music performance: Design and evaluation.* Oral presentation presented at the 2011 Doctoral Workshops Conference, The Open University, Milton Keynes, UK.

- Xambó, A. (2011, June 16). *Tabletop groupware for music performance: Design and evaluation.* Oral presentation presented at the CRC PhD Student Conference 2011, The Open University, Milton Keynes, UK.

- Xambó, A. (2011, June 17). *Designing and evaluating interactive systems: Musical tabletops for collective music performance.* Poster session presented at the CRC PhD Student Conference 2011, The Open University, Milton Keynes, UK.

- Xambó, A. (2011, 4 July). *Collaborative music interaction on tabletops: An HCI approach?* Oral presentation presented at the BCS HCI 2011 Workshop on When Words Fail: What can Music Interaction tell us about HCI?, Newcastle Upon Tyne, UK.

- Xambó, A. (2012, 10 January). *Tangible Additive Sound Synthesis (TASS)* Demonstration presented at the Welcome to the French Embassy, The Open University, Milton Keynes, UK.

- Xambó, A., Roma, G., & Bovermann, T. (2012, 15 April). *Tangible musical interfaces with SuperCollider*. Workshop presented at the SuperCollider Symposium 2012, Goldsmiths, University of London, London.

- Xambó, A. (2012, 12 June). *Collaboration on interactive tabletops for music performance: An exploratory study*. Oral presentation presented at the CRC PhD Student Conference 2012, The Open University, Milton Keynes, UK.

- Xambó, A. (2013, 2 June). *Tabletop groupware for music performance: Design and evaluation*. Oral presentation presented at the CRC PhD Student Conference 2013, The Open University, Milton Keynes, UK.

- Xambó, A. (2013, 11 November). *Tabletop tangible interfaces for music performance and implications for tabletop research*. Oral presentation presented at the School of Computing, University of Kent, Kent, UK.

- Xambó, A. (2014, 9 April). *Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface*. Oral presentation presented at the London Knowledge Lab, London.

- Xambó, A. (2014, 30 April). *Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface*. Oral presentation presented at the CHI 2014 Conference, Toronto, ON.

- Xambó, A. (2014, 1 July). *SoundXY4: Supporting tabletop collaboration and awareness with ambisonics spatialisation*. Oral presentation presented at the NIME 2014 Conference, London.

Contents

List of Figures

List of Tables

List of Abbreviations

CSCM	Computer Supported Collaborative/Cooperative Music
CSCW	Computer Supported Cooperative Work
DAW	Digital Audio Workstation
DI	Diffused Illumination
DIY	Do It Yourself
DMI	Digital Musical Instrument, Digital Musical Interactions
FSR	Force-Sensing Resistor
FTIR	Frustrated Total Internal Reflection
GUI	Graphical User Interface
HCI	Human-Computer Interaction
IR	InfraRed
MCRpd	Model-Control-Representation physical digital
MIDI	Musical Instrument Digital Interface
MUI	Material User Interface
MVC	Model-View-Controller
NIME	New Interfaces (for) Musical Expression
NVC	Non-Verbal Communication
OSC	Open Sound Control
SD	Standard Deviation
SMC	Sound (and) Music Computing
SMI	Smart Material Interface
TTI	Tabletop Tangible Interface
TUI	Tangible User Interface
VC	Verbal Communication
WIMP	Windows Icons Menus Pointer

To my parents Leonor and Sebastian,
and to Gerard, my half

Chapter 1

Introduction

Group collaboration has been always important in traditional ensembles ranging from pop-rock bands, to jazz combos, to classical music. It has not been until recent years that computers have supported collaboration, in particular collaborative music. The design and evaluation of *collaborative musical interfaces*, also known as *computer-supported collaborative music (CSCM)* [Barbosa, 2003], has been a topic of research in recent years within the field of New Interfaces for Musical Expression (NIME).

Tabletop tangible interfaces (TTIs) for music performance or ***musical tabletops*** are a special case of CSCM systems, and a type of TTIs. They provide a physical shareable interface for music performance that every user can see and manipulate, using tangible objects or fingertips, in the same way as a group would work on a table, but enriching the experience with digital technologies. The TTI approach contrasts with another type of CSCM systems, *network music* systems [Barbosa, 2006; Weinberg, 2005], that generally rely on interconnected individual interfaces or virtual shared spaces. They also differ from other tabletop group activities, in that musical interaction can be highly computationally demanding.

CSCM systems are generally designed either for supporting beginners [Blaine & Fels, 2003], or experts [Bongers, 1998; Collins et al., 2007], or mixed groups with predefined hierarchical musical roles [Weinberg et al., 2002]. TTIs are

1

a versatile technology that can be used to support both novices and experts because: 1) starting to manipulate tangible objects does not require computer literacy, which can benefit beginners; 2) the interface design can offer a complex level of interaction, which can benefit experts; and 3) the tabletop system offers a shared space, which can benefit mixed groups simultaneously.

A number of tabletop systems have been developed for music performance, and as with other CSCM systems, TTIs can be experienced in live venues or in museums. Even though they are suitable for collaboration, there is little research on how to best design and evaluate collaboration in TTIs for music performance. Moreover, since these systems are generally designed by practitioners, and for practitioners, there is little research focusing on the collaborative practice itself. In particular, the thesis focuses on understanding the nature of collaboration during musical performance using a shared interface, and how a TTI might be designed to best support both beginners and experts.

1.1 Motivation

The motivation of this research is twofold:

1. To find *new approaches to music making*[1] that can involve beginners and experts with no predefined roles.

2. To find points of connection between a complex and open activity, such as tabletop musical interaction, and collaborative systems in Human-Computer Interaction (HCI).

These are detailed respectively in two sections: democratic tabletop music (§1.1.1), and musical tabletop interaction and HCI (§1.1.2).

---

[1] *New approaches to music making* refers here to both the potential creation of new music from a novel musical instrument, and also to new ways of creating music shaped by interactions with a novel musical instrument. The thesis focuses on collaboration using musical tabletops, in particular on the interactions resulting from the use of this technology. It is out of the scope of this thesis to assess the resulting quality of the musical output, and thus we exclude the aesthetic value of the musical outcome.

1.1.1 Democratic tabletop music

My personal experience of traditional music education in classical music was that it was based on mechanical learning of music theory concepts and techniques, and less on personal creativity. Traditionally, the focus remains on becoming a skilful interpreter of the musical works of others and learning about their work. The creative activity of composing or performing one's own music is less encouraged, and then only after a certain level of musical knowledge has been attained: for composition, for instance, specialised composition courses, built upon music theory, are recommended. By contrast, in jazz improvisation and more contemporary musical practices there is more room for creating one's own music, but this again requires mastering a musical language and an instrument for a successful group performance [Bailey, 1980/1993]. However, digital technologies make it possible to create one's own music, without extensive knowledge of musical theory or hours of practice, by means of more intuitive and appropriately constrained musical interfaces.

Computer music, understood as producing music using a computer system, has increased the number of practitioners who do not necessarily have a specialised music education. Technology has democratised the access to music creation. This democratisation of musical creation with computers has partly been the result of the advent of powerful and affordable desktop computers and audio technologies, along with more contemporary practices of music making [Cox & Warner, 2004; Spiegel, 1998]. However, these practices are still very much oriented to personal digital audio workstations (DAWs). The use of DAWs implies, generally, a certain level of computer literacy at the expense of creativity. Learning how to control a complex DAW may interfere with the creative process.

Musical tabletops are a promising and exciting new environment for exploring collaborative real-time music performance[2] that can attract beginners, experts, or both. A detailed examination of musical tabletop interaction, focusing on the interactions and demands of beginner, expert, and mixed groups, may be useful for the NIME community for a better understanding of how to support collaborative practices under democratic principles on a shared interface.

1.1.2 Musical tabletop interaction and HCI

Music interaction with traditional musical instruments is an example of rich and complex human-artefact interaction. Buxton's seminal paper proposes that understanding musicians' demands and expectations can help designers of computer music systems, and can lead to the development of more effective systems in areas outside music that are also physically demanding (e.g. creative applications):

> I also discovered that in the grand scheme of things, there are three levels of design: standard spec, military spec and artist spec. Most significantly, I learned that the third, artist spec, was the hardest (and most important). If you could nail it, then everything else was easy. [Buxton, 1997, p. 10]

A follow-up of Buxton's ideas in computer music can be found in Jordà [2008]'s motivation for designing *digital musical instruments* (DMIs), including musical tabletops; or in recent publications such as the *Music Interaction* book [Holland et al., 2013]. In these examples, the fields of HCI and NIME can mutually inform each other.

---

[2]With *real-time music*, any change done to the music creation is heard simultaneously while playing. In contrast, *non real-time music* would show changes in the music creation asynchronously. Given their different nature, it is beyond the scope of this research to approach non real-time music.

One core element within the NIME and HCI disciplines is *interaction*, which can be understood as: 1) interactions with the system; and 2) interaction between players, mediated by the system. Interaction is thus the point of connection between users and a computing system, allowing for real-time mutual actions. An example of complex interaction is found in design: shaping an idea in the early stages of a creative process is called *sketching* [Buxton, 2007]. An analogy to sketching in music is *musical improvisation*. Musical improvisation, as sketching, is a collaborative open activity, and itself a creative process. In tabletop musical improvisation, we can consider that the role of a TTI is supporting the activity of *sketching musical ideas*. The possibility of sketching the musical ideas of a group on a TTI entails complex interactions, including interactions with the system as well as interaction between players, both of which ideally the system should be able to support. The bodily interactions emerged from operating with TTIs links to the notion of *embodied interaction* [Dourish, 2001] to refer to interaction with tangible user interfaces (TUIs) that involve both tangible but also social interaction.

A detailed examination of musical tabletop interaction, focusing on the particularities of group musical practice, may be useful to the HCI community for a better understanding of tangible interaction within highly demanding contexts. The findings of such a study of musical interaction on a TTI would inform the design of future tabletop systems. In particular, this approach can be useful for designing creative tabletop applications that computationally implement, adapt, or expand creative group activities.

In the following sections we present the scope of this research (§1.2), terminology used (§1.3), research questions (§1.4), and thesis structure (§1.5).

1.2   Scope and limitations

The scope and limitations of this study are closely related to the influential framework for real-time groupware by Gutwin & Greenberg [2002], and how these authors stated their study limitations, which include: *types of systems*,

types of environment, *types of tasks*, *people's location*, and *types of groups*. In addition, our limitations include the *application domain* and *types of interaction*. Each of the following limitations comprises the area of focus complemented with examples.

- **Systems: *real-time collaborative interactive systems for music performance*.** Here, we refer to systems that support real-time music performance [Jordà, 2007], in particular looping systems. Chadabe [1984] coined the term *interactive composing systems*, in which "the computer responds to the performer and the performer reacts to the computer, and the music takes its form through the mutually influential, interactive relationship" [Chadabe, 1984, p. 23]. We focus here on collaborative interactive systems applied to real-time music performance, in which interfaces allow *multithreaded* (i.e. various computational tasks in parallel) and *shared control* [Jordà, 2007], where there is a mutually influential relationship between performer and interface.

- **Workspace environment: *horizontal tabletop shared workspaces*.** Horizontal tabletop systems contrast with vertical displays [Müller-Tomfelde, 2010]. Here, we work with horizontal table shape systems, in which participants are generally standing.

- **Application domain: *improvisational and open tasks in tabletop settings*.** This research focuses on music, although it shares characteristics with other areas in tabletop research related to open and collaborative tasks and that involve the use of sound: for example, story-telling, sonic design, or collaborative design augmented by auditory feedback.

- **Interaction: *tangible interaction with constructive blocks*.** *Tangible constructive blocks* allow users to build structures by connecting tangible objects. For example, Zuckerman et al. [2005]'s work explores computationally enhanced building blocks that can simulate dynamic behaviours. We here focus on using tangible constructive blocks used to build 'musical instruments' with autonomous behaviours on tabletops.

- **Tasks:** *generation, execution, and discovery*. Musical performance relates to generation and execution activities, a term discussed in Gutwin & Greenberg [2002]. In particular, this research investigates musical improvisation using looping systems, entailing creativity, performance, and discovery tasks (e.g. physical manipulation of objects for generating sounds with or without effects).

- **Time and location:** *synchronous and co-located*. Barbosa [2003] proposed a classification space for computer-supported collaborative music based on the user's location (*co-located* vs *remote*) and group's interaction time (*synchronous* vs *asynchronous*). We here focus generally on groups located in the same space, interacting at the same time on the same interface.

- **Groups:** *small groups, mixed-focus collaboration, beginners and experts*. Small groups of between two to four people, occasionally five to seven, perform on our medium-size tabletop workspaces. These groups of beginners, experts, or both collaborate by switching between focusing on individual and shared work [Gutwin & Greenberg, 2002].

1.3 Terminology

As explained in the background chapter (Chapter 2), the terminology and methodology used is derived from literature in several research fields:

- *Co-located network music* for understanding the practice of co-located collaborative music (§2.2).

- *Tangible user interfaces (TUIs)* for understanding tangible interaction (§2.3).

- *Computer-supported cooperative work (CSCW) and tabletops* for understanding collaborative practices using tabletops (§2.4).

A summary of key terms used in this thesis is presented next, which will be described in more detail in the following chapter (Chapter 2):

- ***Tangible user interfaces (TUIs)***. Interfaces that "computationally augment physical objects by coupling them to physical data" [Shaer & Hornecker, 2009, p. 4], and that can support simultaneous actions in real time, which facilitates collaboration. TUIs can display real-time feedback: visual, auditory or haptic. Both control (input) and visual feedback (output) happen at the same position like in the physical world, also known as *direct input*.

- ***Tabletop tangible interfaces (TTIs)***. TTIs are a subset of tangible user interfaces (TUIs) because they use computationally augmented physical objects on a tabletop surface. Interaction is possible via finger touch (e.g. single or *multitouch* input data) or object manipulation (e.g. tangible objects with generally markers as input data) or both (hybrid systems). Tabletop technology combines aspects of both TUIs and interactive multitouch surfaces [Shaer & Hornecker, 2009]. Thus, although multitouch surfaces use virtual objects rather than physical objects, they are conceptually considered a subset of TTIs. Both are based on direct input.

- ***Computer-supported collaborative music (CSCM)***. NIME systems built for supporting cooperation and collaboration.

- ***Musical tabletops***. Musical tabletops are a type of TTIs and a subset of CSCM systems that allow users to create music.

- ***Social interaction***. In this thesis we use the term *social interaction* to refer to how players and groups of players communicate and interact among themselves using both verbal and non-verbal communication mediated by a TTI within a music creation context.

- ***Workspace awareness (WA)***. Also referred in this thesis as *awareness*, this term is widely used in tabletop studies. We here use the definition by Gutwin & Greenberg, used in HCI and CSCW, which refers to "the up-to-the-moment understanding of another person's interaction with a shared workspace" [Gutwin & Greenberg, 2002, p. 412]. In tabletop studies, a shared workspace includes not only TTIs, but also traditional

tables [Gutwin & Greenberg, 2002; Scott et al., 2004]. In this thesis, we focus on WA in musical tabletops.

1.4 Aim of this thesis

This thesis addresses the following overarching research question:

> *What are the challenges and opportunities provided by TTIs for music performance among beginners and experts?*

The question is motivated by the observation from the literature and from practical experience, that TTIs are a promising platform for collaboration, particularly musical collaboration, that can attract both beginners and experts. However, it is still unclear how best to support collaborative music performance with heterogeneous groups on TTIs.

Our approach is to investigate how groups use TTIs for music performance in different contexts (e.g. short vs long term use, music studio vs public space), in order to understand better their current issues, and their potential. Our aim is to propose principles for the design of future TTIs that promote more effective musical collaboration in music performance for both beginners and experts.

Our approach focuses on a number of gaps found in the literature: a lack of studies of long-term real uses of complex TTIs for music performance because most studies focus on one-off interactions and 'low-entry level' (§3.5.2); a lack of studies that explore how TTIs for music performance can support collaboration between both beginners and experts (§3.5.1); a lack of studies that investigate the nature of social interaction with musical tabletops (§2.3.1); and little research on how to best support WA in musical tabletops (§2.4.2).

The above observations lead to the following subsidiary research questions:

1. ***What is the nature of group tabletop interaction (GTI) on TTIs during music performance among beginners and experts?***

2. ***In TTIs for music performance, how is workspace awareness (WA) manifested among beginners and experts?***

We have studied both subsidiary research questions in two distinct sets of circumstances:

1. *Long-term use* (e.g. music rehearsals).

2. *Short-term use* (e.g. museums, casual set-ups).

The goal of this research is therefore fourfold:

1. To identify the potential opportunities offered by musical tabletops as a democratic setting for music performance for beginners, experts, or both.

2. To identify the potential opportunities that emerge from a system that enables highly demanding hands-on group interaction, and the implications of this for HCI design.

3. To shed light on conceptual, methodological and practical aspects of supporting collaboration in musical tabletops for music performance.

4. To produce a set of conceptual, methodological and practical tools for practitioners, researchers, and *digital luthiers*.[3]

In particular, the thesis aims to shed light on the design of TTIs and CSCM systems for collaborative musical practices. Three studies were conducted, focusing on understanding collaborative musical practice of three different kinds: expert musicians improvising music over multiple sessions using the Reactable, a commercial TTI and one of the most popular tabletops (Chapter

---

[3]A digital luthier is a term coined by Jordà [2005a] to refer to a person who makes or re-makes digital musical instruments.

4); casual visitors creating music with the Reactable in a science centre (Chapter 5); and mixed groups improvising with a bespoke TTI, the design of which was influenced by the findings of the previous two studies (Chapter 6).[4]

Lastly, our results have wider implications for HCI design in collaborative systems that support organisational work, layout design, brainstorming, or entertainment. Studies of tabletop collaboration often tend to focus on improving the efficiency of a specific technique or performing a task on a novel interface, rather than understanding the nature of a practice enabled by the interface. This thesis brings a different approach to understanding computer-mediated collaboration in physically demanding activities, such as music performance.

1.5 Thesis roadmap

The subsequent chapters of this thesis are structured as follows:

- *Chapter 2* maps the literature on existing collaborative interactive systems related to our research, focusing on salient design aspects of tangible music and TTIs for collaboration.

- *Chapter 3* presents an overview of the methodology used for collecting and analysing data in the three studies included in this thesis.

- *Chapter 4* presents a multi-session exploratory study with expert musicians using the Reactable in a casual lab set-up.

- *Chapter 5* presents an exploratory study with naïve users using the Reactable in a science centre.

- *Chapter 6* presents a study with a bespoke TTI, SoundXY4, with heterogeneous groups in a casual lab set-up.

---

[4]Given the novelty and scarcity of musical tabletops, it is difficult to find expert players of these instruments. In music, as in language learning, the more musical background you have and musical instruments you play, the easier to learn how to play new instruments. In this thesis, experts refer to those with solid musical background, with the exception of Chapter 4 in which it also refers to expertise with the Reactable.

- ***Chapter 7*** summarises the results achieved and sheds light on future research in TTIs for music performance.

Chapter 2

Background: A Survey of Collaborative and Tabletop Tangible Interaction

In this chapter, we survey the literature on collaborative music and tabletop tangible interaction from the perspectives of HCI, CSCW, and NIME. The survey includes systems that support collaborative practices, focusing on CSCM for music performance, and particularly on musical tabletops. This survey helps us to better understand the state of the art of collaborative systems design for real-time music.

2.1 Introduction

In the previous chapter, we discussed the motivation for this thesis and presented research questions and goals. In this chapter, we explain certain key terms and characteristics of musical tabletops and related systems. We first survey the literature on computer-supported cooperative work for music (CSCM) systems, in particular co-located network music and tangible user interfaces (TUIs) for music performance. We argue that co-located network music is relevant here, as it has a long tradition of supporting collaborative

practices. Although TUIs and musical tabletops support collaboration, there is little research on actual collaborative practices in these contexts. Thus, we borrow some terms from well-established CSCM practices, for example network music, as some of the terms used in this area are relevant to our research (§2.2). As stated in the previous chapter, musical tabletops are a subset of TUIs and tabletop tangible interfaces (TTIs). We present general concepts in the discipline of Human-Computer Interaction (HCI) of TUIs (§2.3) and TTIs within the field of computer-supported cooperative work (CSCW) (§2.4). Finally we map the literature on musical tabletops (§2.5), and detail the Reactable, which has been used in two studies in this thesis. Figure 2.1 shows the overall taxonomy of the systems and key terms presented in this chapter.
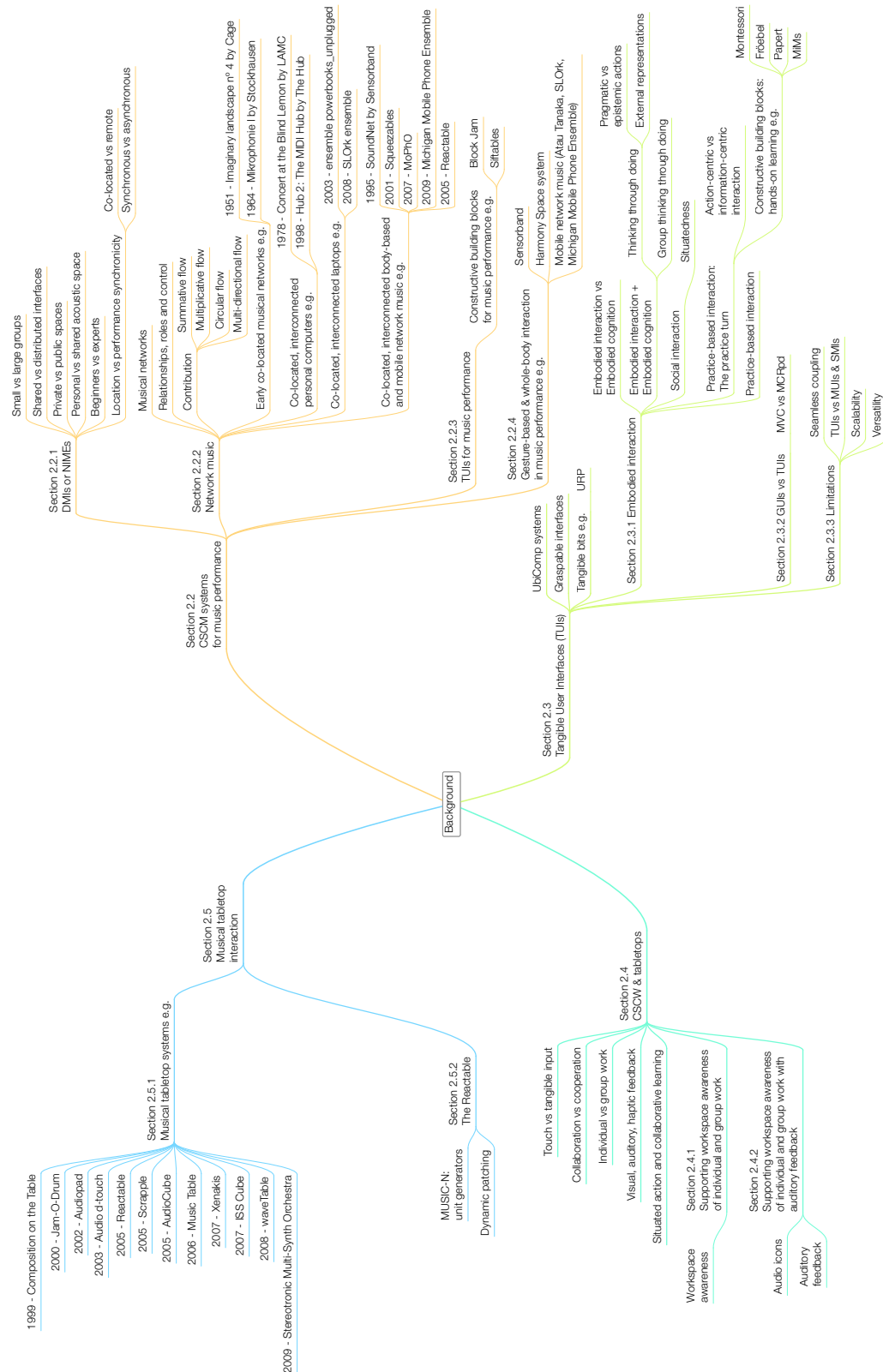
FIGURE 2.1: Diagram of the taxonomy of surveyed systems and key terms from this chapter, by sections

2.2 CSCM systems for music performance

2.2.1 DMIs or NIMEs for collaboration: key terms and characteristics

Recently, music technology has become more affordable, with a wide range of musical software and hardware available. In addition, practitioners tend to create their own DIY musical systems: for example, *Digital Musical Instruments* (DMIs) [Miranda & Wanderley, 2006], *Digital Musical Interactions* (DMIs) [Gurevich & Cavan Fyans, 2011], or *New Interfaces for Musical Expression* (NIMEs) [Fels, 2004]. In this paper, we use the NIME and DMI terms interchangeably, because we are only interested in the collaborative aspects of these systems in general.

By contrast with the rigidity of acoustic and traditional musical instruments, which usually need years of practice for mastery, less demanding, novel musical instruments can be designed which still produce satisfactory sound (although, naturally, this is not always the case, as novel instruments requiring expert technique can also be designed). Novel musical instruments explore new design criteria, such as adding more constraints to the interface [Gurevich et al., 2012; Magnusson, 2010], or offering a modular design adapted to the user' needs [Jordà, 2008]. With DMIs or NIMEs, it is also possible to challenge some principles of traditional musical instruments, such as musical interfaces designed for individual use, and explore new aesthetics: for example, with CSCM systems in particular, shareable interfaces for music [Jordà, 2008; Weinberg, 2005].

The main trade-offs of CSCM systems include *group size* (small vs large groups), *interface control* (shared vs distributed interfaces), *privacy* (private vs public spaces), *audio delivery* (personal vs shared acoustic spaces), *expertise* (beginners vs experts), and *space vs time* (location vs performance synchronicity). We now review each of these concepts:

- **Group size:** *small vs large groups.* CSCM systems are designed for a minimum of two or more players [Blaine & Fels, 2003]. Weinberg [2005]

distinguishes between *small-scale local systems* (three to ten players) and *large-scale local systems* (more than ten players). The main difference highlighted by Weinberg is the level of intimacy: in small-scale systems, interpersonal interactions can be more fine-grained and subtle. Blaine & Fels [2003] discuss the term *scalability* and how factors such as turn-taking procedures and gesture-sound mappings can vary depending on the group size.

- **Interface control:** *shared vs distributed interfaces*. With a few exceptions, such as the piano, it is unusual to share the interface of a conventional instrument; and even in the case of the piano, there still needs to be a certain compromise (i.e. distribution of space or roles). This also happens within CSCM systems based on a co-located shared interface [Jordà, 2005b].

- **Privacy:** *private vs public spaces*. The addition of private spaces to a CSCM interface allows players to have an individual work space within which to make their musical contributions [Fencott & Bryan-Kinns, 2010].

- **Audio delivery:** *personal vs shared acoustic space*. Options range from using headphones, loudspeakers, or both. Using headphones corresponds to providing a personal, individual acoustic space [Morris et al., 2004]. Using speakers can correspond to moving from a personal to a more shared acoustic space depending on: 1) the position of the speakers (e.g. individual vs global positions); and 2) the channel layout of the delivered audio (e.g. individual vs mixed channels). Some examples of experiments with different audio delivery set-ups can be found [Fencott & Bryan-Kinns, 2012] including TTIs [Blaine & Perkis, 2000; Hancock et al., 2005].

- **Expertise:** *beginners vs experts*. The expectations of beginners and experts are different: beginners prefer highly constrained interfaces and prioritise the collective experience, whilst experts prefer more complex

interfaces that allow virtuosic expression [Blaine & Fels, 2003]. Designing for both requires a trade-off. Some of the CSCM systems distribute roles according to expertise (e.g. Beatbugs by Weinberg et al. [2002]).

- **Space vs time:** *location vs performance synchronicity*. The seminal work by Barbosa [2006] presents *computer-supported cooperative work for music applications*, which adapts concepts of CSCW to collaborative music systems (CSCM). Barbosa proposes a classification based on the dimensions of space and time, and inspired by a classification space for CSCW: *location* of players (*co-located* vs *remote*), and *performance synchronicity* (*synchronous* vs *asynchronous*), as shown in Fig. 2.2. Barbosa focuses on examples of networked music systems, so there is little detail on the specifics of musical tabletops, although the classification is also relevant to tabletop systems.

  - *Co-located vs remote*. *Co-located* refers to a mode in which players are located in the same physical space (e.g. same room or venue), whilst in the *remote* mode, players are located in different places. Barbosa [2003] refers to these two conditions as *off-line* and *on-line* modes respectively. In this thesis we avoid the terms off-line and on-line because co-located networks can be also connected to the Internet, so the distinction is blurred.

  - *Synchronous vs asynchronous*. As defined by Barbosa [2006, pp. 8–9], in *synchronous* mode players are active simultaneously on a joint project, whilst in *asynchronous* mode players do not need to be active simultaneously, although the system should support synchronous activity, even if players want to remain inactive at times.

In this section we reviewed the characteristics of NIMEs for collaboration and their trade-offs: what is the group size, how the interface is controlled, whether there are private spaces, how the audio is delivered, whether it is designed for beginners or experts or both, and what is the location and synchronicity of performers. These characteristics are helpful for understanding how wide can be the range of NIMEs for collaboration, and for later
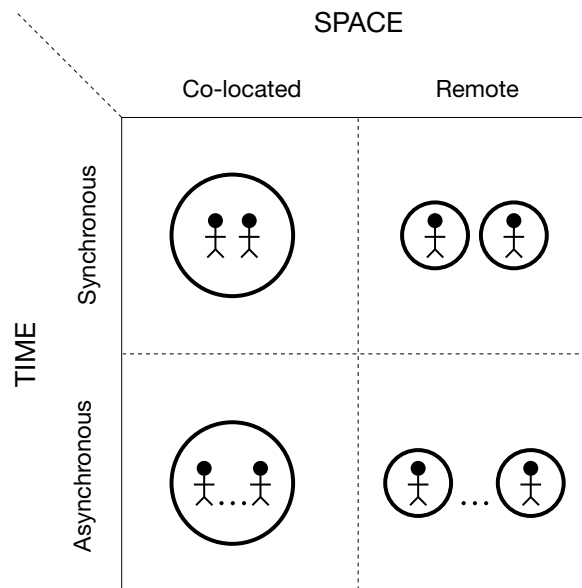
FIGURE 2.2: Adaptation of Barbosa's classification matrix for computer-supported collaborative music [Barbosa, 2006]

understanding similarities and differences between existing systems within network music, tangible music, and tabletop tangible music: *Network music* (§2.2.2) and *tangible music* (§2.2.3) are two of the most common approaches to CSCM. In particular, *tabletop tangible music* borrows concepts of both areas as explained later in this chapter (§2.5).

2.2.2 Network music

Network music is a common and probably the most well-known approach to CSCM. Here, we discuss the key terms and characteristics of network music, such as *musical networks*, types of *relationships, roles and control*, and types of *contribution*. We also present a range of examples from early co-located musical networks with analogue technologies; to the advent of computers and their use; currently including laptops, mobile technologies, sensors, or the Internet.

Musical networks. In computer music, the term *musical networks* refers to those environments with musicians and computers connected by a network,

independent of the musicians' location [Barbosa, 2003; Jordà, 2005b; Rohrhuber, 2007; Weinberg, 2005]. Weinberg refers to *interdependence* as a key term of a musical network because players can influence, share, and shape each other's music in real time [Weinberg, 2005; Weinberg et al., 2002], and, hence, they have partial control of the overall result. Musical networks can be remote or local [Barbosa, 2003; Weinberg, 2005]. Here, we focus on local musical networks for their close association to the topic of co-located musical tabletops and their players' interdependencies, although in musical tabletops there is no network strictly speaking. *Local musical networks* are defined by Barbosa [2006] as "groups of performers who interact in real-time, in the same physical location, on a set of musical instruments, with the possibility of sonic interdependence provided by a fast local computer network" [Barbosa, 2003, p. 57]. In musical networks, musicians supported by computers tend to be identified as nodes of the network. Accordingly, research in this field tends to explore the nature of mutual interactions between nodes. Musical tabletops are generally co-located, although some exceptions include e.g. the Reactable, when it is played remotely [Jordà et al., 2005]. Furthermore, musical tabletops generally use one table only, with some exceptions, such as Iwai [1999]'s Composition on the Table, which includes four tables in a local setting. The effect of an interdependent musical network is described in the following way by the members of the network music band The Hub: "when the elements of the network are not connected, the music sounds like three completely independent processes, but when they are interconnected, the music seems to present a 'mind-like aspect'"[Bischoff et al., 1978, p. 28]. Like a musical network, a musical TTI can promote interdependence, where structures created with tangible objects (or musical instruments) can be mutually modified on a shared interface. The main difference between musical tabletops and network music systems is that a co-located musical tabletop is a single node that allows for multiple interconnected instruments managed by a single computer. Thus a number of musical network principles can be adapted to TTIs.

Relationships, roles, and control. Weinberg [2005] presents a theoretical framework for musical networks, introducing a set of principles that focus
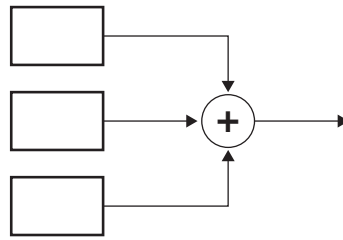
FIGURE 2.3: Summative or parallel flow
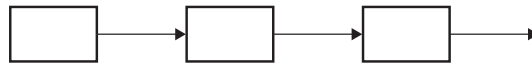
FIGURE 2.4: Multiplicative or serial flow

on the socio-political organisation of the networks (e.g. centralised vs decentralised). This framework is useful for an understanding of the relationships, roles and multi-dimensional control aspects between users in musical networks.

Contribution. In collaborative music, a distinction can be drawn between *summative* contribution and *multiplicative* contribution [Jordà, 2005b]. There is also a circular contribution [Bischoff et al., 1978], and a multi-directional contribution [Gresham-Lancaster, 1998]. Here, we use the term *flow* instead of contribution, to refer to something that emerges from a number of contributions:

- ***Summative flow***. The sum of individual contributions, or independent processes, where there is little mutual interaction, as shown in Fig. 2.3.

- ***Multiplicative flow***. The product of individual contributions, or the serial processing of interdependent processes, with more mutual interaction than in the summative approach, as shown in Fig. 2.4.

- ***Circular flow***. As shown in Fig. 2.5, an instance of multiplicative flow, in which there is a circular sequence of the serial processing of interdependent processes.
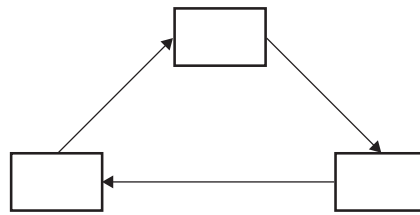
FIGURE 2.5: Circular flow
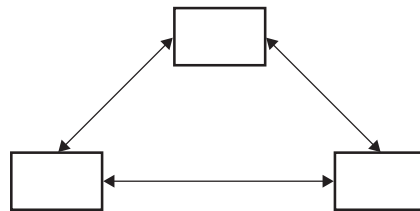
FIGURE 2.6: Multi-directional flow

- *Multi-directional flow*. Players interact with each other by sending and receiving data, and therefore more mutual interaction is possible compared to in the above approaches, as is shown in Fig. 2.6.

After the above overview of the key terms and characteristics of network music, we next present different examples of network music, from using analogue technologies or computers, to more bodily-based interactions led by mobile technologies or sensors.

Early co-located musical networks. Early examples of musical networks include classic avant-garde music pieces based on shared control, as well as interdependent and egalitarian roles, with the composer as the director of the piece, which contrasts with more current approaches to CSCM. Here, we present two examples: *Imaginary landscape No. 4* (1951) and *Mikrophonie I* (1964):

- *Imaginary landscape No. 4 (John Cage, 1951)*. This represents an early co-located face-to-face musical network that dates back to the 1950–60s using analogue technologies. The composer, John Cage proposed Imaginary Landscape No. 4 (1951) to be played by 24 performers with 12

transistor radios [Cage, 1961]. Each pair of musicians was in charge respectively of changing the frequency and amplitude of one radio unit (see Fig. 2.7). Changes were based on a set of instructions provided by the composer [Cage, 1961, pp. 57–59]. This approach is summative: the interdependency emerges from playing the radio unit as a shared interface.

- ***Mikrophonie I (Stockhausen, 1964)***. Another example is *Mikrophonie I* (1964) by Karlheinz Stockhausen, where six performers divided into three pairs were in charge of, respectively: 1) playing a tam-tam instrument; 2) capturing the resulting sounds with microphones; and 3) applying filters and potentiometers to the microphone output sounds [Burns, 2002; Jordà, 2005b]. This approach is thus multiplicative because each pair of musicians affects the other two pairs in a three-step chain of play, record, and filter.

In both examples, the actions of the musicians tend to be interdependent, and it seems that nobody takes the lead. Having said that, in these early works there is still a clear hierarchy between the composer and the performers, because the latter have to follow a script when performing the musical composition. In these two examples we can foresee some of the common issues in collaborative musical networks, such as the fact that nobody takes complete control of the musical piece (neither performers nor the composer). This is related with the notion of unpredictability and uncertainty in the musical process as described by Cage [1961], because the control is shared among a group. Furthermore, in the two pieces the parameters that performers can actually manipulate are very constrained and are interrelated. This means that the interdependencies between the performers are high, because the results of their actions depend on others' actions. Thus, the musical output is the product of the actions of the group as an interconnected team.

Co-located, interconnected personal computers. With the advent of affordable personal computers in the mid-1970s, groups like the League of Automatic Music Composers [Bischoff et al., 1978] and, later in the mid-1980s, The

FIGURE 2.7: Summative approach: two people sharing control of, respectively, the frequency and amplitude of a radio unit

Hub [Gresham-Lancaster, 1998] performed live using computer networks, e.g.

- ***Concert at the Blind Lemon (1978) by The League of Automatic Music Composers (LAMC)***. Three players performed with three interconnected KIM-1 microcomputers, with a different music program on each that offered musical input and output. There was a circular flow, in which each computer sent data to another computer and received data from another computer [Bischoff et al., 1978].

- ***Hub 2: The MIDI Hub (1998) by The Hub***. With this configuration, each player had a MIDI[1] port assigned with one input and one output. There was multi-directional flow, in which players could exchange (send to and receive) from one another MIDI channel messages with the identity of who was sending the message, that allowed participants to play the set-ups of others [Gresham-Lancaster, 1998].

The approach of The Hub and later groups to network music consisted of interconnecting individual players with their own technologies using protocols such as MIDI or OSC[2] [Wright & Freed, 1997], with an additional option of modifying others' compositions by networked mutual interactions in

---

[1]MIDI is a protocol defined in 1982 that allows communication between electronic musical instruments, controllers or computers by event messages.

[2]OSC is another protocol that allows communication between computers and other devices through networking technology. OSC is used as an alternative to MIDI, allowing more flexibility thanks to multiple data types (not only integers as MIDI) and a configurable naming scheme.

a more multi-directional approach. The Hub also experimented with remote networks using the Internet for communication and synchronisation. Egalitarian organisations emerged from this, an approach that can be useful for understanding the emergence of roles in musical tabletops because different musicians are, in principle, equally sharing the same interface.

Co-located, interconnected laptops. Laptops were also introduced into network music, and had a world-wide impact, as there are now ensembles and orchestras all over the world, including an international association of laptop orchestras.[3] Two representative approaches to laptop ensembles are the *ensemble powerbooks_unplugged* and the *SLOrk* ensemble:

- ***ensemble powerbooks_unplugged (2003)***. The Republic is a project that started in 2003 based on collaborative live coding [de Campo, 2014] using laptops and their built-in speakers. *Live coding* practices are based on the use of scripting languages for real-time music improvisation [Brown, 2006; Collins et al., 2007; Freeman & Van Troyer, 2011; Rohrhuber, 2007]. The Republic is based on an available extension library (with the same name) written in SuperCollider, a real-time audio synthesis environment and programming language [McCartney, 2002]. The Republic's principle is to create a symmetrical network, in which each player can access and modify each other's code, as if it were, in the words of Rohrhuber et al., *purloined letters*: "networked live coding makes code a means of communication and of collective musical thought. More literally, the program becomes a letter, a letter that circulates amongst the participants" [Rohrhuber et al., 2007, p. 4]. The approach is thus multi-directional. Both the ensemble *powerbooks_unplugged* and the ensemble *Republic 111* use the Republic's ideas and technology.

- ***SLOrk ensemble (2008)***. The Stanford Laptop Orchestra (SLOrk) was founded in 2008 and is directed by Ge Wang [G. Wang et al., 2009]. With the involvement of more than 20 laptops, participants also use controllers and custom multi-channel speaker arrays. The musical pieces

---

[3]`www.ialo.org/doku.php` (accessed 30 September 2014).

are primarily written in ChucK, a programming language also developed by G. Wang [2008]. The ensemble also uses the Small Musically Expressive Laptop Toolkit for rapid prototyping. In contrast with the egalitarian approach of the ensemble *powerbooks_unplugged*, in *SLOrk*'s performances there is a more hierarchical structure, because the ensemble conductor also has a role. The approach is, therefore, summative.

A major criticism of co-located musical networks with laptops is the lack of bodily interactions and the lack of transparency of the performer's action. This also applies by extension to network music with personal computers and the genre of *laptop music*. For example, a member of the audience commented about a performance of The Hub "The music was fantastic, but you looked like a bunch of air traffic controllers" [Gresham-Lancaster, 1998, p. 43]. Live coding represents a step forward to avoid *obscurantism* by projecting the performer's screen showing the code.[4] However, the performer's action can still be difficult to understand for an audience that is not code-literate. In contrast, musical tabletops promote bodily interactions that—combined with a projection of the interface—can arguably be understood by a wider audience, because it is a more familiar object-based representation and interaction. However, it will also depend on the clarity of the interface.

An opportunity brought about by laptop ensembles, and similar to acoustic musical instruments, is the exploration of the performers' positions among the audience, and thus performers can be treated as as a sound source distributed in the physical space of the room, which contrasts with solo laptop performances, where the performer usually sits on stage. Ensembles such as the powerbooks_unplugged or SLOrk explore the aesthetics of mixing performers with the audience within the space. This is possible because there is a number of group members. For example, powerbooks_unplugged expands the idea of creating a symmetrical network, by sitting among the audience with laptops and using their built-in speakers, so that the distinction between the performers and the audience is blurred. This ensemble's approach is to

---

[4]`http://toplap.org/wiki/Read_me_paper` (accessed 30 September 2014).

maintain *individual agency* in an intimate acoustic performance, which contrasts with the traditional approach of amplifying the sound using a set of multichannel speakers. SLOrk usually opts to build additional multi-channel speaker arrays of six speakers arranged in a hemisphere, each next to a performer, with the aim of enriching the laptop-based acoustic sound source. SLOrk's performances explore the distribution and location of performers, ranging from indoors to outdoors, on stage to mixed with the audience, or co-located to remote settings. The exploration of multiple locations of sound sources in performance is still underdeveloped in the case of musical tabletop's performances. This is partly because the table is a bulky artefact and it generally represents a single sound source point in a physical space. Yet, in our final Study 3 (Chapter 6) we explore the spatialisation of sounds for four performers using four loudspeakers, which potentially can be expanded to an audience who could also benefit from the immersive musical experience.

Co-located, interconnected body-based and mobile network music. Other approaches to co-located network music are based on smaller or larger scale technologies than laptops. Using devices other than laptops can facilitate bodily movements, as well as the exploration of the performance space. Such devices include systems that are based on body movements, for example: the work of Sensorband with *SoundNet*; interconnected special purpose electronic musical instruments, such as *Squeezables*; mobile devices such as *MoPhO*'s work or the Michigan Mobile Phone Ensemble; and interconnected musical tabletops such as the Reactable:

- ***SoundNet by Sensorband (1995)***. Sensorband was a group of three musicians (Atau Tanaka, Zbigniew Karkowski, and Edwin van der Heide) who together played the multi-user instrument Soundnet, built by Bert Bongers, together [Bongers, 1998]. Soundnet is a large web structure of interconnected ropes with 11 sensors at the top that could detect stretch and motion. Like spiders climbing over a web, the musicians had to climb the structure in order to produce music, leading to a high degree of mutual interaction, given that changes to a sensor depended

on the movements of the others. This approach is thus termed multi-directional.

- *Squeezables (2001)*. This system comprises a set of six squeezable balls, which can be controlled by continuous squeezing and pulling hand gestures [Weinberg & Gan, 2001]. The set is mounted on a rounded podium. Information about pressing and pulling the balls from force-sensing resistor (FSR) pressure sensors is converted to MIDI and transmitted to a central computer, with a patch built in Max (cf. [Puckette, 2002]), which maps the data into sound output. There are five accompaniment balls (synth, voice, theremin, arpeggio, and rhythm) and a melody ball: three of the accompaniment balls control the timbre-oriented parameters, while the other two control the rhythmic parameters. The melody soloist ball controls *contour*—which refers to the overall tonality or pitch curve of a scale and not to the actual pitches—along with the timbre parameters. The system provides different levels of interdependency depending on the musical role. The musical roles are marked by the type of ball, in which the five accompaniment balls have autonomous control and can influence the melody ball, yet input from other balls has no effect on their output. The melody ball can influence the behaviour of the accompaniment balls as well. It is thus a role-based circular flow.

- *MoPhO (2007)*. The Stanford Mobile Phone Orchestra (MoPhO) started in 2007 and is based on mobile music [Oh et al., 2010; G. Wang et al., 2008]. Their current set-up includes iPhones and wearable speakers (over the past their set-up included Nokia N95 smartphones). MoPhO usually perform pieces developed by the ensemble members using the MoMu toolkit. The pieces provide examples of summative and multi-directional flows.

- *Michigan Mobile Phone Ensemble (2009)*. The Michigan Mobile Phone Ensemble was founded by Georg Essl in 2009 [Essl, 2010]. Their set-up includes iPod Touches and iPhones, as well as custom-built wearable speaker systems. Their UrMus tool is used for rapid prototyping. Some

of the pieces are led by a conductor, and their approach is generally summative.

- ***Reactable (2005).*** This system is a musical tabletop [Jordà et al., 2005], and is further described in §2.5.2. The instrument allows performers to work on independent audio threads or channels (summative), as well as to share audio threads (multiplicative) [Jordà, 2005b]. This TTI can be played simultaneously in collaboration, both co-located and remotely.

As with the laptop ensembles, some of the above mobile ensembles also explore alternative performance set-ups to that of a traditional performance on stage in front of an audience. For example, in an eight-channel surround sound piece by the MoPhO ensemble, performers were surrounded by the audience and moved around with an iPhone and a wearable speaker each. Furthermore, Lee & Freeman explore the notion of the audience as performers by creating a musical performance for large-scale audience participation using networked smartphones [Lee & Freeman, 2013], an approach that is beyond the scope of this thesis. Mobile devices facilitate this exploration of different concert settings. Gesture-based and whole-body interaction is also explored in mobile network music, as further discussed in §2.2.4.

Table 2.1 summarises the different approaches to network music discussed in this section.

This section presented network music's characteristics that are useful for understanding the nature of sharing an interface on a musical tabletop:

- The level of interdependence between performers that exist in co-located musical networks.

- The possible performers roles organised as socio-political structures from hierarchical to egalitarian.

- The possible types of contribution or flow from independent to interdependent processes with more mutual interaction between performers.

TABLE 2.1: Network music systems

Year	System/Musical Piece/Ensemble	Contribution	Location	Synchronicity	Target	Roles	Technology	Support of privacy	# Performers[*]	Reference
1951	Imaginary landscape No. 4 by Cage	Summative	Co-located	Synchronous	Beginners, Experts	2 roles: frequency *vs* amplitude	12 transistor radios	None	24	Cage [1961]
1964	Mikrophonie 1 by Stockhausen	Multiplicative	Co-located	Synchronous	Experts	3 roles: tam-tam, microphones, effects	Tam-tam instrument, microphones, sound effects mixer	None	6	Burns [2002]
1978	Concert at the Blind Lemon by LAMC	Multiplicative (circular flow)	Co-located	Synchronous	Experts	3 roles: 3 diff. music programs with I/O	KIM-1 microcomputers, mixer	Yes (microcomputers screens)	3	Bischoff et al. [1978]
1995	SoundNet by Sensorband	Multi-directional	Co-located	Synchronous	Experts	Flexible	Large web structure with 11 sensors (stretching & movement)	None	3	Bongers [1998]
1998	Hub 2: The MIDI Hub by The Hub	Multi-directional	Co-located	Synchronous	Experts	6 roles: 6 diff. softwares, 1 lead player indicates new sections & hardware set-ups with MIDI ports (I/O)	Hub2: OpCode System's MIDI interface	Yes (MIDI channel messages)	6	Gresham-Lancaster [1998]
2001	Squeezables	Circular flow (role-based)	Co-located	Synchronous	Beginners	2 roles: 5 accompaniment balls, 1 melodic soloist ball	6 squeezable balls with FSR pressure sensors	None	3	Weinberg & Gan [2001]
2003	ensemble powerbooks_unplugged	Multi-directional	Co-located	Synchronous	Beginners, Experts	Flexible	Laptops, wireless network, SuperCollider	None	6	De Campo [2014]; Rohrhuber et al. [2007]
2005	Reactable	Summative, Multiplicative	Co-located, Remote	Synchronous	Beginners, Experts	Flexible	Musical tabletop	None (only headphones)	1–4 (larger groups in public venues or with remote settings)	Jordà [2008]; Jordà et al. [2005]
2007	MoPhO	Summative, Multi-directional	Co-located	Synchronous	Beginners, Experts	Flexible	iPhones, custom wearable speakers, MoMu toolkit	Yes (mobile phones screens)	From individual to large group (+12)	Oh et al. [2010]; G. Wang et al [2008]
2008	SLOrk ensemble	Summative	Co-located, Remote	Synchronous	Beginners, Experts	Flexible	Laptops, controllers, custom multi-channel speaker arrays, ChucK, Small Musically Expressive Laptop Toolkit	Yes (laptops screens)	Large group (+20)	G. Wang et al. [2009]
2009	Michigan Mobile Phone Ensemble	Summative	Co-located	Synchronous	Beginners, Experts	Flexible	iPod Touches/iPhones, custom wearable speakers, UrMus	Yes (mobile phones screens)	Large group (11–20)	Essl [2010]

[*] The number of performers is either defined in or derived from the description of a particular musical piece or network music system.

Then, we surveyed a selection of co-located network music systems from early pieces in the 1950s to present, systems closely related to experimenting with available technology, including radio units, microphones, microcomputers, the Internet, laptops, sensors, mobile devices, and tabletops. It is noticeable that, even though egalitarian principles have been present from the early pieces, over time systems have offered more flexible roles and more multi-directional communication between performers. In particular, this thesis investigates musical tabletops as a promising democratic platform for both beginners and experts. With the advent of mobile and sensor technologies in network music, bodily interactions have become more present contributing to reducing the criticised obscurantism of personal computer network music, along with the projection of performers screens. Bodily interactions are also key in tabletops and even more so in musical tabletops, which is an aspect of study of this thesis. Furthermore, mobile devices allow performers to explore the location of different sound sources in the performance space, an aspect that is also explored in this thesis.

2.2.3 TUIs for music performance

Tangible music is another approach to CSCM. An overview of tangible music systems is provided by Martin Kaltenbrunner on his website,[5] and by Shaer & Hornecker's survey of TUIs [Shaer & Hornecker, 2009]. Tangible music systems are grouped by the type of tangible objects used, including *building blocks*, *token-based sequencers*, *tangible musical artefacts*, or *toys*. Tangible objects have advantages over virtual representations because of their direct use of physical artefacts. Tangible music can encourage new forms of participation, as it usually requires no computer or musical literacy. TTIs for music performance can be seen as a subset of musical TUIs, because they add a tabletop surface as the centre of operations. We give a broader overview of TUIs in §2.3, and explain TTIs for music performance in §2.5. Here, we provide an overview of the TUIs that are relevant to our research, in particular constructive building blocks for music performance:

---

[5]`http://modin.yuri.at/tangibles` (accessed 30 September 2014).

FIGURE 2.8: Closeup of Sifteos

Constructive building blocks for music performance. Some musical projects, such as Sony's Block Jam [Newton-Dunn et al., 2003], or the musical application of Siftables [Merrill et al., 2007], have been investigated for the building of musical sequences. Block Jam works as a controller, where the input is operated by the tangibles, but the output is displayed on a separate screen. Siftables (currently also known as Sifteos, see Fig. 2.8) is a multi-purpose platform, which includes a music sequencer. Tangible objects with an embedded computer contain input and output that occur in the same place. In contrast, tabletop systems such as the Reactable [Jordà et al., 2005] incorporate a tabletop surface merging both, and where the input is operated by the tangibles, while the output is projected onto the table creating a 'digital shade' (see §2.5.2).

Of the myriad of available TUIs for music performance, we focus on constructive building blocks because their behaviour resembles and can inspire the design of sets of tabletop tangible objects. On the one hand, each block of a collection has an autonomous, audiovisual behaviour. On the other hand, it is possible to interconnect the blocks to build more complex musical structures. This modular approach is suitable for both beginners and experts, as discussed throughout this thesis.

2.2.4 Gesture-based and whole-body interaction in music performance

There exist a number of examples of interactive music systems where the interaction with the interface is based on gestures, including networked systems. For example, Sensorband performed computer music by using a range of gesture-based sensors (e.g. ultrasound, infrared (IR), and bioelectric sensors) [Bongers, 1998]. Another example is the Harmony Space system, in which harmony (e.g. bass notes or chords) can be explored using the whole body by stepping on the floor-projected interface [Holland et al., 2009]. According to Holland et al.'s study, the use of the whole body is a promising interaction style for collaboration in producing music. These examples show how collaborative music can be performed by means of interconnected bodily gestures, which is a relevant aspect of TTIs for music performance. However, there is little mention of the social aspects of these bodily interactions, which is of interest to us in this thesis.

In another example of systems for collaborative mobile music that promote gesture-based and whole-body interactions, Tanaka explores how collaborative musical creations can be performed in mobile and remote conditions with users connected to the same network [Tanaka, 2004a,b]. There are also examples of co-located settings with ensembles such as SLOrk or the Michigan Mobile Phone Ensemble, as presented above (§2.2.2). These ensembles offer a number of performances that experiment with the possibilities of the mobile medium using laptops or mobile devices, and in which organisational and performative aspects of these practices may inform the design of TTIs in similar real-time contexts.

2.3 Tangible user interfaces (TUIs) and HCI

It is worth noting that tangible user interfaces (TUIs), also known as *tangible interaction*, is a broad term that refers to the design of physical forms that

embed digital information. TUIs include a wide range of interesting inter-
faces from arm's reach scale (e.g. Topobo, a 3D construction kit with kinetic
memory [Raffle et al., 2004], or MoSo Tangibles, a set of artefacts for under-
standing sound concepts such as pitch, tempo and volume [Bakker et al.,
2011]) to body's reach scale (e.g. a tangible programming space [Fernaeus
& Tholander, 2006], or Buildasound,[6] a sound building blocks game). Given
the massive literature on the topic, this thesis specifically focuses on TUIs that
are arm's reach scale and that are based on constructive building blocks with
audiovisual feedback, as detailed in this section, as well as in §1.2 and §2.3.2.

We now present the context of tangible user interfaces (TUIs), or tangible bits
as a subset of UbiComp systems, and their origin stemming from graspable
interfaces.

UbiComp systems: *the invisible computer*. In the early 1990s, Weiser [1991]
coined the term *ubiquitous computing*, also known as UbiComp, to refer to
human-computer interactions in which the computer is less visible during the
interaction (e.g. wireless, embedded, wearable and/or mobile technologies).
In UbiComp systems, information can be accessed everywhere and is context
aware. Tangible user interfaces (TUIs) can be seen as an example of UbiComp
systems (cf. Shaer & Hornecker [2009]).

Graspable interfaces: *space-multiplexing input*. Graspable interfaces is a
term coined by Fitzmaurice [1996] in the mid-1990s to refer to the use of phys-
ical objects (*graspables*) to control a GUI interface. These are designed as an
alternative to the mouse. The approach here is to expand the features of the
virtual GUI objects to physical objects in a space (similar to the controls of an
audio mixing console). However, the results are still based on an individual's
manipulation of a GUI using physical objects, in which each object represents
one function: for instance, manipulating the position of a digital square by
dragging an object on a tabletop surface. Graspable interfaces are a precursor
of TUIs.

---

[6] `http://buildasound.info` (accessed 4 March 2015).

Tangible bits: *coupling representation and control.* The term *tangible user interfaces* (TUIs) refers to physical objects that can both control and represent digital information. In particular, the term *tangible bits* is used as opposed to *painted bits*, which refers to the representation offered by GUIs.

> Our attempt is to change "painted bits" into "tangible bits" by taking advantage of multiple senses and the multimodality of human interactions with the real world. We believe the use of graspable objects and ambient media will lead us to a much richer multi-sensory experience of digital information. [Ishii & Ullmer, 1997, p. 8]

URP [Underkoffler & Ishii, 1999] is one of the first examples of a TUI. URP is a system for urban planning, where physical objects represent buildings, and the system allows urban planners to create urban simulations with shadows or wind for visualisation and discussion. With TUIs there is direct interaction, which allow users to directly manipulate the digital information. This contrasts with traditional pointer-based desktop interfaces, a style of indirect interaction generally known as Window, Icon, Menu and Pointing devices (WIMP). With tabletop TUIs, instead, digital information (bits) is more manipulable, and representation and control are at the same location. In other words, the representation and control of data are combined in a single unit of meaning. An expected characteristic of TUIs is a *seamless coupling* between the digital and physical, between representation and control. The fact that digital information can be contained in physical objects makes data more shareable, as a physical object is accessible to anyone in its vicinity. As pointed out by Shaer & Hornecker [2009], the physicality of the interface in TUIs contrasts with the notion of the invisible interface in UbiComp systems, although TUIs are considered a subset of such systems. Thus the importance of materiality in the interface is a salient characteristic of TUIs that may not be found in other UbiComp systems. Thus, a TUI will offer an affordance that users may consciously prefer to work with.

We next discuss the novel interaction aspects introduced by TUIs in terms of the physical and social aspects that differentiate them from other kinds of

computer systems (§2.3.1): in particular, how TUIs differ from GUIs (§2.3.2); and the limitations of TUIs (§2.3.3).

2.3.1 Embodied interaction: tangible, social, situated, and hands-on interactions

Marshall & Hornecker's review of the theories of embodiment in HCI contains two main strands: *embodied interaction* and *embodied cognition* [Marshall & Hornecker, 2013].

Embodied interaction was a term first used by Dourish [2001] to refer to novel interaction experiences with computers other than desktop interfaces, as well as the implications for interface design of these systems (e.g. TUIs [Ishii & Ullmer, 1997], UbiComp interfaces [Weiser, 1991]). In these novel interactive systems, physical actions replace the point-and-click interaction typical of single-user PCs. Dourish presents *social interaction* and *tangible interaction* as related research areas within this approach, which are two aspects of key interest in this thesis.

Dourish emphasises how meaning is co-constructed through *making* (i.e. hands-on activities) within a social context and mediated by the technology: "Embodied interaction is the creation, manipulation and sharing of meaning through engaged interaction with artefacts" [Dourish, 2001, p. 126]. Here, the notion of embodiment refers to the creation of meaning while bodily interacting with these systems. It also refers to interacting (and socialising) with other people through the use of these systems.

Embodied cognition refers to theoretical perspectives that, influenced by concepts of human-computer interaction, rethink the Cartesian view of mind and body as distinct entities. Examples include studying the role of the environment related to the situated body and mind [Hutchins, 1995], or studying the role of physical experience related to the generation of abstract concepts [Lakoff & Johnson, 1999]. The former develops the notion of *distributed cognition*, in which human cognition is distributed between individuals, artefacts, and the environment, bringing to the fore the social aspects of cognition from

an individual perspective. The latter, even though it is not particularly influenced by computers, has influenced authors in computing. In particular, in the domain of music, Wilkie et al. [2010] looked into how *image schemas* and *conceptual metaphors* can inform music interaction design based on sensory-motor experiences, for example using metaphors of space, force, containment or orientation. Personal full immersion in an activity or *flow* [Csikszentmihalyi, 1975], as well as *group flow*, are terms used for understanding creative engagement. Group flow in collaborative music making has been investigated by Sawyer [2007] focusing on innovation; and also by Bryan-Kinns & Hamilton [2009], particularly *mutual engagement*.

In line with Dourish, our interest is in the kinds of social interactions arising in situated group activity mediated by TUIs. We are interested in what these social and interactional aspects can tell us from a *social dynamics* perspective, rather than focusing on the individual cognitive experiences that emerge from working in a group, which is a perspective appropriate to psychology and cognitive science. In some circumstances, problems are more effectively solved by groups than by solitary individuals (for example, think of how an academic paper is improved with the help of a number of researchers during the academic peer-review process).

Klemmer et al. [2006] discussed the implications for the design of TUIs integrating ideas from both major strands of embodiment in HCI, including theories about embodiment from psychology, sociology and philosophy. The paper focuses on how bodily movements have greater significance (i.e. whole-body interactions) in TUIs than in traditional desktop systems, as the use of the body differs when riding a bicycle compared to when writing an email. The paper presents five themes: *thinking through doing*, *performance*, *visibility*, *risk*, and *thick practice*. As an example of one of these themes, which resonate with our research, we here particularly focus on the theme of *thinking through doing*.

In *thinking through doing*, perception, cognition, and action are connected through bodily interactions. This theme integrates gestures, physical objects,

and actions: the user manipulates artefacts to accomplish tasks (*pragmatic action*) or to explore the context of the tasks (*epistemic action*). Using *external representations* for supporting thinking has been discussed by Kirsh [2010]. A sub-category of this theme is *thinking through prototyping*, where products are designed by working iteratively on a prototype, immersed in a continuous creative process based on a collection of dialogues with materials: sketching on paper or shaping clay, for instance. These activities are similar to music performance, which also generally demands bodily interactions and a thinking through doing approach, depending on the musical instrument. Our position is that thinking through doing or *tinkering* is manifested in tabletop interaction. This approach is aligned with the embodied cognition strand discussed in Klemmer et al. [2006]'s paper. However, as stated above, our focus is influenced by the embodied interaction strand, and so we approach thinking through doing as a group activity, in which there is also *group thinking through doing*. We are interested in the social implications of performing a hands-on collaborative activity.

Social interaction is a key element in understanding interaction with TUIs, in which the physicality of manipulating objects and the possibility of multi-user interaction predominates. This term is generally used in social sciences that broadly refers to relationships between individuals. Other disciplines interested in understanding group interaction, such as tangible interaction studies [Hornecker & Buur, 2006; Shaer & Hornecker, 2009], or museum studies [Hindmarsh et al., 2005; vom Lehn et al., 2001], borrow this term. For example, in museum studies, social interaction is an element of study for understanding how people make sense of exhibits, both between companions and strangers [Vom Lehn et al., 2001].

Hornecker & Buur [2006]'s tangible interaction framework, along with Shaer & Hornecker [2009]'s survey argue that tangible interaction is social and collaborative, and they highlighted the importance of supporting social interaction and collaboration when designing TUIs. As pointed out by Fernaeus et al., the shareable nature of TUIs involves "designing for collaboration, sharing and social interaction" [Fernaeus et al., 2008, p. 226], which contrasts

with more individual design perspectives for desktop interfaces. Hornecker
& Buur's framework is a general-purpose analytical tool for understanding
embodied interaction that brings physicality (e.g. materiality, body, represen-
tation) and the social aspects to the fore. With a focus on the interaction, this
framework helps us to understand the characteristics and nature of TUIs, and
to compare between systems. In this research, we expect to understand the
nature of social interaction with TTIs for music performance, which should
shed light on tangible computing studies.

The survey by Shaer & Hornecker [2009] stresses that TUIs encompass *situ-
atedness*. This term is connected to the philosophical stance of embodiment
and phenomenology, particularly Heidegger and Merleau-Ponty (cf. Shaer &
Hornecker [2009]), and the notion of being in the world, perceiving with, and
learning from, our bodily actions, within a social and cultural practice. More-
over, Shaer & Hornecker point out that because TUIs are part of the physical
world, similar to other daily objects, they are used in situated contexts, in
which the meaning related to their use can change: "Situatedness further-
more implies that the meaning of tangible interaction devices can change de-
pending on the context in which they are placed, and reversely, they can alter
the meaning of the location" [Shaer & Hornecker, 2009, p. 98].

Finally, with the advent of tangible interaction, the *practice turn* has been high-
lighted [Fernaeus et al., 2008]. In the practice turn, *action-centric* interaction
replaces *information-centric* interaction, because TUIs promote more physi-
cal and bodily actions compared to WIMP interfaces. Constructive building
blocks exemplify practice-based interaction:

Constructive building blocks: *hands-on learning*. Using constructive
building-blocks is an approach influenced by the hands-on education ideas of
Papert [1980] in the digital domain, and was first introduced by Montessori
[1912] and Fröbel [1885]. In this approach, participants create diverse struc-
tures from *modular* building blocks, and are encouraged to work as a team.
Research on TUIs and education has investigated computationally enhanced
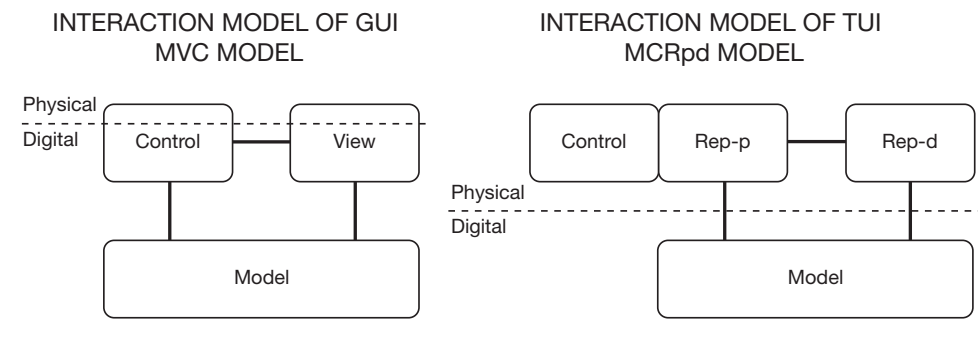
FIGURE 2.9: Left: GUI interaction model. Right: TUI interaction model, based on a figure from Ullmer & Ishii [2000]

tangible building blocks: for example, Zuckerman et al. [2005] explored digital *Montessori-inspired Manipulatives*, also known as MiMs, and endorsed their suitability for hands-on learning by doing, and for promoting group work.

2.3.2 GUIs vs TUIs

Ullmer & Ishii [2000] presented the TUIs interaction model *Model-Control-Representation physical and digital* (MCRpd). This model differs from the classical *Model-View-Controller* (MVC) used for GUIs: as the TUI model employs physical artefacts to both control and represent digital information; while the GUI model separates control from representation (see Fig. 2.9). In the TUI model, representation is separated into: 1) physical representation (*rep-p*), understood as the physical artefacts or tangible objects; and 2) digital representation (*rep-d*), which are media elements that support the visualisation (e.g. video projection).

According to Ullmer & Ishii [2000]'s model, TUIs include visual feedback. Having said that, since this model was presented, a number of TUIs have been developed, and there is now a range of TUIs that lack visual feedback: for example, tangibles for feet with haptic feedback [Schmidt et al., 2014] or the squeezable balls mentioned above with force feedback for music performance [Weinberg & Gan, 2001]. Thus, TUIs do not necessarily require visualisation elements. Feedback is helpful for a better control of a TUI system,

and other senses, apart from just visual, can also be used (e.g. auditory, haptic). We use Ullmer & Ishii [2000]'s TUI model because it is useful for understanding the systems used in the empirical work and literature review of this thesis (see terminology in §1.3). In particular, we focus on TUIs with visual feedback, in addition to the auditory feedback needed in tangible musical interfaces.

2.3.3 Limitations

A limitation of TUIs is that the physical properties of the objects that comprise them cannot be changed. Typically, visual output is overlaid along the object. Projecting a digital image on to a physical object produces an 'intangible' representation, which reduces the seamless coupling between the digital and the physical, creating a more disruptive user experience. An exception is to provide multitouch input to the visual output that is overlaid onto the object, such as in the Reactable. Ishii et al. [2012] refer to TUI systems that are based on transformable materials, such as *radical atoms* or *material user interfaces* (MUIs). For example, research into *smart material interfaces* (SMI) investigates new materials that can change their properties (e.g. colour, shape, or texture) under external stimuli (e.g. electricity, magnetism, light, pressure, or temperature) [Minuto et al., 2012]. In SMIs, the material integrates both the input and the output as part of the physical object and interaction.

Shaer & Hornecker [2009] discuss the challenge of scalability in TUIs in terms of the difficulty of working with large collections of tangibles, as well as the limitation of working with objects of fixed size, when with digital objects it is possible to easily scale and modify object properties. Furthermore, there is a lack of versatility of the TUI interface: usually TUIs are based on a single-purpose program rather than offering the flexibility of a set of different programs that may be available in desktop systems [Shaer & Hornecker, 2009]. An exception is Zuckerman et al. [2005]'s MiMs, a TUI of building blocks with a level of abstraction that permits users to build varying simulations such as probability distributions or dynamic behaviours.

Thus, the future direction of research on TUIs may concentrate on improving the coupling between the digital and the physical domains using novel materials, along with supporting better scalability and flexibility of the TUIs.

In summary, TUIs have brought not only interactive systems with physical objects that control digital data, but also novel approaches to social interaction, collaboration, and hands-on learning. The nature of these novel approaches to collaboration mediated by musical tabletops is investigated throughout this thesis. Moreover, how to best support physical-digital coupling and modularity is an aspect related to TUI design that is also researched in this thesis.

2.4 Computer-supported cooperative work (CSCW) and tabletops

The term *computer-supported cooperative work* (CSCW) refers to the understanding of how computer systems can support group activities. The term *groupware* is also used to describe software that supports group work. CSCW includes real-time collaboration systems that support synchronous vs asynchronous collaboration, as well as co-located vs remote group work. *Real-time collaboration* refers to the support of concurrency or parallel actions and interactions in interactive systems (e.g. network music, collaborative writing using individual interconnected mobile devices, or TTIs).

Weiser [1991] envisioned three sizes of UbiComp systems: "Ubiquitous computers will also come in different sizes, each suited to a particular task. My colleagues and I have built what we call tabs, pads and boards: inch-scale machines that approximate active Post-It notes, foot-scale ones that behave something like a sheet of paper (or a book or a magazine), and yard-scale displays that are the equivalent of a blackboard or bulletin board" [Weiser, 1991, p. 98].

Accordingly, there are currently three main *form factors* for multitouch and tangible interaction: *smartphones* (inch size), *tablets* (foot size), and *tabletops*
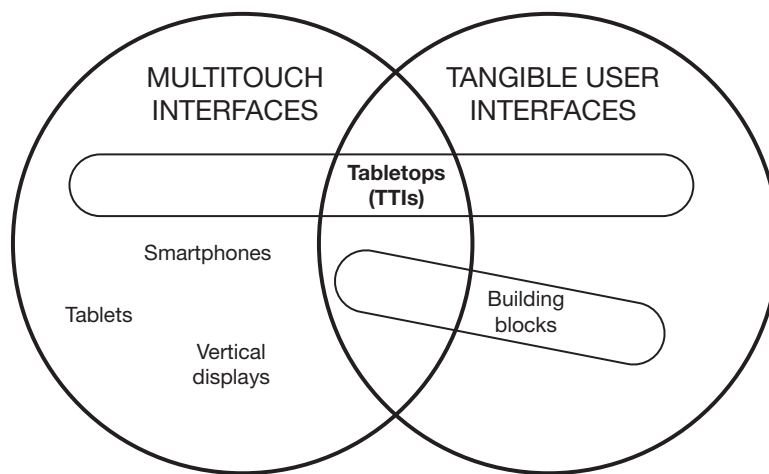
FIGURE 2.10: Diagram of multitouch and tangible user interfaces

(yard size). There are a number of applications for these devices, and some of them support collaborative work. In particular, interactive tabletops support collaboration, because the shared interface resembles the properties of traditional tables over which people work face-to-face. Besides, collaboration is computationally possible as parallel interactions are usually supported, and co-located or remote group work can be supported as well (cf. [Müller-Tomfelde, 2010]). Multi-user actions on the shareable display may vary from using hand gestures (i.e. multitouch interaction) to objects (i.e. tangible interaction), to both. Figure 2.10 shows a diagram of TUIs and multitouch interfaces classified by their form factor and type of input, which tends to be multitouch with the exception of tabletops that can include both types, and building blocks that are tangible but not necessarily multitouch.

The main characteristics of CSCW systems, particularly tabletops, include a classification considering time (synchronous vs asynchronous) and space (co-located vs remote) already discussed in CSCM systems (§2.2.1); and that these systems support real-time interaction. In addition, we find *input style* (touch vs tangible input); *work coupling* (collaboration vs cooperation); *territoriality* (individual vs group work); *real-time feedback* (visual, auditory, haptic feedback); and *group collaboration* (situated action and collaborative learning). Here, we give an overview of these characteristics:

- **Input style**: *touch vs tangible input*. Interactive tabletop systems are large surfaces that permit direct interaction, using multitouch or tangible object interaction. With multitouch interaction, gestures govern the interaction with digital content. If tangible objects are used, the objects control the interaction with digital content and tend to have visual information laid over or below them. Interactive tabletops use a range of input styles, from multitouch using fingers [Friess et al., 2012; Hornecker et al., 2008] or objects such as a stylus [Nacenta et al., 2010; Tuddenham & Robinson, 2009], to object manipulation [Pontual Falcão & Price, 2010], to both [Gallardo et al., 2008; Jordà et al., 2005; Julià & Jordà, 2009]. In the domain of music, there are considerable differences between using a pure multitouch interface vs a tangible interface. Touch input involves the finger manipulation of digital objects: for example, pressing and moving a visual representation of a digital slider from left to right changes values across a range (e.g. *skeumorphic systems*, in which the input entities are designed to look like, as much as possible, their physical counterparts). Whereas, tangible input, involves the manipulation of objects: for example, rotating a physical knob from left to right changes values across a range. Both cases use haptic information related to human tactile feedback. However, the user perceives in a different way the information of the object, because with tangible manipulation the world is perceived through tools (i.e. tangible objects), which contrasts with the abstract digital representation when using pure touch. This resonates with Gibson [1966]'s notion of *haptic perception* as an active exploration, and the different perceptual experiences between using the body to explore the world, or using a tool as an extension of the body to explore the world. When using a tool, there are haptic characteristics, such as size, temperature, texture, volume, shape, or weight, which enrich the haptic experience. Since the popularisation of touch-based devices, such as smartphones and tablets, users are familiar with both types of interaction. However, the work of this thesis focuses on tangible input, because this input style better resembles the interaction with physical musical instruments.

- **Work coupling**: *collaboration vs cooperation*. In CSCW, *work coupling* refers to the level of intensity of communication required during the performance of a task, related to the level of interaction between group members [Neale et al., 2004]. According to Neale et al., and in order from less to more demanding work coupling, there are five levels: *lightweight interactions*, *information sharing*, *coordination*, *collaboration*, and *cooperation*. Collaboration entails a common goal, yet tasks are interdependent and work is done individually. By contrast, cooperation demands the highest level of work coupling and thus the highest quality of communication. Goals and tasks are shared, and consultation with others is needed before proceeding with the work. Work *coupling styles* or collaborative coupling styles between individual and group work were discussed by Gutwin & Greenberg [2002]; Tang et al. [2006]. Supporting fluid transitions between individual and group work is a recurring theme in tabletop studies of collaboration [Tuddenham & Robinson, 2009]. In this thesis, we focus on music performance as a collaborative task, in which musicians are interdependent, and combine individual and group work.

- **Territoriality**: *individual vs group work*. Tabletop research reveals that there are territorial divisions of the surface space into *personal*, *group*, and *storage* territories [Scott & Carpendale, 2010; Scott et al., 2004]. This relates to work coupling from the perspective of how the space is distributed. Design recommendations for supporting collaboration in tabletops include supporting transitions between personal and group work [Scott et al., 2003]. In the tabletop literature it is reported that users of multi-user interactive tabletops have difficulties in differentiating between individual and group actions [Scott et al., 2003], even though multitouch interaction seems to facilitate more awareness of individual and group work than other input techniques, such as mouse input [Hornecker et al., 2008]. During the performance of collaborative tasks on tabletop interfaces, it becomes difficult to identify individual work from group work. The question of collaboration on tabletops

is a live research issue; and there is a need to understand how interfaces can best support group collaboration, while at the same time facilitating individual work [Nacenta et al., 2010; Tuddenham & Robinson, 2009], including how to best support interferences considered as beneficial for group collaboration [Hornecker et al., 2008; Pontual Falcão & Price, 2010]. With a few exceptions [Hancock et al., 2005], solutions to this issue have tended to focus on visual feedback, with no mention of other modes of feedback, such as auditory or haptic.

- **Real-time feedback**: *visual, auditory, haptic feedback*. In interactive tabletop environments, users' actions are generally supported with feedback. In multitouch interfaces, there is a greater reliance on visual feeedback (e.g. Jordà [2008]). This approach has some limitations, such as finger occluding issues [Vogel & Casiez, 2012], or the touch feedback ambiguity problem [Wigdor et al., 2009]. In TTIs, visual representations have been generally used as a mechanism to provide feedback about people's actions. The role of the tangible objects is to both control and represent digital data [Ishii & Ullmer, 1997], and so features of the objects such as position, orientation, identity, or the relationships between objects are visually represented via projection or display (i.e. visual output is overlaid). In TTIs, we find explorations with haptic feedback, with some examples in the music domain [Fiebrink et al., 2009; Kaltenbrunner, Geiger, & Jordà, 2004]. However, using the sense of touch is usually poorer than other senses, and it has some clear limitations, including the difficulty of dealing with complex pieces of information compared to other senses [Marquardt et al., 2009], although auditory feedback has barely been explored. Our Study 3 (see Chapter 6) aims to address this and explore this gap in the literature.

- **Group collaboration**: *situated action and collaborative learning*. Suchman [1987] introduced *situated action* as a term to describe the way

users act in a particular context (in Suchman's case when using a photo-copier), where shared meanings are constructed according to the situation, which depends on the people involved and the particular technology used. A similar term is *situated learning*, where knowledge is shared and co-constructed within a context and in a *community of practice* (CoP), understood as a group that shares an activity [Lave & Wenger, 1991]. Roschelle [1992], for example, introduced the term *convergent conceptual change* to describe a collaborative learning and hands-on approach using interactive tabletops. *Collaborative learning* has been a frequently studied aspect of CSCW systems [Dillenbourg, 1999]. These concepts are relevant here because the present research focuses on understanding particular social and technological contexts, based on a hands-on collaboration on a tangible interface. In particular, these concepts allows us to understand the action as situated, and the practice as a collaborative activity in which the knowledge is transferred by doing in a group.

An important design aspect in TTIs is supporting the workspace awareness of individual and group work, which we discuss in the next section (§2.4.1). We then review the literature on auditory feedback in tabletop studies (§2.4.2).

2.4.1 Supporting workspace awareness of individual and group work

As outlined at the beginning of this section (§2.4), studies have revealed the importance of supporting both individual and group actions in collaborative interactive systems, and the difficulties experienced by group members in differentiating between them [Dourish & Bellotti, 1992; Gaver, 1991; Gutwin & Greenberg, 2002].

In CSCW and HCI, *awareness* is a widely used term. An early definition of awareness was provided by Dourish & Bellotti: "an understanding of the activities of others, which provides a context for your own activity" [Dourish & Bellotti, 1992, p. 1]. For successful collaborative applications, Dourish &

Bellotti suggested that mechanisms should be provided that could address 1) group coordination, 2) data sharing, and 3) the provision of information about group and individual actions. More recent definitions of awareness include the notion of real time. For example, Gutwin & Greenberg provide a framework for the usability design of CSCW systems, in which awareness is narrowed to *workspace awareness*, defined as "the up-to-the-moment understanding of another person's interaction with a shared workspace" [Gutwin & Greenberg, 2002, p. 412]. Yuill & Rogers define the *awareness of others* as "the degree to which awareness of users' ongoing actions and intentions is present or made visible moment-to-moment" [Yuill & Rogers, 2012, p. 1:4], particularly emphasising the intentionality of individuals. All of these three definitions foresee the necessity of understanding the activities of others, in real time.

Supporting workspace awareness (WA) in musical tabletops is also a challenge, because it becomes difficult to simultaneously distinguish between the individual's voice and other participants' voices emanating from the shared workspace [Blaine & Perkis, 2000; Hancock et al., 2005]. The nature of workspace awareness in musical tabletops and how to best support WA is an open issue that is investigated in this thesis throughout the empirical studies.

2.4.2 Supporting workspace awareness of individual and group work with auditory feedback

Early work on using auditory feedback in CSCW systems is exemplified by Gaver [1991], who investigated using *audio icons*, associated with everyday sounds, for supporting collaboration in remote systems within an office, and who highlighted the importance of supporting smooth transitions from individual work to collaborative work and back again, including serendipitous communication. The use of sound is justified as an unobtrusive medium, which can deliver sophisticated information. This work particularly inspired

our third study (see Chapter 6) in using everyday sounds. The main differences are that we work in co-located settings, and within a music creation context: in musical tabletops, there is music as auditory feedback, and so our solution for strengthening auditory feedback is to add spatialisation. Transitions between individual and group work is relevant here, and is related to collaborative coupling, discussed above (§2.4).

More recent studies of auditory feedback in CSCW systems include those of Gutwin & Greenberg [2002], who observed the sounds produced when manipulating artefacts as a mechanism to determine individual and group awareness; and Hancock et al. [2005], who explored the use of auditory feedback in co-located interactive tabletops, and found that supporting both individual and group awareness is a difficult endeavour. Hancock et al.'s approach was to associate sounds with individual users, but at the cost of reducing the awareness of group work.

This section overviewed the most relevant terms from the CSCW literature that resonate with this thesis:

- The difference between multitouch and tangible interaction, and how the physical characteristics of the latter resembles playing physical musical instruments.

- The levels of work coupling, and the importance, yet difficulties, of supporting transitions between individual and group work, particularly relevant in collaboration as there exists a common goal and interdependent tasks.

- The territorial areas that appear in tabletop collaboration, connected to the issues of how to best distribute the space to facilitate both individual and group work.

- The role of real-time feedback in CSCW systems for understanding simultaneous actions and how auditory feedback has been little explored compared to visual or haptic feedback.

- The relevance of concepts such as situated action and collaborative learning for understanding group collaboration as a hands-on learning activity within a particular context, which can be useful in musical tabletops.

Then, we showed the relevance of workspace awareness for individual and group work, and identified existing open issues for supporting simultaneously both. This thesis examines the WA issue, and provides a suitable approach based on auditory feedback.

2.5 Musical tabletop interaction

We here review tabletop tangible music focusing on co-located systems designed for small groups. *Tabletop tangible music* is understood here as the activity of performing music on tabletop systems in real time, by manipulating physical or virtual objects using direct input, including both tangible and multitouch input. This definition excludes tabletop systems that use the auditory channel for any other end than producing sound or music, such as augmenting touch interaction through acoustic sensing for manipulating collections of images [Lopes et al., 2011].

There are a number of TUIs for music performance. Most of them include both tangible and multitouch interaction (*hybrid systems*). Next, we present a representative selection of TTIs for music performance including tangible systems, multitouch systems, and hybrid systems. Then, we detail the Reactable, a TTI for music performance that implements the unit generator concept of MUSIC-N.

2.5.1 Musical tabletop systems

A range of musical tabletops have been developed. We here overview a number of them:
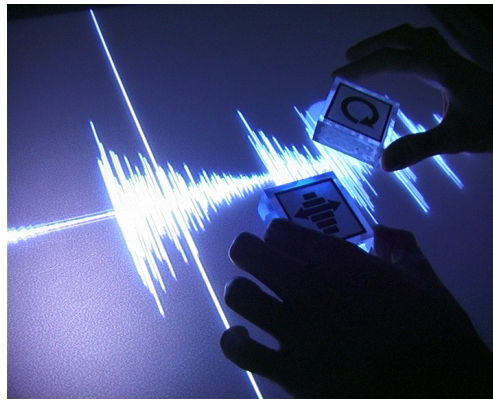
FIGURE 2.11: Close view of the waveTable

- *Composition on the Table (1999)* is an audio-visual installation by Iwai [1999]. It consists of four tables that can be touched — two circular, one square, and one rectangular — in contrast with the most common approach of a single musical tabletop. Top projectors project images on to the tables. Each table allows only a single form of interaction: push, twist, turn, or slide, with images of switches, dials, and turntables. By touching the projected images, both sounds and images are triggered.

- *Jam-O-Drum (2000)* is a drum table for musical improvisation by both beginners and experts, designed for up to twelve players [Blaine & Fels, 2003; Blaine & Perkis, 2000]. It has six drum pads distributed on a circular table, which also has an integrated projection surface. Each pad has a piezoelectric sensor that identifies audible strikes. There are two computers, one responsible for the audio module, and another responsible for the visual module. Initially, Max was used for the audio module, and Director for the visual module. The whole program was later rewritten in C. Each pad is connected to the speaker next to it. There is also a sampler to play the sounds, and a drum trigger module connected to the drum pads. Different audio delivery set-ups were tried including a stereo global mix in two speakers, distributed sound sources in six pairs of headphones, and six individual speakers combined with surround sound speakers.

- *Audiopad (2002)* is a pioneering work by Patten et al. [2002], a hybrid TTI with audiovisual feedback. Visual feedback is projected from above, and objects are tracked electromagnetically using radio-frequency (RF) tags. The position and orientation of each tangible object is tracked by the system. Each tangible object can have a set of samples. It is possible to select the sample, change the volume, and apply effects. Tangibles objects are used individually.

- *Audio d-touch (2003)* is a seminal TTI for music performance developed by Costanza et al. [2003]. It works with blocks on a piece of paper and a top webcam. Three musical applications are demonstrated: sequencing with the physical sequencer, drum editing with the tangible drum machine, and collaborative composition with the augmented musical stave. This system explores concepts such as the manipulation of objects in real time for music performance, in which the physical objects are only used for control, without real-time visual feedback. In the physical sequencer, tangible objects represent sound samples that can be positioned on a cyclic timeline. Sounds can be recorded and several audio effects can be applied in real time. The rotation of a block determines the playback speed of the sound sample; its vertical position determines the volume; and its horizontal position determines the moment when the sound is triggered within the loop.

- *Reactable (2005)* has become one of the most popular multitouch and tangible tabletop instruments [Jordà, 2008; Jordà et al., 2005]. The table is circular shaped, and visual feedback is projected from below. See §2.5.2 for further details.

- *Scrapple (2005)* by Golan Levin [2006] is a sequencer that allows the generation of a real-time *spectrographic score* (a score based on the frequency content of a sound over time) using tangible objects laid on a long rectangular table, with projected visuals. Dark rubber and felt objects are used. The system runs on a PC, with custom software written in C++. For the spectrographic synthesis, an additive synthesizer is used,

and OpenGL is used for the visuals. There is a top IR camera that captures the table in real time, along with a top video projector that projects a grid, a current time indicator, and glowing haloes around each object. Objects are sequentially played in time, from left to right, whilst pitch is defined by the position of the objects, from bottom to top with a variation of eight octaves. Volume is controlled by the darkness of the object (the darker the louder), and the size of the object corresponds to the wider range of frequencies played. Finally, the tempo of the looping sequence can be varied via an external knob.

- *AudioCube (2005)* is an installation that allows visitors to create a soundscape, developed by Daniel Dobler and Philipp Stampfl for the company Audite, which specialises in the development of interactive exhibits.[7] The installation contains four cubes, each one representing a category of sounds: drums, base, leads, and string. Each side of a cube represents a different sound in its category. The cubes can be positioned on a table with a glass plate, and the position of the cube on the table surface is mapped to the position of the sound in the room.

- *Music Table (2006)* allows users to compose musical patterns by placing cards on a table [Berry et al., 2006]. The movements and position of the cards are tracked through a top camera and displayed on a separate screen. The object recognition and visual rendering is done using the Augmented Reality Toolkit, and sounds are managed by a MIDI sequencer written in PD (cf. [Puckette, 2002]). The position of a note card from left to right determines the position of the sound in time within a loop, and its position from bottom to top determines the pitch. In addition to the note cards, cards with other functionalities are provided, such as a copy card, a phrase card, a phrase-edit card, or an instrument card. This allows copying an existing musical phrase, storing a musical phrase, editing a stored musical phrase, or modifying the instrument sound, respectively.

---

[7]`http://modin.yuri.at/tangibles/?list=1` (accessed 30 September 2014).

- *Xenakis (2007)* is a tangible sequencer, inspired by the work of the composer Iannis Xenakis, that uses a stochastic model to automatically compose music that can be modified by the user [Bischof et al., 2008]. The algorithm generates a series of MIDI notes from the probability model. Horde3D, an open source 3D rendering engine, is used for the visuals, and MatraX, an image processing software, for tracking the markers. There is an IR camera and projector below the table and three types of tangibles: instruments, pitches and rhythms. Tangible stones on the table corresponding to an instrument can influence the rhythmic and pitch tangibles that are next to it, following the rules of proximity.

- *ISS Cube (2007)* is a surround sound mixer in which objects associated with a single sound can be positioned in space by moving the objects, creating a soundscape.[8] The audio is delivered with a surround sound system of four speakers.

- *waveTable (2008)* is a real-time waveform editor that combines multi-touch and tangible interaction [Roma & Xambó, 2008] (see Fig. 2.11). The direct manipulation of a sound sample is possible using tangible objects that represent items of a desktop toolkit, including editing tools, such as an eraser, a pencil, copy, paste, or gain. Also included are effects tools, such as delay, a resonant low pass filter, tremolo, reverb, or bit crush. The system uses the computer vision software reacTIVision [Kaltenbrunner & Bencina, 2007]. The waveTable software is written in SuperCollider. There is a bottom IR camera and a bottom projector for projecting the visuals.

- *Stereotronic Multi-Synth Orchestra (2009)* is a multitouch rotary sequencer.[9] It works on a Microsoft Surface tabletop or DIY tabletops. Performers can add notes to concentric rings, and their position determines when are they played in the sequence. The application has been built on the XNA framework, and DirectSound is used for the synthesis.

Table 2.2 summarises the different musical tabletops discussed in this section.

---

[8] www.youtube.com/watch?v=j3UPCUpoIAk (accessed 30 September 2014).
[9] https://vimeo.com/6859653 (accessed 30 September 2014).

TABLE 2.2: Musical tabletops

Year	System	Contribution	Target	Roles	Technology	# Performers	Territoriality	Audio delivery	Input	Output	Reference
1999	Composition on the Table	Summative	Beginners	4 table-based roles	Top projectors	From individual to large group	Flexible	Speakers	Touch	Audiovisual	Iwai [1999]
2000	Jam-O-Drum	Summative	Beginners, Experts	6 drum-based roles (flexible)	Piezos, Max/Director, C	1–12	Individual	Headphones vs speakers (individual, surround sound)	Touch	Audiovisual	Blaine & Perkis [2000]
2002	Audiopad	Summative	Beginners, Experts	Flexible	RF system, top projector	From individual to small group	Flexible	N/D	Hybrid	Audiovisual	Patten et al. [2002]
2003	Audio d-touch: Physical sequencer	Summative, Multiplicative	Beginners	Flexible	Top webcam	From individual to small group	Flexible	Computer speakers	Tangible	Audio	Costanza et al. [2003]
2005	Reactable	Summative, Multiplicative	Beginners, Experts	Flexible	DI system	1–4 (larger groups in public venues or with remote settings)	Flexible	Up to 4 speakers, headphones	Hybrid	Audiovisual	Jordà et al. [2005]
2005	Scrapple	Summative	Beginners, Experts	Flexible	PC, C++ custom software, OpenGL, top projector, top IR camera	From individual to large group	Flexible	N/D	Tangible	Audiovisual	Levin [2006]
2005	AudioCube	Summative	Beginners	4 roles (by cube)	N/D	From individual to small group	Flexible	Speakers (surround sound)	Tangible	Audio	None
2006	Music Table	Summative, Multiplicative	Beginners	Flexible	PD, Augmented Reality Toolkit, top camera, screen	From individual to small group	Flexible	N/D	Tangible	Audiovisual	Berry et al. [2006]
2007	Xenakis	Summative, Multiplicative	Beginners, Experts	3 roles (by type of tangibles)	MVC program, bottom projector, bottom IR camera, MatraX	Small group (1–3)	Flexible	N/D	Tangible	Audiovisual	Bischof et al. [2008]
2007	ISS Cube	Summative	Beginners	Flexible	N/D	From individual to small group	Flexible	Surround sound system (4 speakers)	Tangible	Audiovisual	None
2008	waveTable	Multiplicative	Beginners, Experts	Flexible	SuperCollider, bottom IR camera, bottom projector, reacTIVision	From individual to small group	Flexible	N/D	Hybrid	Audiovisual	Roma & Xambó [2008]
2009	Stereotronic Multi-Synth Orchestra	Summative, Multiplicative	Beginners, Experts	Flexible	Microsoft Surface, XNA framework, DirectSound	From individual to small group	Flexible	Stereo	Multitouch	Audiovisual	None

* The number of performers is either defined in or derived from the description of the musical tabletop.

FIGURE 2.12: Close view of the Reactable

This section surveyed a selection of musical tabletops using criteria borrowed from network music, TUI, and CSCW: contribution, target, roles, technology, number of performers, territoriality, audio delivery, input mode, and output mode. Beyond technological differences among them, musical tabletops are designed for co-located small groups, with a shared interface that usually lacks private spaces. The use of speakers for audio delivery is a common option yet little explored, an aspect of interest in this thesis. The flexibility in roles and territoriality offered by these systems is another aspect examined in this thesis. Next, we present in detail the Reactable, a TTI based on constructive building blocks.

2.5.2 The Reactable: from MUSIC-N to dynamic patching

The Reactable [Jordà et al., 2005] is a commercial tabletop and tangible user interface (TUI) with multitouch input for electronic music performance (see Fig. 2.12). The system is a collaborative instrument that implements a virtual modular synthesizer, and can be controlled using tangible objects, in this case constructive building blocks that can be interconnected. It was invented by a team in the Music Technology Group in Barcelona, presented at the ICMC Conference in 2005, and commercialised in 2009. As explained in the methodology chapter (Chapter 3), two of our three studies in this thesis explored the use of the Reactable.

Although influenced by analogue and digital modular synthesizers such as Robert Moog's or Donald Buchla's voltage-controlled synthesizers [Moog 1965], the sound synthesis and control method implemented in the Reactable

interface can also be seen as a tangible representation of the unit generator paradigm invented by Max Mathews [Roads, 1996, pp. 89-90].

Matthews created a series of sound synthesis programs in the 1950s–60s, i.e. the so-called MUSIC-N, in which he developed the concept of *unit generators* (UG). These are snippets of code with different functions (e.g. generators or modifiers of sounds). UGs work as building blocks: they can be interconnected (whereby, the output of a UG can be connected to the input of another UG) to create instruments (*patches*), using a range of structures from basic to complex, and produce sound [Roads, 1996, pp. 787-788]. The interconnections are possible because the UGs have inputs and outputs, and as UGs can emit either audio or control signals.

MUSIC-N, and particularly the UG concept, influenced much of the next generation of computer music software: for example, SuperCollider, and PD or Max/MSP. One of the main differences between MUSIC-N and more contemporary programs is that the latter handle sound synthesis in real time. SuperCollider, which is based on the object-oriented programming (OOP) approach, implements the unit generator paradigm using *UGens*, which are objects (or stand-alone snippets of code) with different functions for producing or modifying audio signals that can be interconnected for creating instruments. Similarly, PD or Max/MSP implements the unit generator paradigm using 'objects', but in this case the interconnections are done by wiring visual objects. The Reactable is a step beyond newer instances of the UG paradigm, such as the visual programming languages (e.g. PD or Max/MSP), because it simulates the physical wiring of the older analogue modular synthesizer, with unit generators taking the physical tangible form (as opposed to virtual 2D form) of objects that can be interconnected.

Unlike visual programming musical languages for real-time audio synthesis (e.g. PD or Max/MSP cf. [Puckette, 2002]), the Reactable supports real time *dynamic patching* [Kaltenbrunner, Geiger, & Jordà, 2004], permitting users to edit and play at the same time, instead of having two separate modes. An analogue modular synthesizer traditionally has two modes of usage: 1) *sound design* stage, and 2) *performance* stage. In the sound design stage, it is more

common to wire and unwire connections, and serendipitous actions can benefit the creative process. In the performance stage, it is more common to work with the control of a patch with fixed connections. The Reactable merges these two stages digitally via dynamic patching. This mechanism promotes creative actions by facilitating automatic connections between objects in real time. It thus brings the serendipity of the sound design stage into the performance stage. The Reactable's complexity, combined with a reasonable degree of variability and unpredictability, mean that complex behaviours can emerge [Jordà, 2004], allowing for fortuitous events that can benefit musical improvisation.

2.6 Discussion

This chapter surveyed relevant literature on musical tabletops and collaboration. Musical tabletops are a novel instrument for musical collaboration, yet little is known about the challenges and opportunities that these systems offer to music performance among beginners and experts. To summarise the main open questions, and how this thesis will address them:

- *Musical tabletops and CSCM systems*. The implications of the flexibility of roles offered by musical tabletops are investigated between beginners (Chapter 5), experts (Chapter 4), or heterogeneous groups (Chapter 6). The implications of using auditory feedback are also addressed, in particular the use of speakers and headphones (Chapter 5), and the use of speakers with localised sound sources (Chapter 6). The implications of co-located collaboration and different levels of synchronicity are researched in Chapter 5 (collaboration over time) and in Chapter 6 (synchronous vs asynchronous).

- *Musical tabletops and CSCW systems*. The implications of lack of territorial constraints within musical tabletops are addressed in Chapter 4. Little is known about between-group interaction in public settings, which is investigated in Chapter 5. How WA is manifested in musical

tabletops is considered with respect to experts in Chapter 4, beginners in Chapter 5, and heterogeneous groups in Chapter 6. How best to support transitions between individual and group work within a tabletop musical activity is investigated in Chapters 4 and 6.

- *Musical tabletops and TUI systems*. The implications of physically interacting with a TUI as a group are addressed in Chapters 4, 5, and 6. The implications of seamless coupling and the relation between the physical and digital domains are addressed in Chapters 4 and 5. The implications of musical tabletops as a hands-on collaborative environment are explored in Chapters 4, 5, and 6.

2.7 Summary

The aim of this chapter was to review the relevant literature on musical tabletops, focusing on design aspects for collaboration. We first mapped the key terms and characteristics of CSCM systems. In particular, we surveyed network music, because of its long tradition of co-located collaborative practices. We also reviewed TUIs for music performance. Then, we presented relevant concepts from tangible interaction and tabletops in CSCW. Finally, we surveyed musical tabletops. Borrowing concepts from network music has been useful for understanding a range of potential collaborative musical practices, in particular those related to organisational and aesthetic decisions. Similarly, concepts from HCI, TUIs and tabletop research have been useful for understanding the key terms and issues describing collaborative activities on TUIs and TTIs.

In the next chapter, we present a rationale and overview of the methodology used in this thesis for understanding and evaluating collaboration in TTIs for music performance, using the concepts presented in the literature.

Chapter 3

Methodology

This chapter is a general overview of the methods and materials used in this thesis in order to provide a better understanding of tabletop collaboration in music performance. The chapter starts by explaining our choice to use video in data collection and analysis as the most suitable research tool for understanding social hands-on activities. The thesis research design is organised into several sections. The chapter explains the commonalities and differences between the three studies to be presented, which are based on observing group interactions with musical tabletops. Subsequently, we give details on participants, tasks, procedures, ethics, methods for data collection and analysis, and study limitations. Further details of the methodological approach undertaken for each of the tabletop studies are elaborated in Chapters 4, 5, and 6.

3.1 Introduction

As stated in Chapter 1, the purpose of this thesis is to identify the challenges and opportunities that beginners and experts face in relation to the use of TTIs for music performance in different contexts. In particular, we aim to investigate:

- The nature of ***group interactions*** with these artefacts during short- and long-term use.

- Manifestations of ***workspace awareness*** in musical tabletops for music performance, between beginners and experts, and during short- and long-term use.

Video is used as a research tool for identifying the main challenges and opportunities users face when of using these artefacts. Video is a suitable research tool to overcome the *say/do problem* [Jordan, 1996] investigated in social science disciplines such as anthropology of analysing differences between what people say and what they actually do:

> One situation for which video provides optimal data is when we are interested in what 'really' happened rather than in particular accounts of what happened, including people's recollections and opinions. [Jordan, 1996, p.35].

Accordingly, the analysis of video material, in contrast with other qualitative methods such as field notes or interviews, provides a more detailed account of what happened, compared to participants' reports on what happened. We use video to capture and examine group interaction phenomena (what people do) in a musical context where there is potentially little verbal communication.

Video has its own limitations in terms of using chosen points of view for capturing interaction, or reflecting a past event, in contrast with in-situ observation. However, video can capture in-depth detail of collaborative interaction; it can be replayed as often as needed; and it can be discussed in co-located or distributed teams, which helps the researcher to reach a consensus of opinion among collaborators. These features provide benefits for understanding collaborative and social interaction and make video the most suitable tool for this research. As reflected in §3.5, to our knowledge, examining group interaction using video is new to research on musical tabletops.

In the following section, we distill our methodological approach and compare it with other approaches (§3.2) to give an overview of the methodological points in common between the three studies in terms of: participants and settings (§3.4); activity (§3.5); apparatus (§3.6); type of data collected (§3.7); procedures for data collection (§3.8); ethics involved (§3.9); and methods used for data analysis (§3.10).

Further detail on the methodological particularities of Studies 1, 2, and 3 is provided in the relevant chapter on each study.

3.2 Our methodological approach: understanding social hands-on activities using video as a research tool

Video has been widely used in social sciences as a research tool for understanding social settings, including group interactions with technology [Heath et al., 2010]. For example it has been used for studying operating theatres [Korkiakangas et al., 2014], peer programming [Plonka et al., 2011], and serious games design [Kreitmayer et al., 2012]. In tabletop studies, the use of video is a common method to capture tabletop interaction (e.g. among others, Hornecker et al. [2008]; Pontual Falcão & Price [2010]; Scott & Carpendale [2010]; Tuddenham & Robinson [2009]), resulting in video data that can be examined in depth later. The possibility of capturing human-computer interactions from different scales (e.g. hands scale, body scale, room scale) and synchronising different video sources, makes it a good fit for the aims and objectives of this thesis, that is understanding group-computer interaction. Hence, video as a research tool is the principal method used.

Furthermore, we used video because it is suitable for recording verbal and non-verbal communication (e.g. body posture or gaze) in complex interactions, such as musical interaction. In the studies presented, the role of the researcher and the position of the camera were designed to be unobtrusive so

as to reduce interference with the setting and the data obtained. The use of video has limitations that were addressed in a number of ways:

- The influence of the camera on the behaviour of participants: as it is common to video-record performances in a musical context, we argue this influence is minimal for research.

- Data analysis can be highly time-consuming: to overcome this issue the analysis used specific research questions to focus in on the video data.

- Analysis by one researcher can raise issues in relation to interpreter reliability: to reduce interpreter bias the video and analysis results were shared and discussed with the supervisors.

Alternative social science research methods based on naturalistic observation, which are non-interventionist, were rejected because their approaches did not enable the generation of the data necessary to answer the study's research questions on group interactions. Ethnography, for example, is a qualitative research method based on naturalistic observation. It consists of taking written notes about a community in everyday situations, over long periods of time [Lofland et al., 2006]. This method is commonly used in software engineering [Easterbrook et al., 2007]. However, this approach was not appropriate for the purpose of these studies: using ethnography was unable to capture the complex and fast-paced group interactions that happen in short periods of time with a novel community. Furthermore, group interactions, which are the unit of analysis of this thesis, are characterised by a myriad of actions happening at the same time. Taking field notes of these actions is a difficult and inaccurate endeavour. In contrast, video recordings record and allow later replay of the same action at different speeds (e.g. slow motion or fast forward).

Fencott & Bryan-Kinns [2010, 2012] researched co-located group interaction using interconnected desktop-based computers. They used a mixed methods approach of qualitative and quantitative by using video and log files for capturing group interaction, and post-questionnaires and group discussions

for capturing user opinions. For data analysis, the authors combine statistical analysis and qualitative analysis of video transcripts. The approach used in this thesis, by contrast, focuses on capturing and analysing the activity of group interaction. The character of interaction with TTIs can be a richer and more complex stream of data if compared to mouse-and-click group interaction. For example, in TTIs the interaction of the body in relation to the space plays a greater key role if compared with the body-space interaction when sitting in front of a monitor. With the aim of understanding bodily interactions with TTIs we focused on analysing what people do, capturing as much detail as possible using multiple video cameras, and, where possible and appropriate, system log files of these interactions were also collected. This allows the capture of both verbal (e.g. dialogues) and non-verbal (e.g. hand gestures, glances, bodily motion, sounds, etc) interactions happening at the same time. Although video started to be used in musical tabletop studies [Franceschini et al., 2014; Klügel, Hagerer, & Groh, 2014], to our knowledge, this is the first study on musical tabletops that uses video extensively.

3.3 Research design: commonalities and differences among the three studies

Three consecutive studies were conducted with groups of users. Table 3.1 summarises the commonalities and differences among these studies. We observed group interaction, working with experts (Study 1), beginners (Study 2), and both (Study 3). The first two studies were exploratory using the Reactable, in two representative contexts: a lab resembling a rehearsal room with expert musicians, and a science centre with visitors. We then conducted an experimental study with heterogeneous groups in an informal lab setting in order to keep a balance between an open task and a controlled experiment. In this last study, we focused on particular aspects of group coordination using a DIY tabletop system. In the two lab studies we asked participants to improvise music; in the field study we investigated natural interaction.

TABLE 3.1: Research design across the three studies in this thesis

	Study 1	Study 2	Study 3
Aim	Exploratory	Exploratory	Experimental
Apparatus	Reactable	Reactable	Woodentable + SoundXY4
Setting	Lab	Science centre	Lab
Group profile	Experts	Visitors/Beginners	Heterogeneous
Data collection	Video	Video	Video
			Log files
Data analysis	Interaction analysis	Thematic analysis	Thematic analysis
		Statistical analysis	Statistical analysis
Time	Multi-session (4 days)	One-off session	One-off session

Video recordings of the group interactions were captured for all studies. Information on the tangible interface interaction was also collected via the interaction log files (text files) for Study 3 (§6.3.4). Generally, we used an unobtrusive approach, avoiding direct observation of group interactions, in order to mitigate researcher interference and the *Hawthorne effect* [Forsyth, 2010]—possible influence of the researcher on participants' behaviour—during the studies. We had a mostly qualitative approach to video analysis focusing on our research questions on group interaction and workspace awareness, using the interaction analysis framework in Study 1 (§4.3.5), and thematic analysis in Studies 2 (§5.3.4) and 3 (§6.3.5), combined with basic quantitative analysis when interaction log files were available (§6.3.5).

Our methodological approach, using video recordings and interaction log files, was inspired by tabletop studies that combine data about what is happening above the tabletop ('what people see') (e.g. video transcripts or field notes in Scott & Carpendale [2010]) and data about what is happening below the tabletop ('what the system sees') (e.g. system's activity maps in Nacenta et al. [2010]). A similar approach is found in Tuddenham & Robinson [2009]. This contrasts with studies that opt to include in their analysis users' opinions (e.g. Nacenta et al. [2010]). Our approach is suitable for understanding non-verbal interaction focusing on interactional patterns.

3.4 Participants and settings: volunteers in casual set-ups

The following section provides details of the study participants and settings in terms of group sampling design, location, number of participants, the number of groups, and duration of the activity. Figure 3.2 gives an overview:

- In Study 1 (§4.3.2), a call for voluntary participants was made in the Music Technology Group at the Universitat Pompeu Fabra in Barcelona (Spain) using the group mailing list. We gathered 12 international participants, who were allocated to one of four groups depending on their expertise with the tabletop system: resulting in three expert groups and one beginner group. The number of participants in each group varied from 2–4. They participated over four sessions.

- In Study 2 (§5.3.3), we attended the INTECH Science Centre in Winchester (UK), the only UK public centre/museum with a Reactable. We collected data during a weekend (two days); a random sample was drawn from four different time slots distributed between the morning and afternoon. A weekend was chosen because museum facilitators reported that this is the time when families and a diversity of individuals visit the museum, in contrast with scholar groups during weekdays. The visitors participated naturally with the Reactable exhibit, and some interacted with it twice or several times.

- In Study 3 (§6.3.2), a call for voluntary participants was made at the Open University in Milton Keynes (UK) using mailing lists (including the Centre for Research in Computing, the Faculty of Arts and the post-graduate students list). These mailing lists were used to attract a varied range of people from academics to students to staff, both music experts and beginners. No musical background was required. We gathered 32 international participants, who were randomly allocated in eight groups of four people each. They participated over one session.

TABLE 3.2: Samples used in the three studies

Study	Sampling design	Location	Participants	Groups
1	Non-random	Universitat Pompeu Fabra	$N = 12$	$G = 4$
2	Random	INTECH Science Centre	$N = 174$	$G = 54$
3	Random	The Open University	$N = 32$	$G = 8$

In order to maintain the balance between an open task and a controlled experiment, we adapted casual lab set-ups similar to a rehearsal room in Study 1 (§4.3.3), or to an isolated exhibit space in Study 3 (§6.3.3). In Study 2 (§5.3.2), the set-up was already configured in the science museum as a permanent exhibit.

3.5 The activity: why musical improvisation in lab studies?

Musical improvisation is an important aspect of each of the studies: in the two lab studies participants were asked to improvise together, while in the museum the observed interactions were also improvisational in character. In the next section, we discuss the limitations of lab studies, and how improvisation can balance these limitations.

In the lab vs in the field. Rogers [2012] compares controlled experiments in the lab with in the field or *in-the-wild* studies: lab-based studies are convenient for observing specific effects under similar conditions with pre-booked participants, although this approach lacks the real-world substance provided by field studies. Analysing wild data can be more difficult and time-consuming, however, because there is less control about factors and result in longer periods of interaction data than in a conventional one-off lab study.

Improvisation in the lab and the 'third wave' in HCI. In lab studies, improvisation is a hybrid approach that produces wild data in a controlled setting. This approach fits into the 'third wave' in HCI, which reveals how bodily

actions matter in post-WIMP computing systems [Bødker, 2006; Harrison et al., 2007; Jacob et al., 2008], in which the body is used differently compared to individual desktop interaction. By contrast, the other two non-mutually exclusive HCI waves were stemmed from engineering and human factors – ergonomics ('first wave'), and cognitive science ('second wave') [Harrison et al., 2007]. The 'third wave' in HCI provides a different lens for understanding alternative computing systems to window-icon-mouse-pointer (WIMP) systems; such as UbiComp systems. Harrison et al. [2007] use the term *situated perspectives*, to refer to a sum of perspectives that investigates HCI interaction as a situated context, which connects to qualitative disciplines such as ethnography or practice-based research. With this change of perspective in HCI, user experience has increased in relevance. We find a range of studies that embrace more open-ended tasks. For example, some studies have made use performance and theatre techniques in HCI and interaction design, as reported by Jacob et al. [2008]. In particular, Medler et al. [2010] studies have compared the use of improvisational theatre vs role-playing to inform HCI design methods. The authors discuss the benefits of improvisation: "The lack of constraints and group cohesion allow an improv scene to produce a novel performance that other acting techniques could not achieve without explicit preplanning and practices" [Medler et al., 2010, p. 486]. This novelty can help HCI designers to discover a wider range of opportunities for design and use of HCI systems. Our approach is similarly about exploring interaction opportunities but focusing on music performance.

Musical improvisation can provide relevant information for both DMIs and HCI design. In the section below we describe the potential of musical improvisation in two areas of research—HCI and DMI (§3.5.1). This is followed by a review of methods for evaluating collaborative CSCM systems based on musical improvisation (§3.5.2).

3.5.1 Musical improvisation: complex interaction in DMI studies

Musical improvisation is a well-known practice in music performance, which is based on spontaneous communication between musicians. The decision to focus on musical improvisation within the three studies contrasts with the more traditional approach of working with guided and, usually, more closed tasks. Musical improvisation is chosen, however, as it is a versatile approach that was observed in the lab and in the field; one which can inform both HCI and DMI design; and one that is well suited to interaction with tabletop systems and with DMIs; and which attracts both beginners and experts more easily.

Informing HCI design about complex interactions. Studying the *high ceiling* or expertise and complex interactions (e.g. see Chapter 4) contrasts with asking participants to perform specific and constrained tasks. The latter can be a reductionist approach and can miss the bigger picture of interactions. Observing group interaction during musical improvisation can inform HCI design. Buxton's quote (see §1.1.2) in the article "Artists and the art of the Luthier" [Buxton, 1997] reflects the complexity of supporting the 'artist spec' in HCI design, in which musical improvisation is included.

Jordà [2005a], one of the inventors of the Reactable who considers himself as a *digital luthier*, refers to the notion of *complex interaction* as related to skilled, expressive, and explorative interaction. In particular, Jordà refers to tangible and tabletop systems for music performance as interfaces that promote complex interaction via explorative and expressive activities:

> "This is why performing arts (. . .) and music performance in partic-
> ular constitute excellent realms for deeply exploring and fully exploiting
> the potential of this type of interaction." [Jordà, 2008, p. 272]

In lab studies, musical improvisation is well suited to exploratory studies, as participants have the freedom to behave as they would do in a 'real' setting. Musical improvisation has been used in lab experiments with novel technology as a suitable activity for obtaining less constrained data, in terms of being

richer and more exploratory, as compared to the data generated by specific task-based studies. Using musical improvisation in lab settings produces situations that are closer to the spontaneity of the real world and thus allows the investigation of new DMIs and potential uses. For example, Swift [2013] used musical improvisation to understand the experiential aspect of "being 'in the groove'" of groups of musicians using a collaborative music system based on mobile technology. In this research, we observed musical improvisation in the lab (e.g. Studies 1 (§4.3.4) and 3 (§6.3.4)), and the rich stream of data this provided informed a set of design considerations.

Real environments can also inform HCI and DMI design. For example, Stowell & McLean [2013] argued that a rich open task such as live coding can shed light on interface design in terms of supporting complex gestures. For example, we observed tabletop musical improvisation in a public setting (e.g. see Chapter 5), and this informed a set of proposed design considerations.

Tabletop musical improvisation. Promoting the practice of musical improvisation in the context of tabletops aligns with the notion of complex interaction. For example, the Reactable was designed for musical improvisation, in the words of Jordà "the more 'interactive' an instrument is, the less sense it makes writing completely deterministic pieces for it, thus the best suited for improvisation it should be" [Jordà, 2005a, p. 203]. Musical improvisation on tabletops, given its characteristics of spontaneity and unpredictability, is a suitable phenomena for understanding the nature of complex interaction within a situated context.

A range of approaches to musical improvisation with DMIs. We find a number of DMIs, targeting beginners, experts, or both. DMIs that are designed for experts tend to: 1) expand existing concepts of music theory such as Harmony Space [Bouwer et al., 2013], or Hex Player [Milne et al., 2011]; 2) require a certain level of computer music knowledge such as live coding [Brown, 2006; Stowell & McLean, 2013], or playing the Reactable [Jordà et al., 2005]; or 3) require experts with traditional virtuosity and practice-based skill for robots to mimic and learn from them [Weinberg & Driscoll, 2006], or expert improvisers for agents to improvise with them [Collins, 2006; Linson, 2014]. Other

DMIs explore alternative avenues by: 1) using samples such as in Freesound Radio [Roma et al., 2009]; or 2) exploring novel and constrained interactions such as in Squeezables [Weinberg & Gan, 2001]. These two examples do not necessarily rely on traditional musical knowledge, and attract both beginners and experts starting from a similar entry level. Having said that, within this continuum from beginners to experts, some of the above expert-like DMIs also support novice users, who can interact with the system irrespective of their musical background, although background knowledge is desirable for better control (e.g. Reactable, Harmony Space).

Musical improvisation with DMIs for beginners and experts. Traditional values persist within this novel terrain of DMIs and musical improvisation: "A great deal of skill and training is required to participate in improvisational group music-making at a high level" [Swift, 2013, p. 86]. Improvisation shares a common ground with more traditional contexts such as jazz in terms of an open task activity that requires coordination using non-verbal communication; and the use of certain protocols for starting, developing, and finishing a session. However, there is a need to rethink other aspects, that can become opportunities, such as roles or techniques when using novel interfaces. For example, traditional musical skill and training are less required for improvising with novel collaborative musical interfaces, as the group can learn together in the practice of doing within an exploratory collaborative task; or separately explore techniques with the DMI at home and then join the group.

3.5.2 Evaluating CSCM systems using musical improvisation

A survey on evaluating NIMEs and DMIs borrowing methods from HCI is provided by Jordà & Mealla [2014]; Kiefer [2012]; O'Modhrain [2011]. We particularly focus here on those evaluations of CSCM systems borrowing methods from HCI centred on: 1) musical improvisation; and 2) collaborative practices.

Influenced by usability evaluation and CSCW studies, most studies of musical improvisation with novel interfaces for collaboration tend to be one-off

studies situated in labs. Fencott & Bryan-Kinns [2010] focused on public and personal spaces for users of individual computers who accessed a shared virtual representation while co-located in the same room lasting for one session; Bryan-Kinns [2013] studied the distributed use of visual shared representations; Klügel, Lindström, & Groh [2014] explored musical collaboration on a multitouch surface using genetic algorithms and a collaborative voting system for supporting evolving musical forms; and Pugliese et al. [2012] investigated situated interaction and collaboration during mobile group improvisation, focusing on video analysis of participants' comments when viewing their own videoed session. Swift [2012] carried out a longer-term lab study that addressed musicians' insights into the experience of co-located improvisation on mobile devices. This study used an ethnographic approach, video recordings, and post group interviews. Also Booth & Gurevich [2012] used an ethnographic approach, over three months, and video recording to provide thick descriptions of collaborative work practices in a laptop ensemble.

Our approach builds on this work, but is based on video analysis of participants' interactions using *interaction analysis* [Jordan & Henderson, 1995] or *thematic analysis* [Braun & Clarke, 2006], which are explained in detail in §3.10. The study of participants' interaction in unstructured musical improvisation during either a single session (Studies 2 and 3), or multiple sessions (Study 1), is a new direction in tabletop and NIME research.

3.6 Apparatus: TTIs for music performance

As explained in Chapter 1, TTIs can be seen as an opportunity for collaboration because these systems emulate the characteristics of traditional tables, that is, sharing the same space for individual and group activities. Particularly in music performance, TTIs can gather together both beginners and experts. In our three studies we worked with two tabletop systems: the Reactable, which is a commercial tabletop system that includes its own software; and the Woodentable, a DIY tabletop system. For the study involving the Woodentable, we also used the bespoke software SoundXY4 written

in SuperCollider. We divided our studies into an exploratory phase and an experimental phase, which helped us to understand tabletop systems from different perspectives, and was useful for strengthening the bigger picture in relation to our research questions.

3.6.1 Tabletop tangible technologies for object tracking

Most of the tabletop systems have a configuration of:

1. A suitable surface for either multitouch or object tracking.

2. A suitable illumination for illuminating either objects or fingers, so a computer vision software can identify them.

3. A camera for capturing the object or fingertip input data.

4. Usually, a projector for displaying output display data.

Where projections are used, there is a tension between using a surface dark enough to display the projected images, and illuminated enough to let the computer vision software track the objects, or fingertips, via the camera. A consensus solution is to use the IR spectrum for managing the objects, and a visible light for the projected visuals. Using the IR spectrum is achieved by using IR illumination and an IR camera, in which the visible light lens is removed. This solution allows us to keep illuminated the objects or fingertips for tracking them in the IR spectrum (the sensing side), and a dark surface for projection in the visible light spectrum (the display side). Appendix A.1 describes in depth tabletop tangible technologies available, focusing on the diffused illumination (DI) technique.

In this study we use two tabletops, the commercial product Reactable, and the DIY tabletop Woodentable. Both implement the rear DI technique. As shown in Fig. 3.1, the IR camera is connected to a computer vision engine, in this case reacTIVision, for object tracking (see Fig. 3.2).
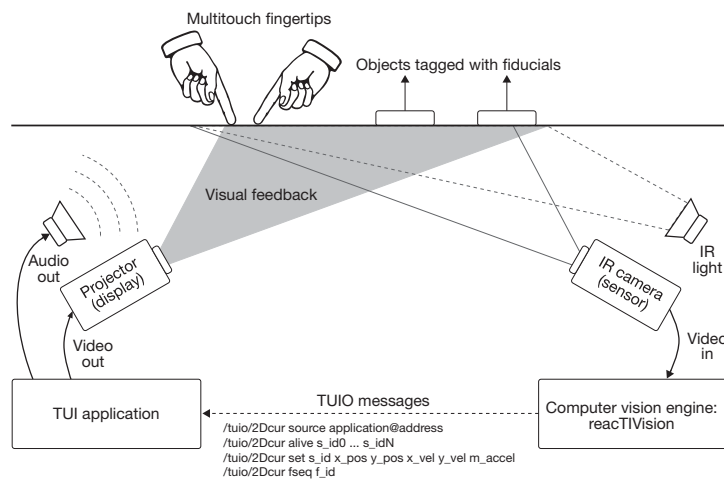
FIGURE 3.1: System design overview of a musical TTI implementing the rear DI technique
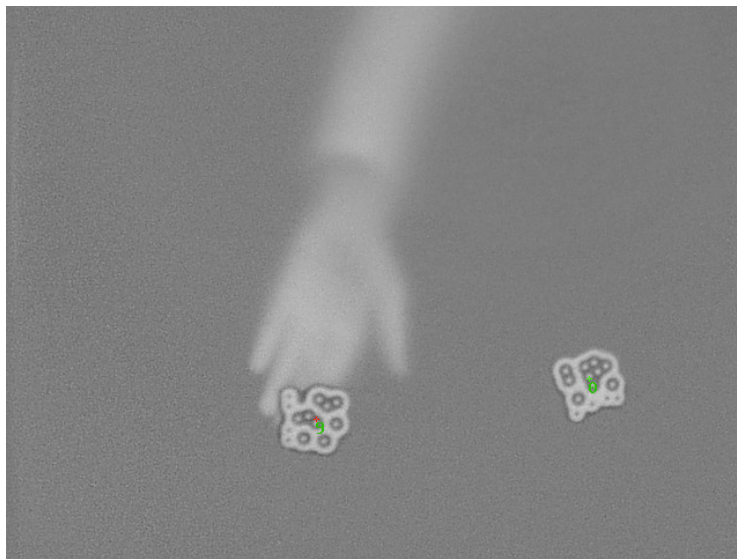
FIGURE 3.2: A screenshot of reacTIVision: identifying fiducial markers from a rear DI's tabletop

TUIO [Kaltenbrunner et al., 2005] is a protocol for TTIs. Reactivision sends TUIO messages about the tracked objects (e.g. identity, position, or rotation) to a TUI application (e.g. the Reactable). TUIO is implemented using OSC. TUIO works on platforms that support OSC such as SuperCollider, PureData or Max/MSP. The TUI application, in turn, dictates the system behaviour and output, including visual feedback from the projector.

The technologies available shape the methods chosen to address our research questions. Our choice to use the Reactable and the Woodentable systems, both of which use the open source tool reacTIVision, means that no information about who is operating with what object is provided (via the system). This information is, however, relevant in group tabletop interaction. Commercial multitouch technologies such as the DiamondTouch, are able to track the identity of the users within a group of four people positioned on each side of the tabletop surface, via information stored in system files. In our case, video data was used to obtain subject identity in relation to object identity. Using video to transcribe users' actions is more time-consuming than harvesting this information from the system files. However, the video-based manual approach used in this study is more flexible as the number of users per group and their position around the table is less fixed. This approach is therefore, we argue, more suitable for researching open tasks such as musical improvisation.

A description and rationale for the choice of the hardware and software used in our three studies is given in the next section; it is divided into two sections: 1) an exploratory stage with the Reactable (§3.6.2); and 2) an experimental stage with the Woodentable (§3.6.3).

3.6.2 Exploratory stage: the Reactable

In our Studies 1 (§4.2) and 2 (§5.2), we used a commercial tabletop system, the Reactable, in order to observe the challenges and opportunities of using these ready made commercial systems in real environments. The Reactable used in the two studies of this thesis measures $100\,\mathrm{cm}$ diameter including a rim area, i.e. a non-interactive outer area of $10\,\mathrm{cm}$, in which tangible objects can be allocated (see Fig. 3.3).

During the first stage of the research, we aimed to explore the challenges and opportunities of a tabletop system in two different real situations: expert musicians in the long-term use, and visitors in the short term use.

The benefits of studying a commercial system are threefold:
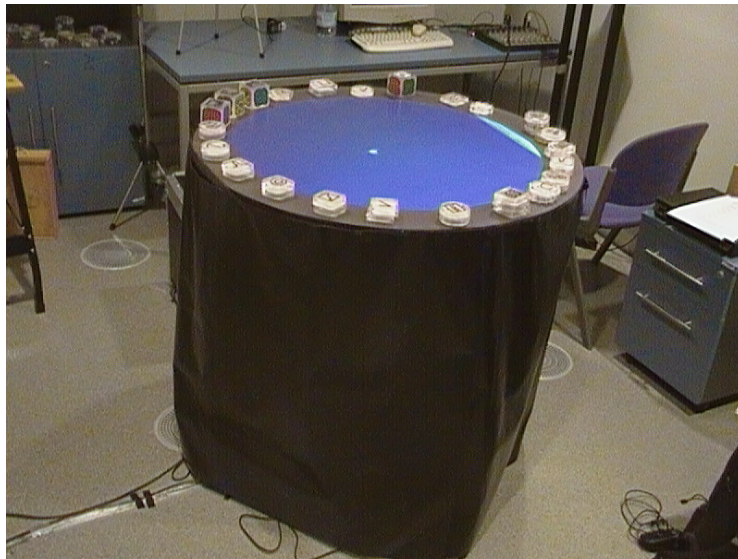
FIGURE 3.3: The Reactable

1. Observation of real uses by users who already have knowledge of the system, albeit in a lab setting.

2. Observation of interactions in a real setting including both one-off interactions, and multiple interactions.

3. Research did not need to involve a system development phase, saving valuable research time within a PhD timeframe.

The main limitations of this choice of a commercial system were:

1. Using a commercial software that is not open source prevents any modification for research purposes.

2. To gain access to a commercial system, such as Reactable, requires coordination with external institutions that house it, which incurs travel and accommodation costs, which in turn make it difficult to maintain a long-term study.

These disadvantages partly motivated the second stage of our research: building a DIY testing environment for Study 3 (§6.2).

3.6.3 Experimental stage: the Woodentable and SoundXY4

In Study 3 (§6.2), we built our own tabletop system for testing particular aspects related to workspace awareness and group tabletop interaction that emerged from previous Studies 1 and 2.

The benefits of building our own testing environment after an exploratory phase were threefold:

1. Obtaining a controlled observation of particular aspects.

2. Using open source software that allows the modification of the code when needed.

3. Independence of external bodies when conducting the research.

The main limitation was the difficulty of building a system with the same level of complexity and accuracy as a commercial system, particularly when time reduces cycles of iteration, and resources mean a small design and development team. Within a PhD timeframe, we iteratively built a bespoke software, SoundXY4, for Study 3. This software was suitable for evaluating the concepts of auditory feedback using a spatialisation technique on a TTI for music performance. We also built our own DIY tabletop system to deploy our software, which allowed us to keep the budget low, be independent, and have full control of the apparatus during the PhD research. See Appendix A.2 for an inventory of the components used in our DIY table, and Appendix A.3 for an explanation on how to build a rear DI tabletop.

At The Open University, we built a low-cost wooden interactive tabletop based on the commonly used rear DI technique, and inspired by the Reactable approach. The tabletop has a square surface that measures $87 \times 89 \times 100\,\text{cm}$ including four attached wheels. The table has an inner square frame of $58 \times 58\,\text{cm}$, in which there is an acrylic-diffuser sheet for projection and object detection. The size was designed to keep a balance between mobility and multi-player capabilities. There is a rim area, which is a non-interactive outer area of about $30\,\text{cm}$, useful for laying out objects. Figures 3.4 and 3.5 show

FIGURE 3.4: Early stages of the Woodentable

FIGURE 3.5: Final stages of the Woodentable

different stages of the Woodentable frame design, and Figure 3.6 outlines the parts of the tabletop interface from a top view.
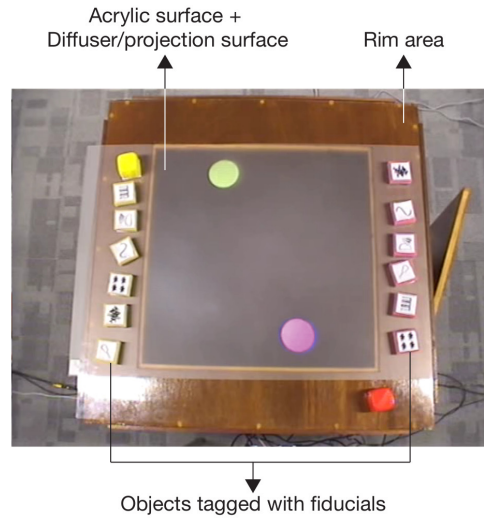
FIGURE 3.6: Woodentable top view
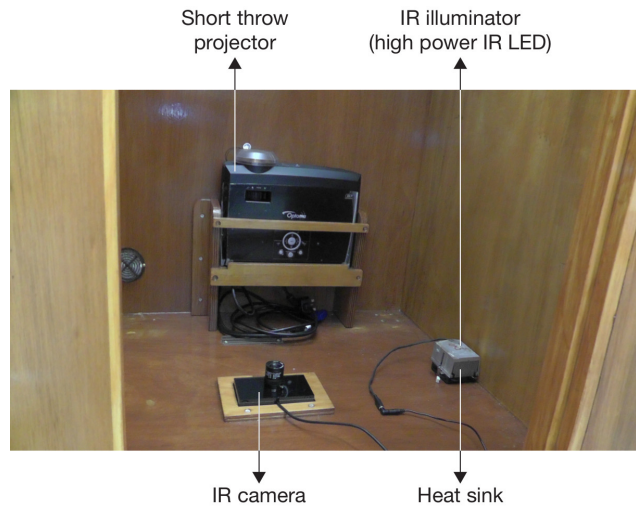
FIGURE 3.7: Woodentable inside view

A short throw projector was mounted inside the box facing the surface. This was suitable for our purposes, as using a lighter and finer-grained projector such as a pico-projector would have increased the cost. For object detection, we used a high power infrared (IR) LED mounted on a heat sink, and a camera with IR lens positioned at the bottom centre of the box (see Fig. 3.7). The object tracking was processed using reacTIVision.

TABLE 3.3: Video collected in the three studies

Study	Video collected	Total time
1	45 min (average) $\times$ 4 sessions $\times$ 4 groups	720 min $=$ 12 h
2	120 min $\times$ 4 sessions	480 min $=$ 8 h
3	15 min $\times$ 2 conditions $\times$ 8 groups	240 min $=$ 4 h

The tabletop frame has an opening door, which in Study 3 needed to be partly open for ventilation. That was not an issue for the performer next to the door, as there was a distance between the performer's body and the afore-mentioned door. Using a square shape surface for Study 3 was suitable for the purpose of working with groups of four people, and easier to implement.

3.7 Data collection

We collected video data in the three studies. We collected system log files in the last study.

3.7.1 Collecting video

We collected a total of 24h of video data. Table 3.3 unpacks the time of video collected for each study by session and group.

We were interested in capturing group interaction, at a body and gesture scale. In Studies 1 (§4.3.3) and 2 (§5.3.3), we used two video cameras: a panoramic view, and a closeup view of the tabletop surface. In Study 2, which was conducted in a public venue, we were particularly interested in capturing the social context beyond the group interacting with the table. In Study 3 (§6.3.3), we used 6 cameras: four dome cameras each facing a corner of the tabletop, and two floor-standing cameras for a closer perspective of the tabletop. We provide more details of the camera set up in the chapter on each of the studies.

3.7.2 Collecting system log files

It is easy to retrieve information on the interactions as a group in the form of scripts (interaction log files) using SuperCollider and reacTIVision in SoundXY4. For example, reacTIVision messages can be tracked. These scripts can be later analysed using statistical packages such as R.

In Study 3 (§6.3.4), we captured interaction log files. Interaction log files included information about the objects on the active tabletop surface (i.e. time, position, orientation, or identity). We were able to obtain this information at 30 fps from the TUIO messages sent by reacTIVision. See Appendix A.4 for an example of an interaction log file, in which we see information such as the timestamp, the ID of the object, the number of instance, the position in x-y coordinates, or the angle of rotation.

We used interaction log files for the experimental stage of this thesis. We aimed at complementing the information seen from below the tabletop surface (log files) with the interactions seen from above (video data). With interaction log files it is possible to understand space use on the tabletop surface over time, by rendering activity plots, for example. In HCI tabletop studies it is common to combine video data with interaction log files data.

3.8 Procedures for data collection

In the lab studies (Studies 1 (§4.3.4) and 3 (§6.3.4)), we observed how participants collaborated in an open task; whilst in Study 2 (§5.3.3), we observed in-the-wild group interactions in a public venue.

The procedures for data collection in the lab studies consisted in:

1. Explaining the activity (written and oral).

2. Asking participants for their written consent (see next Section 3.9).

3. Asking participants to improvise music for a set amount of time.

We conducted unobtrusive observation by leaving the room. We used video cameras positioned unobtrusively to capture the group activity. In Study 1 (§4.3.2), we administered a pre-questionnaire to aid the allocation of participants to the groups, dividing them into experts and beginners with the Reactable, as well as for gathering demographic data. In Study 2 (§5.3.3), we made a public announcement about at what time of the day the study was to be held; and we videoed group or individual activity. In Study 3 (§6.3.4), we administered a post-questionnaire to collect demographic data from participants.

3.9 Ethical approval and informed consent form for participants

Upon arrival, participants were informed of the purpose of the research, which was also explained in a covering letter, in addition the participants' questions were answered face to face.

Participants in all of the three studies were volunteers. In the lab studies, participants were asked to sign a consent form in which they were informed about:

1. Their freedom to withdraw at any time during the research, and to request the destruction of any data that had been gathered from the participant, up to the point at which data was aggregated for analysis.

2. That results of this research constituted personal data in compliance with the Data Protection Act, and so results were going to be kept secure, confidential and not released to any third party.

By signing this consent form, participants were giving their consent to:

1. Being video recorded.

2. Giving their permission for the data collected to be used in an anonymous form in any written reports, presentations and published papers relating to this study.

3. Accepting that their written consent would be sought before using any identifiable material related to them in this study.

The projects conducted in the UK (Studies 2 and 3) had the approval of the OU Human Research Ethics Committee. Study 1 was conducted in Spain, and we collected individual consent forms. See Appendix A.5.1 for an example of an informed consent for use in a lab study.

For collecting data in a public space such as in Study 2 (§5.3.2) we sought consent by using public signs and leaflets. See Appendix A.5.2 for an example of a sign for use in a science centre or museum.

The differences between collecting data in a public space such as a science centre and collecting data in a lab are:

1. The *level of control* differs as in a public centre it is difficult to know who will participate, whether there will be groups or not, and the number of team members; whilst in a controlled lab setting the number of groups and participants is known, as is what the activity will be, and how long the activity will last.

2. The *level of interaction* differs, as in a public centre participants are anonymous, the consent form is in-situ (using leaflets, for example), and there is no prior interaction of the facilitator explaining the activity; whilst in a controlled setting, the facilitator can collect demographic data from the participants.

In-the-wild data can thus offer a level of surprise and unpredictability that, depending on the research question (for example, exploratory research questions), can be interesting; whilst in a controlled lab setting there tends to be a reduction of unpredictable elements.

3.10   Methods for data analysis

This thesis uses mainly qualitative methods for data analysis. Qualitative research is a broad term. Hammersley comments:

> "We can define 'qualitative research' along the following lines: a form of social inquiry that tends to adopt a flexible and data-driven research design, to use relatively unstructured data, to emphasize the essential role of subjectivity in the research process, to study a small number of naturally occurring cases in detail, and to use verbal rather than statistical forms of analysis." [Hammersley, 2013, p. 12]

The author explains that there are many ways of doing qualitative research. A degree of in-depth detail is desired to characterise as much as possible the observed social setting, what Geertz termed *thick descriptions* [Geertz, 1973]. For an example of thick transcriptions from transcribing video see Mavers [2012].

Video in qualitative research is commonly used as discussed above (see §3.2). There are a number of approaches to analysing video. For example, multi-modality [Jewitt, 2013] looks into a detailed micro-analysis of video extracts focusing on modes (e.g. gaze, gesture, posture, speech). Similar to multi-modality, we are interested in transcribing video considering verbal and non-verbal communication (see Mavers [2012] for different written or graphical multimodal representations of transcribing video). A multimodal approach of the different modes by participants would be highly time-consuming and focus on a small percentage of our collected video data. We are interested in getting a sense of the whole dataset, and then focusing on particular examples, rather than focusing in on short video extracts at a micro-level. In particular, we are interested in capturing patterns of interaction focusing on verbal and mostly non-verbal communication.

In summary, our research adopts a qualitative approach to analysing video data. This approach is supplemented with quantitative analysis of interaction system files, as appropriate, for instance when certain aspects need

to be quantified, for example when comparing two conditions in Study 3 (§6.3.1, §6.3.4). In the next sections, we detail the methods used in the two approaches: qualitative (§3.10.1, §3.10.2, §3.10.3), and quantitative (§3.10.4, §3.10.5).

3.10.1 Video analysis: qualitative approach

Our qualitative video analysis is based on a social science approach to analysing collected data (e.g. video) through the identification of emerging themes [Lofland et al., 2006].

In particular, we used:

1. The *interaction analysis* method (§3.10.2) that helps to describe videoed sequences of social events by splitting them into verbal and non-verbal communication [Jordan & Henderson, 1995], used in Study 1 (§4.3.5).

2. The *thematic analysis* method (§3.10.3) for identifying emerging themes from video data, adopted in Studies 2 (§5.3.4) and 3 (§6.3.5).

In video analysis, it is a common practice to exemplify the patterns that are present across the data using *vignettes* [Heath et al., 2010]. Vignettes are representative examples of video extracts that illustrate these patterns (i.e. thumbnail narratives). Vignettes are thus illustrative rather than exhaustive [Lofland et al., 2006]. Transcribing all video data is highly time-consuming, and not necessary for the purposes of this study.

Figure 3.8 shows the workflow for Study 1 whilst Figure 3.9 shows the workflow for Studies 2 and 3. The main difference between these two workflows is that, in Study 1, we looked into the video data in respect of themes informed by the literature review (top-down approach) along with emerging themes from the video data (bottom-up approach). While in Studies 2 and 3 we had a bottom-up approach only. In Study 3, we also transcribed speech of
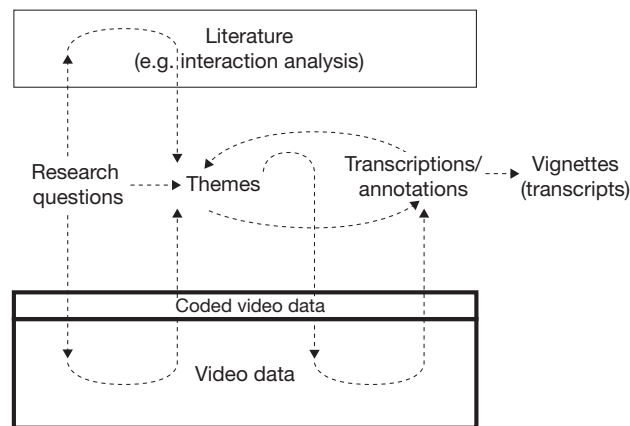
FIGURE 3.8: Video analysis workflow of Study 1: top-down and bottom-up approach
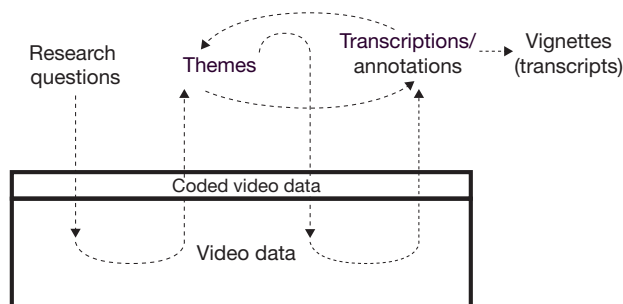
FIGURE 3.9: Video analysis workflow of Studies 2 and 3: bottom-up approach

all video data related to our research questions in order to provide a quantitative comparison between two conditions.

To avoid data overload and to manage the data analysis, we used transcription in a focused manner, we:

1. Coded the video about the main events (by timecode and description).

2. Looked at the video data in respect of the research questions (with themes in mind in Study 1).

3. Identified a set of themes (in all studies bottom-up approach, and also top-down approach in Study 1).

4. Transcribed a set of vignettes by either:

- Selecting a collection of illustrative video extracts for each of the identified themes, and then transcribed those considered to be most representative (Studies 1 and 2).

- Transcribed all of the video data by the identified themes, and then selected a collection of illustrative video extracts for each of the themes (Study 3).

3.10.2 Interaction analysis

The principles of *interaction analysis* as stated by Jordan & Henderson [1995], describe a practice of video analysis for studying interaction between humans, and between humans and artefacts, in an environment. Accordingly, both verbal and non-verbal activities can be the focus of analysis, which the authors divide into *talk-driven interaction* and *instrumental interaction*, respectively. The latter refers to activities mainly driven by the manipulation of physical objects (e.g. technologies or human-made artefacts), where talk may happen as subsidiary to the physical activity:

> "In the course of such instrumental interaction, talk may and usually does occur, but it is not central as in talk-driven activities. It is usually ancillary to, supportive of, and sometimes even coincidental to the main business at hand. Thus, the conversation of two surgeons about the morning's difficult operation is an instance of talk-driven interaction; but the actual surgery (including the talk that happens in its course) is instrumental interaction" [Jordan & Henderson, 1995, p.65]

Music performance on TTIs is a collaborative task, in which non-verbal communication between musicians, and physical interaction with the tangible objects, are both key elements of the activity. Interaction analysis is a suitable approach for understanding group interaction in TTIs for music performance because it provides an appropriate lens for understanding what people do during practical activities, particularly where object-manipulation is a central part of the activity. Interaction analysis was used in Study 1 (§4.3.5) to

help us understanding the role that non-verbal communication played in the development and sharing of ideas between musicians mediated by the Re-actable.

3.10.3 Thematic analysis

Thematic analysis (TA) is a qualitative method widely used in social sciences such as psychology [Braun & Clarke, 2006]. "Thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data" [Braun & Clarke, 2006, p. 6]. Thematic analysis is flexible and this enables it to be used as a bottom-up or top-down approach to data analysis, in both approaches the aim is to identify emerging themes or patterns. We have used thematic analysis in Studies 2 (§5.3.4) and 3 (§6.3.5) as a bottom-up approach. In Study 1 (§4.3.5), we combined bottom-up and top-down approaches for identifying emerging themes as well as using existing themes from the literature. In TA, the selection of themes depends on the criteria of the researcher, which are expected to be consistent. For the themes we focused on our research questions.

3.10.4 Statistical analysis

The analysis of the data included descriptive statistical analysis to afford a sense of the participants' profiles, and comparative statistical analysis of the interaction log files in the experimental study. We used the *Student's t-Test*, which is a suitable test for comparing conditions statistically, as explained in Appendix A.6. We analysed differences between space use in terms of ranges in the x axis or y axis; mean and standard deviation of the x or y coordinates; and overall area of the distance travelled. Although most of the time we did not find statistically significant results in Study 3 (§6.7), we have included the results that were most salient. In Study 2, we also used the normal distribution and exponential distribution statistical tests (§5.5).

3.10.5 Activity maps

In our experimental study, we generated *activity maps* or *activity plots* of the whole interaction. Activity plots provide a graphical representation of the location and amount of group activity, in this case musical improvisation, in terms of the position of the tangible objects on the tabletop surface for a period of time. Each map shows, in a single plot, every time that a group member sets an object, moves an object, or releases an object during a session. Activity plots are used in tabletop studies for understanding patterns of interactions. See for example Nacenta et al. [2010]; Scott et al. [2004]; Tuddenham & Robinson [2009]. A difference between our activity plots and these authors's activity plots is that they trace each of the participants by a different colour, whilst we use the same colour because we focus on group interactions. Another difference is that we use dots of one size that represent tangible objects; whilst Nacenta et al. [2010] plot dots with different colour intensity of magnetically-tracked pens; Scott et al. [2004] tracks traditional media (pen and paper) with different sizes of the dots to represent the amount of activity; and Tuddenham & Robinson [2009] plot dots of same size that represent digital stylus activity. There are no written rules for plotting tabletop activity, plots are designed according to the research questions.

3.11 Methodological limitations

We now turn to discuss the methodological limitations, internal validity, and external validity of this work.

What people do. One potential limitation of this study is that we have focused on understanding what people do through observation, and as this process lies on researcher interpretation, it can be shaped by researcher bias. In order to overcome potential biases related to our subjective interpretation when watching videos of what people do, we sought expert consultancy: we facilitated group meetings in Study 1 (§4.3.5), and participated in two analytical workshops using the video data in Study 2 (§5.3.4). The aim of these

meetings was to discuss and find a common ground about the themes and vignettes. In Study 3 (§6.3.5), we had supervisory meetings which focused on specific aspects. We elected not to conduct participant interviews as the focus was on collaborative music performance on interactive tabletops, and we considered that the constant fast-paced interaction captured on video was rich enough for our purpose.

Sample size. A second possible concern is the sample size. We have worked with a small sample size, mostly in lab studies, and with a larger sample size in the field study. This sample size is suitable when working with a detailed qualitative analysis. In Studies 2 and 3 we combined qualitative analysis with descriptive statistical analysis to strengthen the data analysis as we had larger groups than in Study 1. Our results can shed light on directions for future studies that can include quantitative research questions.

Collaboration with strangers in lab studies. A third possible critique of the study design is our decision to study collaborative music performance through the lens of musical improvisation with groups who not necessarily know each other in the lab (Studies 1 and 3). Our approach was to achieve ecological validity through unobtrusive observation of a variety of musical improvisation in an informal lab setting in order to keep a trade-off between openness and control. Social action has been studied in lab settings in the past [Rooksby, 2013; Suchman, 1987]. Rooksby argues that it is possible to do fieldwork in the laboratory ('wild in the laboratory') drawing on Suchman's photocopier study. Furthermore, we are working with new technologies, which are of difficult to access by musicians and people unrelated to tabletop research. To our knowledge, there are no existing communities (yet) in the real world that we could have conducted naturalistic observation of (with the exception of visitors in public venues, as in Study 2). In addition, these devices are relatively unstable, large, and still difficult to move out of a lab setting. In terms of the groups, most group members did not know each other, and if they did, to our knowledge, they had never played music together before. Even though these situations, in Study 1 we worked with professional musicians who were expected to know the protocols in musical

improvisation. In Study 3 we worked with heterogeneous groups who were expected to know how to do an improvisational activity; and to be familiar with a collection of everyday sounds. The observed musicians' behaviour in the lab studies (e.g. group dynamics, individual actions) indicated that we succeeded in achieving ecological validity.

Internal validity. For internal validity, in each study there were early informal pilot tests to confirm that the procedures and measurements were clear and unambiguous. We assumed a strong internal validity of the exploratory studies (Studies 1 and 2), in which we observed real uses of a commercial product that has been designed iteratively by a professional team over years. Moreover, these studies were consecutive in time. They were also designed building on research experience gained from previous studies.

External validity. External validity relates to what extent the study is generalizable. In social sciences, there is consensus between qualitative researchers in redefining this generalizability, moving from the classical concept of laws that apply universally, typically found in sciences, to results that can be generalizable to some extent to similar situations [Schofield, 1993]. As the studies of this thesis include qualitative research methods, this approach of rethinking generalizability is more suitable than trying to claim general statements. Each of the three studies contains fine-grained descriptions, which allows one to make informed judgements of whether these findings can be useful in other (similar) situations.

3.12   Summary

This chapter presented the overall methodological design of this thesis: the use of video as a research tool for understanding what people do in collaborative music performance on TTIs.

In particular, we outlined the commonalities and differences of our three empirical studies (two exploratory, one experimental) in terms of aims, apparatus, settings, participant recruitment, sampling and group profiles, data

collection and data analysis. We explained our approach of working with participants in casual set-ups using unobtrusive methods, and our focus on musical improvisation using two apparatus: 1) the Reactable for Studies 1 and 2; and 2) the Woodentable and bespoke software SoundXY4 for Study 3. We stated our procedures for data collection, including seeking consent from participants in compliance with the Data Protection Act and university ethical guidelines. We briefly described our methods for data collection using video cameras and system log files. We also highlighted the methods used for video data analysis, including interaction analysis and thematic analysis; and log files analysis including statistical analysis and activity plots. We finally discussed the main methodological limitations of using a qualitative approach to focus in on understanding what people do using video, the small sample size, and situating the research in informal lab settings. We have also argued that despite the limitations of the study design, video-based unobtrusive observation of rich and open activities such as musical improvisation in lab settings is the most apt for initial in-depth understanding of group interaction on a new interface for collaborative music. Furthermore, we have argued that the number of groups observed has been sufficient for identifying both group dynamics and behaviour patterns expected in a larger scale.

The next chapter presents the first empirical study of this thesis, an exploratory study on collaborative music performance by groups of expert musicians improvising with the Reactable, over multiple sessions, in an informal lab setting.

Chapter 4

Study 1: Expert Peer Learning with the Reactable

Our first study investigates groups of expert musicians improvising on the Reactable during four long sessions, taking place over a week in a lab setting. Our approach focused on exploring the social aspects of group interaction and workspace awareness over time, drawing upon ideas of situated, hands-on, and practice-based peer learning. From the lessons learned, design considerations of how to best support long-term interactions, in particular hands-on tabletop collaborative learning, are discussed.

4.1   Introduction

As explained in §3.5.2, even though existing studies show that tabletop systems can be suitable for collaboration in the short term, little is known about how can they be best used by musical experts in the long term. With a few exceptions [Piper & Hollan, 2009; Wigdor et al., 2007], most tabletop studies look into one-off interactions and 'low entry level': that is, how participants interact or what participants learn during a short period using simple prototypes (cf. Shaer & Hornecker [2009]). Thus, we know little about how to support the development of group coordination and work practices using

95

more complex interfaces than simple prototypes. We argue that it is necessary to move beyond studies of simple demonstrator applications. We need to detail how people use tabletops for a variety of real purposes, in order to better understand the strengths and limitations of this novel platform. Thus, we are interested in observing the level of skill acquired over longer periods of time using complex interfaces.

This study described in this chapter constitutes the first detailed examination of collaborative learning over multiple sessions with a tabletop system, and explores group development with a novel interface in an unconstrained, unguided task, within which a group can develop their own musical and social language. We are interested in studying collaboration over time in terms of *group tabletop interaction* and *workspace awareness* (for a definition of the term, see §1.3). We describe findings from video analysis of four groups of expert musicians improvising music using the Reactable over multiple sessions. We investigate change over time looking through four lenses: the *Reactable interface*'s characteristics, *tangible interaction*, *musical interaction*, and *group interaction*.

We anticipate that the study will inform future design and analysis of colocated tabletop interaction for small groups.

4.2   The system

The Reactable is a commercial real-time modular synthesizer [Jordà et al., 2005], which is used in professional music contexts, as well as in public spaces such as museums, science centres and exhibitions. It has a circular tabletop surface, and it combines multitouch with tangible object manipulation as input. For this experiment we used a Reactable Experience model with a 100 cm diameter including a rim area of 10 cm. The Reactable's TUI enables the construction of a variety of configurations of building blocks to produce sound. As shown in Fig. 4.1, a set of physical objects allows users to create music by building *audio threads*, each thread representing an audio channel (i.e. a
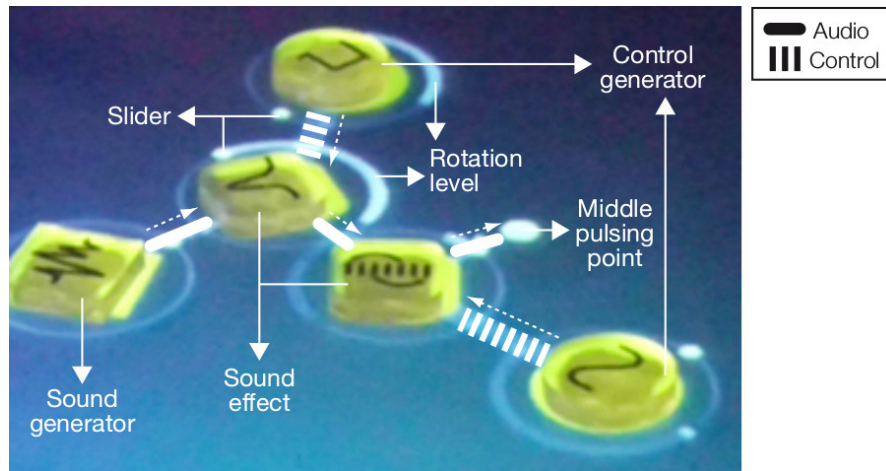
FIGURE 4.1: A Reactable's thread

sound source that can be operated individually). Here, we use the term *audio thread* to refer to the physical connection between objects, whilst *musical voice* refers to the character of the musical sound resulting from this connection (e.g. melodic voice).

A white pulsing point in the middle of the table area represents the sound output, as well as the tempo of the table. Every audio thread connected to this point is audible and synchronised to the same global tempo. Each thread is shown in a different colour, taken from a defined palette, and is built from interconnected objects. The sum of the audio threads constitutes a *patch*, which represents a piece.

The objects have different functions, each represented by different shapes:

- *Sound generators* (SG), represented by squares and cubes, to create sounds, in which there is a subset of *instrument generators*.

- *Sound effects* (FX), represented by rounded squares, to transform sounds.

- *Control generators* (CT), represented by circles, to control other objects.

- *Global controls* (GL), represented by polygons, to control global parameters.

A thread needs a minimum of one sound generator in order to generate sound. A player's own sound files or samples can also be loaded and associated with the different sides of a cube. These sounds are played in a loop.

There is immediate real-time feedback when the tabletop system recognises the objects on the table surface (as explained in §3.6.1), and any change is represented both aurally and visually. In particular, on the right side of each object there is a projected slider that can be controlled with the finger; and on the left side of each object there is a display for the primary object parameter, which changes depending on the rotation angle of the object. While most interaction with the Reactable is carried out via the tangible objects, users can use touch input to, for example, mute or unmute the audio connection within a thread. Usually it is possible to change from one to three sound parameters for each object—by altering the rotation angle, the position of the projected slider, or the distance to the next object in the thread (or the distance to the centre in the case of the last object in the thread). There are also *special objects* (SP) such as the *radar trigger* or the *programmer*. The radar trigger is a special object that works as a local tempo controller with local up to global effects on all objects in its range. The speed of the radar is controlled by rotation, whilst its range is controlled by the projected slider on its right side. The programmer is a special object that is used to reprogram cubes and instrument generators with samples.

In the Reactable, each object has a number of inputs (from none to multiple) and outputs (one or none) depending on its category, which make connections between objects possible. Connections can be either *control signals* (i.e. when the destination is a parameter of a unit generator) or *audio signals* (i.e. when the destination is either an audio input of a unit generator or the global audio output). Figure 4.1 shows both type of signals. For example, if we connect a square low frequency oscillator or LFO (CT) to a noise generator (SG), the square LFO sends control data based on a square wave to the noise generator. If we connect a noise generator to a resonant filter (FX), the resonant filter only lets pass a part of the audio signal.

The Reactable supports *dynamic patching* (see §2.5.2 for further information), a mechanism that automatically connects the inputs and outputs of sound objects that are close to one another, as if by magnetic attraction, instead of requiring the user to connect objects manually. Furthermore, the Reactable interface is designed on a modular and constructive basis [Jordà, 2008]: a set of tangible building blocks, each with a specific function that can be interconnected to create both basic and complex configurations, provides a *modular interface*. Modularity has been identified as a design feature for learning through construction processes with TUIs [Zuckerman et al., 2005].

4.3 The study

Using the Reactable, the study investigates the nature of group tabletop interaction (GTI) and how workspace awareness (WA) is manifested among expert musicians in multiple sessions; in terms of the Reactable's interface, tabletop interaction, musical development, and group development.

4.3.1 Study design

Our study investigated the progress of collaborative learning and improvisation over time on the Reactable. While the Reactable is exhibited in various museums around the world, its primary purpose is to provide expert musicians with a novel interface for musical expression. In museum settings, users tend to be complete novices with the interface, and they rarely have time to gain extended experience with it. Understanding its longer-term use by expert musicians necessitated a study setting that resembled improvisational sessions by groups of musicians. Since our interest is in both 1) group progress in interface understanding, and 2) group coordination in musical improvisation; we chose to study participants who are not already accomplished Reactable performers.

We investigated four groups of co-located musicians collaborating around the Reactable: each of these played with the system in four sessions, which

FIGURE 4.2: Participants jamming the Reactable

were scheduled in close succession over the course of one week. Given that our participants were already expert musicians with theoretical knowledge of sound generation, this enabled us to observe the initial phase of getting accustomed to the Reactable interface and its rapid appropriation into musical improvisation. All sessions were video recorded and took place in a lab setting.

Our approach follows that of other studies (cf. Hornecker et al. [2008]), beginning with Suchman's study of photocopier use [Rooksby, 2013; Suchman, 1987], in conducting a naturalistic observation of activity within a lab setting. As shown in Fig. 4.2, we attempted to create a casual set-up, resembling the settings of rehearsal rooms, where musicians gather to play. Thus, setting and activity type were designed to be familiar for our participants.

4.3.2 Participants

Twelve males aged 22–54 ($M = 32.7$, $SD = 7.4$), from a department of Music Technology, volunteered to participate in the study. We teamed them up in four groups: one of two people, two of three people, and one of three to four

TABLE 4.1: Summary of groups

Group	Participants	Reactable's expertise
G1	M1, M2, M3	Experts
G2	M1, M2, M3	Experts
G3	M1, M2, M3, M4	Experts
Gb	M1, M2	Beginners

people (initially three, with a fourth joining for the last two sessions). Even though the group members knew each other, they had never played together before. All participants had a medium to substantial degree of musical training, being either music technology students, music practitioners or professional musicians. Of these, five were active practitioners of electronic music with synthesizers, electronic devices or computers. Most of the participants were already familiar with the Reactable: five participants reported they had "some" familiarity with the technology and how it works, seven reported themselves as having "a lot" of familiarity. This meant some had played the Reactable before, some had been introduced through a course, some had the mobile version for smartphones and tablets, and some had watched online tutorials and videos. Even though some of our participants were knowledgeable, there was still a contrast with musicians such as Carles López,[1] who have developed, over years, expertise through individual practice. We were interested in observing the earlier stages of such development.

Only one of the four groups had no experience of using the Reactable's TTI version: we named this the beginner group, although its members were still expert musicians. Participants were international (9 from Europe, 1 from North America and 2 from South America). In the following, G1, G2 and G3 are used to refer to the three Reactable expert groups and Gb to refer to the Reactable beginner group; M#musicianG#group to refer to each of the 12 musicians and S1 to S4 to refer to the four successive sessions of each group (e.g. "M1G1 in S2 initiates a new thread" or "shared thread in GbS4"). Table 4.1 outlines the four groups.

---

[1]Carles López performance: `www.youtube.com/watch?v=x8WuWagPTwk` (accessed 30 September 2014).
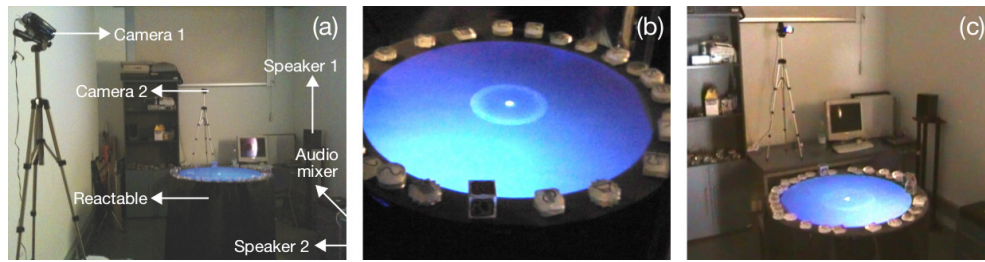
FIGURE 4.3: (a) Music lab set-up (b) Close-up view of camera 2 (c) General
view of camera 1

FIGURE 4.4: Video composition for video analysis

4.3.3 Set-up

We carried out the study in a lab located in the music studio area of the Universitat Pompeu Fabra in Barcelona. The room is isolated from the busy classroom areas and has a permanent Reactable in the centre of the room for rehearsals and user studies. The lab has a sound proof door, which is common in recording studios. We opted for a dimmed light environment, which is common in rehearsal and performance settings.

All sessions were video recorded by two cameras on tripods positioned nonintrusively: one with a more general view, a second with a more close-up view, as shown in Fig. 4.3. We later synchronised these two video sources in a single file and view for easing video analysis. A snapshot of the single view is shown in Fig. 4.4.

4.3.4 Task and procedure

We conducted the study during two weeks in April 2011. In the first session, we asked participants to sign a consent form. Participants were asked to improvise music in group with the Reactable during four consecutive sessions over the course of one week. Sessions were said to last for 15 min–45 min. Participants could check a printed copy of the Reactable user manual if needed, could stop at any time during the session, and were notified one minute before the session end.

The set of Reactable objects for this experiment comprised 39 objects, which were initially distributed around the rim of the table: 12 SG, 10 FX, 10 CT and 7 GL/SP. Within each of the four categories, almost every object was different; although a few were duplicated (FXs, CTs). An electronic version of the Reactable user manual was sent to the participants the day before their first session. Before each session, participants had the option to load their own samples to be used by sending them to the facilitator.

The facilitator intervened at the beginning and end of each session to set up or shut down the system, trying to be as unobtrusive as possible and encouraging participants to act as they would in a real context. The facilitator moved to a room next to the music lab during the sessions, and only checked on the activity from time to time. Otherwise, participants had complete control of the session: for example, they were told they could stop the cameras if they preferred a shorter session, they could control the audio mixer or turn the output of the speakers up or down at any point (these were to one side of the Reactable).

4.3.5 Method

The study was concentrated in four areas of interest:

- **Interface characteristics**: how the Reactable's interface characteristics influenced group behaviour over time, in particular its dynamic patching mechanism and lack of territorial constraints.

- **Tangible interaction**: how the properties of the tangible interface facilitated group progress and the development of expertise; the nature of gestures in group tabletop interaction and learning; and the usage of tangible objects.

- **Musical improvisation**: how tabletop musical improvisation brought new challenges compared to those encountered in traditional ensembles; and how the organisation of improvisation developed over time.

- **Social factors in the group development of expertise**: the nature of collaborative learning through constructive processes on the TTI; and how this supported different group learning styles.

We used *interaction analysis* [Jordan & Henderson, 1995] to analyse the video in detail and identify themes from both bottom-up (emerging from data) and top-down (driven from the literature) approaches: some emerged from watching the video data, such as *interface explorations* or *peer learning*; others were partially developed from our overarching research questions, such as the *musical techniques* employed. Other themes were inspired by Jordan & Henderson [1995], such as our analysis of *beginnings and endings* of sessions, or of *error/repair situations*. We used Elan,[2] which aids video coding, to analyse the videos.

Representative video extracts were selected by the author and repeatedly viewed and discussed by the author and two expert researchers in video analysis and HCI, Eva Hornecker and Paul Marshall. We focused on verbal communication (VC) and non-verbal communication (NVC) themes, of which those related to NVC were divided into musical, physical and interface-related (see Appendix C). We also focused on lower-level categories specific to the Reactable interface, such as *territories and thread ownership*. For space reasons, only the most salient themes related to the above perspectives are discussed in the following sections. Next, we present our *coding scheme*: it includes the most common study-specific terms used in our video analysis.

---

[2]See: `www.lat-mpi.eu/tools/elan` (accessed 30 September 2014), developed by the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands [Sloetjes & Wittenburg, 2008].

In this chapter, we discuss a number of video vignettes. Links to the video extracts analysed are listed in Appendix B.

4.4   Coding scheme

In this section we present the codes used for data analysis. The coding scheme was designed to allow us for the transcription of behaviour and events that concern Reactable's interface (§4.4.1, §4.4.2), tangible interaction (§4.4.3), musical interaction (§4.4.4), and social peer interaction (§4.4.5). For the latter, verbal and non-verbal communication behaviours are included. The themes are defined with more detail in Appendix C.1 and C.2. Themes are discussed in the findings sections (§4.6–§4.9).

4.4.1   The Reactable's interface coding scheme

We observed and analysed a number of Reactable's interface-related behaviours, including:

- *Dynamic patching*: we include here those *intentional* or *unintentional* interaction events related to the automatic connections present in the Reactable interface.

- *Territories & threads ownership*: events related to the Reactable's distribution of territories and ownership of audio threads, including invasions, takes, gives, and individual/shared threads; which are detailed below. Our interest was in how participants dealt with the trade-offs of sharing an interface, and what kind of territorial behaviours might emerge.

- *Use of special objects*: use of global and special objects such as the *radar trigger* or the *programmer*.

4.4.2 Territories and threads ownership

We observed and analysed a number of territory-related behaviours, including:

- *Invasions*: 'interfering' in somebody else's thread via an action.

- *Takes*: taking an object that 'belongs' to somebody else for individual use. Takes can be *active takes* (taken from the surface table area) or *passive takes* (taken from the rim area close to another person who is not currently using it).

- *Gives*: handing an object to somebody else for individual use, which can be *active gives* (given on the surface table area) or *passive gives* (given via the rim area).

- *Individual/shared threads*: a thread can belong to an *individual* (thread built by a single person), or be *shared* (thread built in collaboration) as shown in Fig. 4.5. Scott & Carpendale [2010]; Scott et al. [2004] distinguish between *personal territory* as a workspace close to the person, including storage space; and *group territory*, such as the centre of the surface table area or the spaces between collaborators (on the Reactable the global tempo pulsing point and sound output constitute the centre). We assume that the distribution of personal and shared areas depends on the number of users and the shape of the tabletop surface (see Fig. 4.6). While territories assume a spatial distribution of ownership, it was also necessary to take account of the fact that thread ownership is object-based and may change dynamically as threads change, irrespective of the position of the individuals.

4.4.3 Tangible interaction coding scheme

We observed and analysed a number of tangible interaction-related behaviours, including:
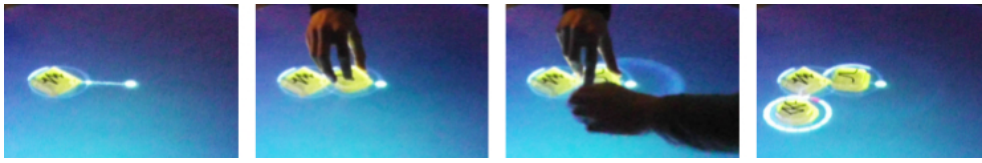
FIGURE 4.5: A shared thread
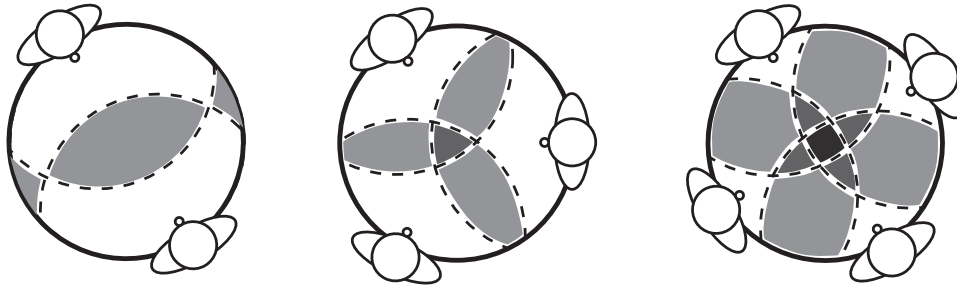
FIGURE 4.6: Imaginary divisions of personal and shared areas

- *Basic techniques*: basic interaction styles with the objects, including:

  - *Dragging*: moving an object along the active surface while maintaining continuous contact with the object.

  - *Rotating/sliding*: either rotating an object; or changing the position of the projected slider with a fingertip.

  - *Strobing on/off*: placing an object on and off the surface table area in a rhythmical or non-rhythmical pattern for a musical effect.

  - *Twisting*: quick and rhythmic rotations of the same object either from left to right, or from right to left.

- *Complex techniques*: complex interaction styles with the objects, including:

  - *Dragging & strobing*: moving an object along the active surface while placing it on and off the surface table area in a rhythmical or non-rhythmical pattern for a musical effect.

  - *Preview technique*: musicians first built silent thread configurations, and then activated them by adding a SG, which activates the thread.

- – *Stacking*: piling objects.

  – *Tossing/rolling*: throwing an object casually, with no sense of which side it will land on.

- **Basic configurations**: basic arrangements of objects in a particular form, including: *1 single SG*, *1 FX at the end of the thread*, and *1 CT with a SG or a FX*.

- **Complex configurations**: complex arrangements of objects in a particular form, including: *2–4 FXs at the end of the thread*, *a CT connected to a SG connected to a FX*, and *two-branch thread*.

- **Coordination**: non-verbal joint action in group.

  – *Bobbing head*: short rhythmical movements with the head, up and down.

  – *Handover*: handing an object over to another musician.

- **Error/repair situations**: a person starts an unintended effect (e.g. a mistake) and then potentially tries to fix it.

- **Gestures**: physical and musical-related events linked to hands and body actions including *sound producing gestures*, *ancillary gestures*, *communicative gestures*, and *sound accompanying gestures*; which are detailed below. In a study of clarinettists' movements, Wanderley & Vines [2006] used the term *ancillary gestures* for those gestures that, without being strictly necessary, support the sound-producing gestures. We here focus on non-mutually exclusive movements made by group members that are relevant to tangible and social interaction. We based our analysis on Jensenius et al. [2010]'s typology—a detailed investigation of musicians' gestures is beyond the scope of this dissertation:

  – *Sound producing gestures*: in our study, sound producing gestures are connected with *instrumental interaction*, constituting "activities that crucially involve the manipulation of physical objects" [Jordan & Henderson, 1995, p. 65]. For example, strobing.

- *Ancillary gestures*: gestures that support sound-producing gestures. For example, coordination gestures, such as a handover.

- *Communicative gestures*: gestures for communication between performers and the audience, including eye contact. For example, coordination gestures, such as *deictic gestures* (pointing gestures).

- *Sound accompanying gestures*: body gestures not related to sound production. For example, coordination gestures, such as bobbing heads.

- **Physical-digital explorations**: physical and/or musical-related events linked to the action of exploring. For example: dragging, rearranging, strobing objects to ones' own and/or others' threads, with both musical and interactional impact. Or also stacking, rolling or tossing objects not necessarily with any musical result.

4.4.4 Musical interaction coding scheme

We observed and analysed a number of musical interaction-related behaviours, including:

- **Solos**: leading melodic and/or rhythmic voice operated from an individual thread.

- **Musical explorations**: physical and/or musical-related events linked to the action of exploring. For example: building serendipitous threads or dragging an object along different threads, with interactional and/or musical impact.

- **Non-participation**: a group member stops for several seconds or minutes (i.e. standing back), leaving his individual space with an active thread, but sometimes with none.

- **Intros/endings**: a group starts or finishes the session. The beginnings and endings of a musical improvisation session refer to what in popular and jazz music are traditionally known as intros and endings. In

interaction analysis, beginnings and endings are considered meaningful units in a structured sequence of events, because they can tell us about collaborative negotiation during the start and end of an activity [Jordan & Henderson, 1995, pp. 57–59]. We identified the following types of ending techniques.

- – *Fade out*: incremental decrease of volume (e.g. using global volume or modifying the volume parameter of a SG).

- – *Global object*: use of an object with global effect, including global feedback, global volume, global metronome (tempo), panning, and radar trigger.

- – *Objects removed sequentially*: starting from one or multiple threads for each musician to objects removed one after the other.

- – *Serendipity*: presence of serendipitous actions (e.g. massive use of all the objects or randomly tossing an object).

- – *Shared threads*: presence of a shared thread (e.g. starting from one thread for each musician to one shared thread).

- *Musical dialogues*: conversation between at least two leading melodic and/or rhythmic voices operated from an individual thread each. Our analysis looked at the nature of these dialogues based on dichotomies such as homophony (i.e. a single voice plays a melody) vs heterophony (i.e. multiple voices playing a melody with variations), and elements of musical forms such as call-response or rhythm vs melody (cf. Pressing [1984, 1988]).

- *Transitions/changes*: the individual or collective process of moving from one motif to another motif (e.g. the process of moving from one individual leading melody to another or the group process of moving from one consistent set of leading melodies and/or rhythms to another consistent set).

4.4.5 Social peer interaction coding scheme

We observed and analysed a number of social peer learning-related behaviours, including:

- *Explicit peer learning*: a group exchanges verbal information to solve a problem in collaboration (e.g. verbal exchange related to solving an error/repair situation).

- *Mimicking*: at least one person imitates another person's actions or behaviours with both musical and interactional impact.

4.5 Findings: overview and summary of the vignettes

The sessions tended to last from 35 to 45 minutes, but sometimes longer: All groups used the full time available for their sessions until the room had to be freed. Groups used cubes differently: G1 and Gb only worked with preloaded samples, whereas G2 (M1 in S4) and G3 (M1 and M2 in S2–S4, M3 in S4) loaded their own sounds.

An environment resembling a rehearsal music room setting was successfully created. Participants demonstrated enthusiasm to start the session, generally arriving on time, sometimes even earlier, for their allocated sessions. There were no comments about whether they had studied the manual in-between sessions, but those who loaded their own sounds took extra time to prepare and send them to the facilitator.

Group dynamics tended to be as diverse as in a real rehearsal: in the beginner group Gb, one participant arrived five minutes after the scheduled time twice, and thus the other group member of the duo group started without him. In G3 (initially a group of three), a further member joined the group for the last two sessions. There was generally a relaxed and informal atmosphere: in Gb one musician left the room during a session to attend an urgent phone call, whilst the other musician kept playing (see Fig. 4.9); and in all

TABLE 4.2: Overview of vignettes

Section	Topic	Vignette	Figure
Reactable interface (§4.6)	Dynamic patching	1. Triggering unintentional effects with dynamic patching	4.7
		2. Triggering intentional effects with dynamic patching	4.8
	Territoriality	3. Dynamic reconfiguration of individual territory	4.9
		4. Complex invasion with the radar trigger object	4.10
Tangible interaction (§4.7)	TUI control	5. Preview technique	4.11
	Gestures	6. Handover	4.12
		7. Bobbing heads	4.13
		8. Physical-digital exploration: Tossing	4.14
Musical interaction (§4.8)	Group musical interaction	9. Basic vs complex intro	4.15
		10. Basic vs complex ending	4.16
Social peer interaction (§4.9)	Mimicking	11. Mimicking	4.18
	Explicit peer learning	12. Explicit peer learning from error/repair situation	-
		13. Explicit peer learning from showing "the sync"	4.19

groups musicians approached either the audio mixer or the speakers to turn the speakers' volume level up or down when needed. All groups asked for a copy of the videos of the sessions and G2 reported that the group would follow-up by meeting and rehearsing together.

In the next sections, we detail the findings relating to the evolution of group tabletop interaction and workspace awareness over time. We looked at four aspects divided into one section each:

- The Reactable's interface control over time (§4.6).

- Tangible interaction development (§4.7).

- Musical development (§4.8).

- Social peer learning (§4.9).

Table 4.2 shows a summary of the vignettes discussed.

4.6 Findings: Reactable's interface control over time

In this section we present findings concerning features specific to the Reactable interface. In particular, we explain the change of use over time of the Reactable's interface, including the use of dynamic patching; individual territories; shared territories; and the object radar trigger related to thread ownership.

It is notable that:

1. Dynamic patching and related serendipitous effects were used more intentionally over time.

2. The position of musicians was dynamically reconfigured depending on the number of musicians in the room.

3. Individual spaces dynamically emerged irrespective of the lack of space divisions.

4. Shared threads were used for musical purposes with greater control over time.

5. The use of the radar trigger in later sessions promoted a more advanced notion of thread ownership than the simple dichotomy between individual and shared threads.

These findings indicate that the Reactable's interface prompted fluid transitions between individual and group work; fluid group configurations around the table; and a greater Reactable's interface control over time within a situated context.

4.6.1 Serendipity with dynamic patching: from unintentional to intentional

We found several examples of participants triggering *unintentional effects* due to the Reactable's *dynamic patching* mechanism. Unintentional interference
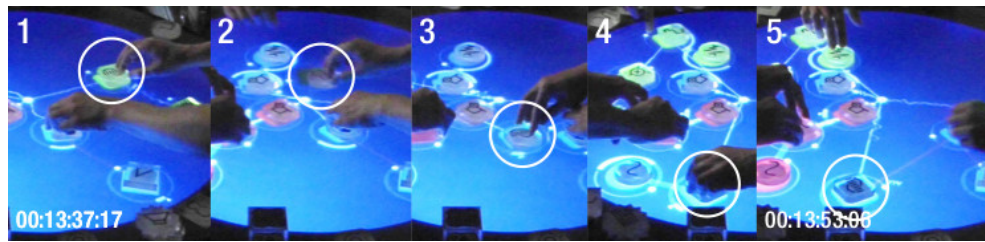
FIGURE 4.7: Triggering unintentional effects with dynamic patching (G1S1)

with other people's threads, or even the entire patch, was not uncommon, especially during early sessions. Sometimes users found it difficult to discern the functionality of an object. For example, see §4.9.2 on error/repair situations and §4.8.1 on misusing objects. If any of these events occurred, the musicians either treated it as serendipitous, integrating it into their ongoing improvisation and building on it, or attempted to repair and revoke it. An example of an unintentional effect of dragging an object in an early session (G1S1) is shown in Fig. 4.7 (see video 4-3A):

> *Vignette 1: Unintentional effects in early sessions* (G1S1). M3 (the musician on the right) has an individual thread of two objects: a slicer effect (FX) and a string bass instrument (SG) (frame 1). He drags the FX towards M2 (the musician at the top of the figure). This action disconnects the FX from his thread, but no connection is established with M2's threads either (frame 2). He moves the FX towards M1's threads (musician on the left of the figure), twice establishing intermittent connections to two different threads belonging to M1 (frames 3–4). Finally, M3 leaves the object in an individual thread of three objects (frame 5).

This vignette lasted 15–16 seconds, and exemplifies unintentional effects with dynamic patching during a first session.

By contrast, Figure 4.8 (frame 3 of Fig. 4.16, bottom row) provides an example of the intentional use of dynamic patching and serendipitous actions at the ending of S4 (see video 4-8B):
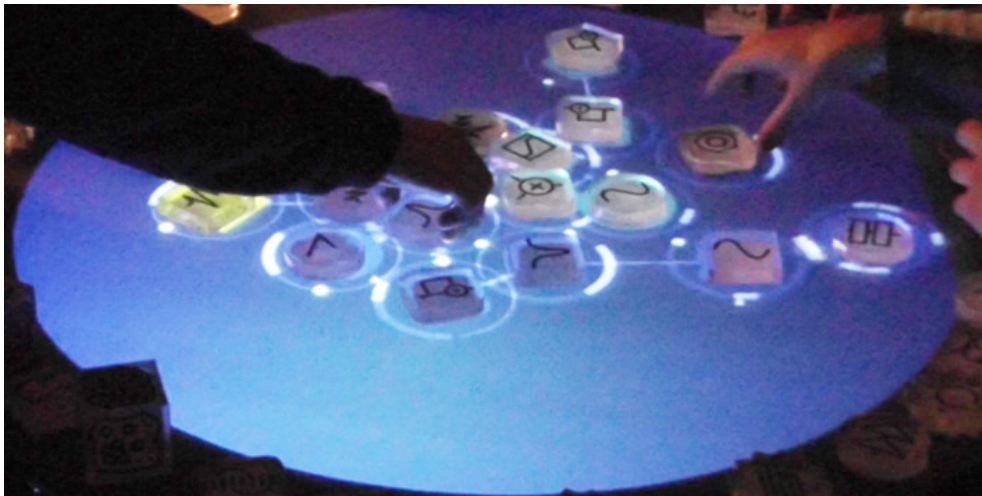
FIGURE 4.8: Triggering intentional effects with dynamic patching (G1S4)

Vignette 2: Intentional effects in last sessions (G1S4). During the ending of S4, there is a shared thread in the centre that musicians playfully contribute to.

This vignette exemplifies how a group in S4 develops a complex ending, in which dynamic patching is used with control.

These examples reveal that dynamic patching of automatic connections is a useful mechanism for promoting hands-on exploration and discovery of the interface in early sessions, and for becoming part of the musical repertoire in later sessions. Dynamic patching together with the lack of territorial divisions of the interface seemed to support exploration and self-regulation of the tabletop space.

4.6.2   Individual territories: flexible reconfigurations

Musicians tended to play within the area nearest to them, confirming a tendency to establish personal areas. The larger the group, the smaller the individual area per person; as shown in Fig. 4.6. These individual areas were reconfigured depending on the number of musicians, as happened with group G3, which grew from three musicians (S1–S2) to four (S3–S4). In general,

musicians finished the session at the same location as they started without switching positions. An exception was the end of session S1 by Gb, where M1 invaded a number of M2's threads, sharing them or taking them over completely; even moving next to M2 to fade the volume of some objects and then remove them. This can be explained as part of the initial exploration of the interface and the available space when playing in duo. In the situations where a participant arrived later (M1Gb in S2 and S3) or left for a short period of time (M2Gb in S3), individual areas were dynamically reconfigured. Yet, if someone momentarily left their position, but remained in the room (e.g. for manual checking or for changing the speakers' volume level), territories did not change.

An example of reconfiguration of individual territory is shown in the vignette in Fig. 4.9 (see video 4-1A). This occurred near the middle of S3 of Gb:

> *Vignette 3: Dynamic reconfigurations* (GbS3). The surface table area was divided into M1 and M2's individual spaces. Suddenly, with no verbal exchange, M2 (musician on the left in frame 1) left the room, leaving his threads playing. Then, M1 (musician on the right in frame 1) interacted with all the threads on the table, fading the volume out of all of M2's threads (frame 2), and even moving to where M2 had been standing (frames 3–4). M1 moved back to his original position. He started two new threads with two cube objects, one of which incorporated an FX from a M2's thread by dynamic patching, and his thread occupied part of M2's space. After approximately two minutes, M2 came back to the room, went to his initial tabletop position and asked "How is mine going?"; M1 replied: "I've faded it out"; and M2 agreed, saying "Okay". Then M1 moved the cube towards himself making the connection with M2's FX disappear.

This example shows how individual spaces and threads could be dynamically reconfigured, depending on the number of musicians in the room.
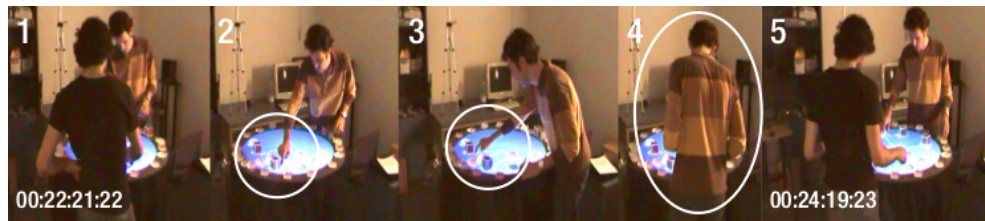
FIGURE 4.9: Dynamic reconfiguration of individual territory

4.6.3 Individual territories: flexible negotiation of takes and gives

Before each session, the facilitator organised the objects in the rim area in no specific order, sometimes stacked in pairs. Only Gb explicitly organised the objects by function and distributed them in the rim area during S1 with the aim of becoming familiar with them, after an initial period of 15–20 minutes of improvisation. Those musicians who loaded a cube object with their own samples tended to have this object close at hand and use it frequently.

Throughout all sessions, musicians tended to play with those objects stored in the rim area nearest to them, although when specific objects were needed, they also took these from the rim area of others' nearest areas or areas in-between musicians: generally these interactions in others' rim area consisted in choosing an object and using it immediately (*passive takes*), without asking for permission. Three musicians of different groups (M1G2, M2G3 and M1Gb) performed passive takes extensively, sometimes leaving the objects again without using them (eventual passive gives). This indicates that the rim area could be used as a shared storage area, where the nearest area to oneself is preferred.

How passive takes evolved over the multiple sessions indicates that the objects and their categories became better known over time. In early sessions, the objects taken generally belonged only to one or two categories: sound generators (G3), sound effects (G1), sound generators and sound effects (G2) or sound generators and global controls (Gb). During the last sessions, objects from all categories were chosen, except for G1 that global controls were not taken. This indicates improved control over the collection of objects.

Passive gives were rare: objects were usually stored, after use, in the nearest rim area to the person, and only sometimes in a fellow musician's rim area when there was lack of space. We rarely observed *active takes*, with some intentional and others unintentional: territorial social protocols of personal spaces and object ownership seem to regulate the use of the surface table area. There were occasional *active gives*: some of these were handovers, which are explained in §4.7.3; others happened when moving threads towards others' areas to create free space within one's personal area. The small number of gives indicates that musicians focused on individual threads.

The data reveals that individual spaces were negotiated flexibly with no need for system level constraints, and that control of the objects collection improves over time.

4.6.4 Shared threads: from unintentional to intentional

Threads tended to be shared by the entire group when they occurred in the spaces between musicians or in the middle of the table. For instance, see Fig. 4.5 for an example of a shared thread at the centre of the table. This indicates the identification of the centre and in-between spaces as shared areas. Shared threads were created either intentionally or unintentionally: during early sessions, unintentionally shared threads that were triggered by dynamic patching were rather common (see §4.6.1), whereas in the later sessions participants had learned to control the system, and shared threads were the result of deliberate actions. For example, shared threads were often used as a resource for beginnings or endings, and their complexity increased in the last sessions, as further shown in Fig. 4.16 (bottom row) with an example of a patch left 'alive' at the centre of the table.

As a group's experience grew, invasions came gradually to be used intentionally for musical effect. Whilst in early sessions they were more often basic interventions into somebody else's thread, or a trial-and-error exploration of effects; during the later sessions the interventions were more complex, using special objects such as the radar trigger.

4.6.5 Shared threads: thread ownership with the radar trigger

The radar trigger object can influence others' threads with no need for physical proximity: the range of the radar can be changed dynamically by moving its slider or moving the object. A representative example of using the radar trigger is shown in Fig. 4.10 (see video 4-2A):

> *Vignette 4: Complex invasion with radar trigger* (G1S4). In S4, M2 puts the radar trigger object next to his personal space. The range of the tempo set with the object covers the whole tabletop surface affecting all the threads (frame 1). M2 rotates the object and thus changes the speed of the radar. He also moves the slider and limits the range to the area next to himself (frame 2). M2 drags the object to the opposite side of the table. The range of the radar trigger is now affecting the personal spaces of M1 and M3 (frame 3). M1 takes now control over the radar trigger, and expands the range to include M2 (frame 4).

This example illustrates a complex invasion of others' threads that was not just related to physical proximity: it depends on the position of the object, but also on the range of the radar. The smooth and swift change of the range of influence in this example indicates the uncertainty of thread ownership when an effect is not related to physical proximity. This vignette raises the question of when an invasion becomes an individual or a shared thread, because the interferences with other objects are more gradual, compared to a standard sound effect with on-off influence. With the radar trigger, there is a radar effect from local to a global range, controlled by moving a projected slider. Another example is shown in Fig. 4.19, where the radar trigger is positioned in the centre of the table, affecting individual threads within its range. In both examples, it can simultaneously affect several threads, creating larger shared threads from smaller shared threads.

In summary, this data shows that apart from individual spaces, there are shared spaces, which seem to be negotiated flexibly with no need for system level constraints. In early sessions, shared threads were the result of exploring the Reactable interface; whilst in later sessions, there was a clearer musical
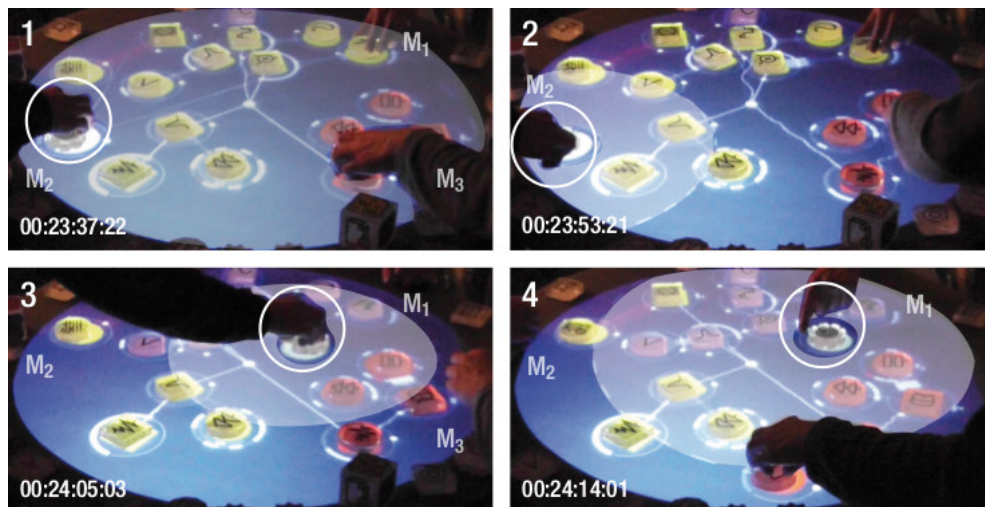
FIGURE 4.10: Complex invasion with the radar trigger object (G1S4). The radar trigger's area of influence appears shaded, and the circle indicates who is operating the object

purpose. It is notable that we only found very few instances in which invasions or sharing of threads were objected to—they tended to be integrated ad-hoc into the musical development.

4.7 Findings: tangible interaction development

In this section we present findings that focus on more generic aspects of physical and tangible interaction. In particular, we describe first how configurations and techniques using tangible objects were built from basic to more complex; second, we discuss the gestural aspects of the interaction, including their performative and communicative value; and, finally, we discuss the way in which tabletop explorations related to the physicality of the tangible objects.

It is notable that:

1. More complex individual and group TUI techniques and configurations with objects were developed over time.

2. There was little overt eye contact and explicit communication, despite the high levels of workspace awareness that developed.

3. Explorations with objects were partly driven by their physical properties irrespective of their musical effects.

4.7.1 Development of TUI control

All groups evolved towards utilising more complex thread configurations and techniques of operating with the tangibles. Initially, groups did not duplicate objects with similar functions within a thread and constructed only basic sequential threads. However, as they gained experience, in later sessions they created more complex patterns—for instance, by using a number of sound effects at the end of a thread, or by connecting objects as a two-branch thread. Also, groups progressed from using single techniques, such as *dragging*, *swapping* or *twisting* objects; to combining two of these techniques simultaneously. Examples of individual and shared complex configurations and techniques are shown in Fig. 4.10 (complex invasion); or 4.19 (frame 3) with use of the radar trigger as a 4-step sequencer.

Usually threads started with a sound generator (SG) followed by or simultaneously used with other objects. Yet groups G1 (M1 in S1–S2, M2 in S3–S4, M3 in S1 and S3) and Gb (M1 in S3, M2 in S1) developed *preview techniques*. Figure 4.11 (which is also frame 2 of Fig. 4.15, bottom row), illustrates a preview technique used by M1 in S2 using two filters (see video 4-7B), after using this technique with one filter in S1:

> *Vignette 5: Preview technique* (G1S2). In S2, M1 starts building a thread with a resonant filter (FX) in the middle of his individual area, which produces no sound (frame 1). Then he adds a second resonant filter in the space between the first FX and the pulsing dot in the middle of the table, and both objects are repositioned closer to the middle with his left hand, while he adds with his right hand a random control between the first FX and himself. The thread remains with no sound effect. He slightly repositions the first FX and the CT to his left (frame 2). Then he

FIGURE 4.11: Preview technique (G1S2)

removes the CT (frame 3). After searching among objects nearest to him in the rim area for over 10 seconds, M1 adds the square wave oscillator (SG), which triggers a filtered sound in a thread of three objects.

This vignette illustrates how in later sessions, there was a growing workspace awareness of individual work. Particularly, groups showed an increasing ability to cope with the Reactable's lack of sound preview. There were no headphones or other alternatives for pre-listening to the sound. So, for example, pre-listening techniques were developed.

Individual actions related to lack of awareness of the individual's own sounds—for example, misusing objects, or trying to identify what object produces what sound by moving an object up and down—seemed to happen more often during early sessions. Some musicians developed special techniques over time (for example, *preview techniques*), or built more complex individual and shared configurations that indicate a greater control of the interface. Therefore there was a greater control of the interface and improved workspace awareness of individual and shared contributions. Interface feedback seemed to help users in the identification and control of their own and others' actions.

4.7.2 Gestures: sound producing gestures

We identified instrumental interactions (for a definition of the term, see §4.4.3) for sound production, including both basic and exaggerated movements, arising from the manipulation of the TTI. Sound-producing gestures were generally performed using hand gestures, one or two-handed, with lack of any movements beyond those necessary to interact with the instrument.

Yet, there were also occasions of exaggerated movements, for instance when performing certain techniques such as strobing (e.g. M2G2 in S3 and S4 or M4G3 in S3). Musicians tended to utilise their whole upper body in this movement, lifting the tangible object high above the surface while moving their upper body in synchrony, emphasising the rhythm. These movements seemed to add bodily emphasis to specific actions, making them more noticeable to other members.

4.7.3   Gestures: ancillary and communicative gestures

We found that musicians played and coordinated, while focusing their attention on the table surface, with little accompanying verbal communication or direct eye contact. Heads and upper bodies tended to be bent forward over the table surface. When searching for objects, musicians tended to focus on their nearest rim area, and, if an object was not found, they would then start to look at other parts of the rim area, with slight turns of the head. In all groups throughout all sessions, there were no collisions of arms or hands, despite a general lack of verbal communication. An example is a handover shown in Fig. 4.12 (see video 4-4A):

> *Vignette 6: Handover* (G1S2). M3, who is playfully exploring effects, drags a resonant filter towards M2 while playing with the serendipity of dynamic patching (frame 1). M3 stops with one finger in a pointing gesture on the object. M2 approaches his hand and takes the object: handover (frame 2). M2 continues the exploration by dragging the resonant filter on the other side of the table. Both are looking down at the table surface without any verbal or overt gestural interaction, or establishment of direct eye contact.

This demonstrates group awareness in instrumental interaction (cf. Hornecker et al. [2008]), which seemed to have been facilitated by the shared visibility of the workspace, which in turn was supported by real-time audiovisual feedback on instrumental actions on the table. Explicit eye contact during verbal exchanges was mostly observed during beginnings or endings,

FIGURE 4.12: Handover (G1S2)

and occasionally in the middle of sessions. Occasional establishment of eye contact can be observed in the video overview when participants lift their heads. Sometimes it was combined with actions that indicated engagement at the end of a session, such as laughter in all groups, including shaking hands in G1S1 or clapping hands in G2S1.

4.7.4 Gestures: sound accompanying gestures

Some groups utilised more sound accompanying gestures than others. Although the extent of this varied between groups, the occurrences indicate a connection between gestural interaction and group dynamics. These gestures were generally full-body movements. In G1, participants remained motionless with only occasional shifts of the upper body towards the table surface when manipulating objects. In G2, all participants appeared to be highly coordinated and engaged, often *bobbing* their heads and occasionally dancing in synch with the music using the whole upper body. Figure 4.13 shows an example of bobbing heads (see video 4-5A):

> *Vignette 7: Bobbing heads* (G2S1). M3 (left) is bobbing his head rhythmically up and down in S1 (frames 1–3), and so does M1 (right), also bobbing his head (frames 3–4).

Bobbing may support the individual musician's sense of rhythm, and also may serve to synchronise the group to this rhythm. In G3, participants occasionally shifted their upper bodies towards the table surface to manipulate objects, and often nodded their heads in time with the music. In Gb, participants remained motionless with occasional shifts of the upper body towards
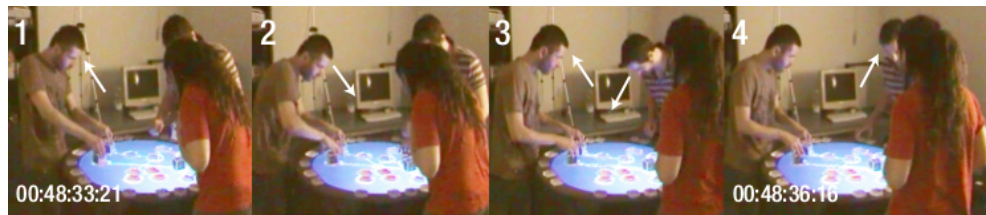
FIGURE 4.13: Bobbing heads (G2S1)

the table surface for object manipulation and also some swaying from side to side. In contrast with traditional ensembles, in which listening and sometimes looking at the other musicians is important to maintain synchronicity and group cohesion, in our study there were a few overt gestures to help synchronise the group: a synched and shared interface and group dynamics seem to support this apparent cohesion of a synched group, irrespective of the lack of visible eye contact.

Overall, a range of coordinated bodily actions and hand gestures were performed with the Reactable, with little direct eye contact and while focusing attention on the tabletop surface. These included exaggerated and non-exaggerated sound producing gestures (e.g. manipulations including gestures such as strobing); ancillary and communicative gestures (e.g. bodily coordinated actions such as handovers); and sound accompanying gestures (e.g. bobbing heads). The use of tangible objects on a shared interface with synchronicity of actions and real-time feedback seems to promote all these types of bodily interaction; the nature of the interface itself makes overt communicative gestures less necessary than with more conventional groups with a set of musical instruments. In the latter, you may find more eye contact, partly because the lack of a shared and centralised view. In the former, communication is mediated by instrumental interaction of a shared interface that everyone can see and operate.
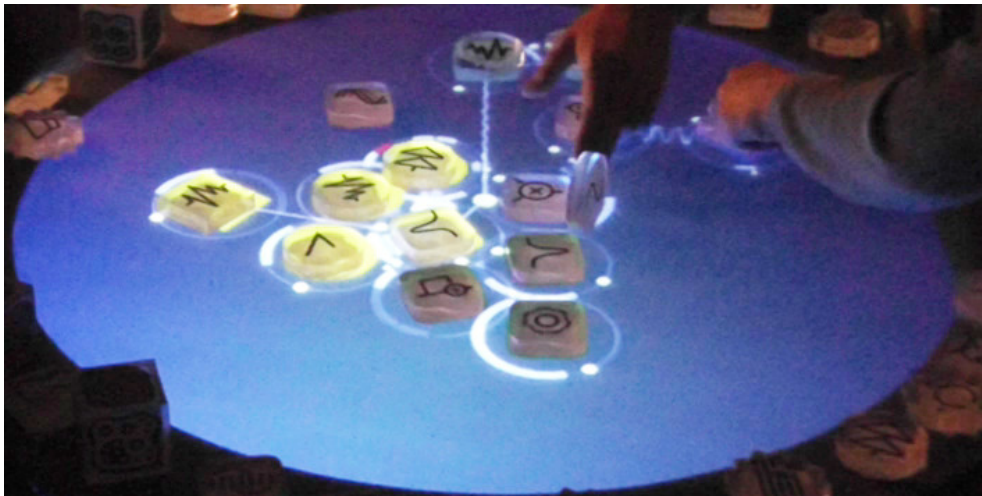
FIGURE 4.14: Physical-digital exploration: tossing (G1S4)

4.7.5 Physical-digital exploration

We noticed several situations where groups explored the limits of the TUI, such as by adding all the available objects to the surface (G3 in S2–S4), stacking objects (M1 and M3 in G3S2), rolling or tossing objects (M1 in G1S4), or adding a mobile phone on the surface as an alternative object—with obviously no sound effect (M3 in G3S4). Figure 4.14 (frame 2 of Fig. 4.16, bottom row), exemplifies an exploratory dialogue between the physical affordances of an object (e.g. rolling a circular object) and the digital interface (see video 4-8B):

> *Vignette 8: Tossing* (G1S4). An example of these physical-digital exploration happens when M1 is making a circular object roll towards the middle of the table, eventually letting it fall down on one side.

An approach of trial and error exploration with the physical tangibles seems to help people discover the digital domain in TTIs, which can, in turn, be seen and imitated by others.

In summary, we discovered that there was a greater control and complexity in creating individual and shared configurations and techniques towards the later sessions, which points to a greater WA of individual and group work

over time. We found that beyond sound producing gestures related to the manipulation of objects (e.g. strobing or twisting), there were few overt ancillary and communicative gestures, because group members tended to focus their attention to the table, and these gestures were mediated by the tabletop interface (e.g. handovers with little overt eye contact). Sound accompanying gestures (e.g. bobbing) varied from group to group. Finally, some explorations with objects were driven by their physical properties (e.g. stacking) not necessarily with any musical effect.

4.8   Findings: musical development

In this section, we focus on the evolution of musical communication within the groups across the sessions. It is beyond the scope of this work to consider the musical style of the groups.

It is notable that:

1. Musicians developed individual voices.

2. Groups developed musical dialogues, intros, and endings with increasing complexity.

3. The exploration and development of the improvisation was different in each group.

4. The nature of tabletop musical improvisation differs from traditional ensembles in terms of roles and communication. These differences, as we argue, are an effect of the shared interface, which enables new forms of musical collaboration.

4.8.1   Individual musical interaction

The video data revealed that issues with workspace awareness mainly arose in the initial sessions, arising from musicians' difficulties with particular features of the interface: for example, misusing the programmer object. This

object can only reprogram the samples of a number of SGs (i.e. the cubes and the instruments), but there is visual feedback when it connects with any other object. On a number of occasions, musicians used this object in the wrong context with no apparent acoustic or visual indication (e.g. G1 and G2).

We observed a number of *solos* in all four groups, which increased in complexity over time, although some groups were more inclined to perform solos (esp. G3 and Gb) than others. Some solos were combined with an existing accompaniment (i.e. a leading voice vs accompaniment voices) and some triggered *musical dialogues*, as detailed in the next section. There were individual *musical explorations*: a subset of these were mimicked and further developed by other peers (see §4.9.1); others were developed individually, such as creating serendipitous threads with numerous objects (M2G3 in S1 and S2), or dragging an isolated object along different threads in an exploratory mode (M3G1 in S1–S4). Observations indicate that, over time, control and workspace awareness improved (i.e. individual work). Occasionally, we observed *non-participation* of musicians standing back and contemplating the patch (e.g. in G1 it happened with musicians M1 in S4, and M2 in S2–S4; or in G3 it happened with musicians M2 in S1–S4, M3 in S1–S4, and M4 in S4).

This data reveals how musicians developed different individual voices within the groups, including solos (similar to jazz ensembles); individual explorations of the interface; and positive non-participation e.g. leaving their patches 'alive' and playing.

4.8.2   Group musical interaction: musical dialogues

Generally, we found equal participation and no evidence of fixed musical roles. The configuration of the objects and the resulting musical output changed constantly. The distribution of roles was dynamic and tended to arise non-verbally: we only found an occasional explicit distribution of musical roles with Gb as melody vs rhythm. In general, there were visible leading

voices in all groups, with either a soloist and accompanying voices emerging, or a leading melodic voice with secondary melodic and rhythmic voices. These voices were dynamically exchanged during musical dialogues, or during intros and endings, as detailed below.

Jazz musicians describe musical improvisation as a conversation between two or more musicians, mediated by open-ended turn-taking, based on a relatively fixed rhythm section and a variable soloist, with a dynamic tension between the two [Monson, 1996, pp. 82–83]. We identified a number of *musical dialogues* between musicians, not necessarily divided into rhythmic accompaniment and a soloist section. Over the four sessions, these dialogues became more complex, using more variations in tempo and a greater use of heterophonic voices. Table 4.3 provides, for each group, one example of a *basic dialogue*, which tended to happen in early sessions; contrasted with one of a *complex dialogue*, which tended to happen in later sessions.

Overall, this data reveals that musicians' participation was egalitarian and roles were flexibly changed, which contrasts with traditional ensembles where roles are more fixed (cf. [Monson, 1996]).

4.8.3 Group musical interaction: intros vs endings

Given the time constraints of a musical improvisation session and the protocols of improvisation, participants had to coordinate *intros* and *endings*. In later sessions, musicians tended to focus more on their individual voices during intros, using more complex configurations, although there were also invasions and shared threads in spaces. For endings, we found that musicians in most groups tended to share voices in a more complex way than in early sessions (G1, G3 and Gb), frequently using the middle of the table. Figure 4.15 illustrates the first (top row) and last (bottom row) intro of G1 (see videos 4-7A, 4-7B); whereas Figure 4.16 illustrates the first (top row) and last (bottom row) ending of G1 (see videos 4-8A, 4-8B). Both figures indicate how musically complex intros and endings could become. For example, in intros a development of greater complexity is shown:

TABLE 4.3: Comparative examples of basic vs complex dialogues

Groups	Basic dialogue	Complex dialogue
G1	In S1 (see video 4-6A) sequential call-response occurred with two textured, melodic voices (M2, M3) and one fixed rhythm voice (M1). There was a lack of role change or variations in tempo. The three voices were clearly audible as separate.	In S4 (see video 4-6B) the leading melodic voice was transferred from M2 to M1 (dynamic role change). The other two voices (M2, M3), added counterpoint to the leading melody as call-response with rhythmic melodies. There were several variations in tempo. The three voices intertwined with one another.
G2	In S2 (see video 4-6C) there were two simultaneous leading melodic voices (i.e. M2 as high pitch voice, M3 as mid-pitch voice), and a bass/rhythmic voice (M1). There was lack of gradual changes in volume. The three voices added counterpoint to one another.	In S3 (see video 4-6D) there was a crescendo or a gradual change in volume with one leading mid-pitch melodic voice (M2), one secondary high pitch melodic voice (M1), and a subtle bass/rhythmic voice (M3). M1 and M3 added counterpoint to M2.
G3	In S1 (see video 4-6E) sequential call-response was observed with one call high pitch melodic voice (M3) and two response melodic voices (i.e. M1 as mid-high pitch and M2 as mid-low pitch). All voices combined melodic exploration with control (e.g. SGs were turned left or right yielding multiplicity of tones or rhythms, which not always combined harmonically and rhythmically).	In S3 (see video 4-6F) there were two simultaneous melodic voices – one high pitch (M1) and one low pitch (M2), which combined harmonically and rhythmically.
Gb	In S2 (see video 4-6G) two simultaneous high pitch melodic voices (M1 and M2) were combined with a homophonic bass/rhythmic voice (also by M2). None of the three voices combined either harmonically or rhythmically (e.g. SGs were turned randomly left or right yielding multiplicity of tones or rhythms).	In S4 (see video 4-6H) the leading melodic voice was transferred from M2 to M1 (dynamic role change): first, simultaneously one high pitch melodic leading voice (M2) was combined with a low pitch melodic voice (M1); then, simultaneously one high pitch melodic leading voice (M1) was combined with a low pitch melodic voice (M2). All pairs of voices combined harmonically and rhythmically.

Vignette 9: Intro (G1S1 and G1S4). In S1, M1 adds a first SG (frame 1). M3 adds and rotates a second SG whilst M1 changes the slider of the first SG (frame 2). M2 adds and rotates a third SG whilst M1 adds a controller to the first SG (frame 3). M2 adds a filter to the third SG while rotating both objects (frame 4). In S4, M3 adds a cube and a programmer to select a sample and M2 adds a SG (frame 1). M2 adds a second SG and rotates one with each hand while M1 adds a FX (preview technique) (frame 2). M3 reprograms the cube, M2 removes the second SG and M1 adds another FX and a CT (preview technique) (frame 3). M1 removes the CT and adds a SG for the thread to sound (frame 4).

In this vignette, a single SG per thread is used with immediate sound output in the first session; whereas more objects per thread are used in the last session. In this last session, the programmer is used for reprogramming the
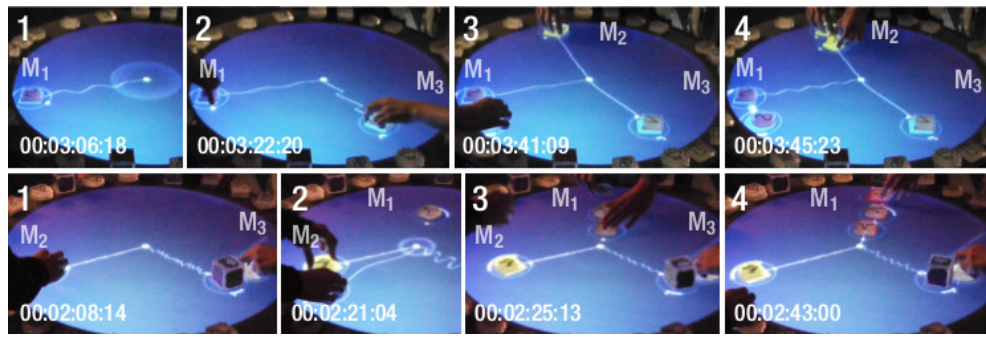
FIGURE 4.15: Basic (top) vs complex (bottom) intro (G1S1, G1S4)

samples of a cube object. Moreover, the preview technique is used for controlling when to trigger the sound output, which shows greater control and workspace awareness.

In endings, the development of greater complexity and workspace awareness is shown, as well:

> *Vignette 10: Ending* (G1S1 and G1S4). In S1, each musician is in charge of one thread (frame 1). They respectively start removing their threads (frame 2). They keep removing their threads (frame 3)... generating silence (frame 4). In S4, there is a shared thread in the centre that musicians playfully contribute to (frames 1-3)... until the patch is left 'alive' (frame 4).

In this example, there is a sequential removal of objects from individual threads in the first session (top row). By contrast, there are voluntary serendipitous contributions to a large shared thread on the centre of the table in the final one session (bottom row).

4.8.4 Group musical interaction: endings

We noted that groups developed a wide variety of endings. They included various combinations of the same set of elements, illustrating how group dynamics and musical practices may differ even in a small set of groups (see Table 4.4).
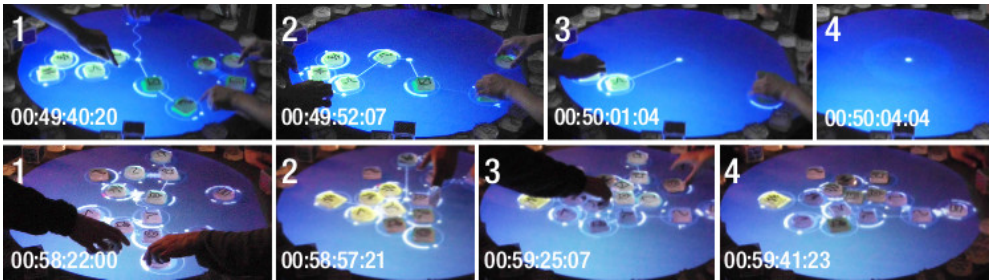
FIGURE 4.16: Basic (top) vs complex (bottom) ending (G1S1, G1S4)

TABLE 4.4: Overview of endings. The icons within the circle represent the different types of ending techniques that were utilised. Filled icons indicate the actual use of this technique in a session

Group	S1	S2	S3	S4

▶ Fade out
⬠ Global object
■ Sequential remove
⬡ Serendipity
★ Shared threads

In summary, we found recurrent elements such as objects removed sequentially; fades of volume; use of global objects; shared threads; and serendipitous actions. The need for coordination in closing each session, as is traditional in popular and jazz music, seems to explain why shared threads are often found in the endings. We observed how markedly different endings could be developed from a relatively small set of elements (see icons in Table 4.4) with combinations varying from session to session within the same group: G1 combined a small number of elements in every ending, with a wide range of

variation, which resulted in combinations that varied from session to session, and which tended to include serendipity. G3 also included serendipitous actions in the endings, but this group combined a greater number of elements instead, with less combinatorial range. Finally, G2 and Gb groups lacked serendipitous actions in the endings; they also combined a small number of elements in every ending with a wide range of variation. This data shows that groups developed their own collective, distinctive sequence of musical events using a variety of combinations of objects.

To summarise, groups developed both individual and group musical interaction over time using different approaches, which points to a situated group tabletop interaction; and a greater control of workspace awareness of individual vs group musical work. Observations also showed that tabletops promoted a more dynamic exchange of roles between musicians during musical improvisation compared to traditional jazz ensembles with conventional musical instruments.

4.9   Findings: peer learning

In this section, we detail the findings regarding social peer interaction in terms of mimicking behaviours and verbal communication.

It is notable that:

1. Mimicking was used as an implicit form of peer learning based on active imitation.

2. Verbal communication was used as an explicit form of peer learning based on sharing knowledge.

4.9.1   Implicit peer learning: mimicking

Mimicking occurred in all groups and throughout all sessions, without explicit talk. We noted that the musicians generally tended not to look up at

each other, and seemed to rely on peripheral vision even when actions were imitated close together in time. However, during active manipulation and problem solving, even if engaged in intense discussion, participants looked onto the shared workspace, and combined this with glances at each other: for example, in response to a joke or an observation.

Generally, in early sessions, musicians tended to mimic basic configurations or techniques: for example, muting or unmuting a thread (G1S1); strobing on and off whilst twisting left and right a filter positioned at the end of the thread (G2S1); exploring and dragging an object on different threads (G3S1); or stacking objects of the same shape (G3S2). By contrast, during the later sessions, musicians tended to mimic more complex configurations or techniques, e.g. building copies or close copies of complex configurations such as the configuration of a CT connected to an SG connected in sequence to two FXs (G1S3); or the techniques of twisting an FX affecting an SG (G2S3), operating two SGs simultaneously (G3S3), or shaking two FXs between two threads (G3S3).

We identified repetitions of different techniques among groups and sessions. A representative example of how an idea was initially used by one musician, and then repeated and reshaped by the rest of the group, is represented in a timeline, as shown in Fig. 4.17. Figure 4.18 illustrates the use of the technique by each of the three participants throughout the sessions (see videos 4-9A, 4-9B, 4-9C):

> *Vignette 11: Mimicking 'strobing and twisting' on the centre* (G2 S1–S4). In S1, M2 operates distortion effect in the centre using strobing and twisting (frame 1). In S3, M3 operates global feedback in the centre using strobing, twisting, and dragging (frame 2). In S4, M1 operates granulator effect in the centre, first twisting it, and then leaving it in the centre for ca. 80 sec (frame 3).

	S₁	S₂	S₃	S₄	
M₁G₂				◆ ◆	◇ Original
M₂G₂	◇	◆◆◆ ◆	◆ ◆	◆ ◆◆◆◆	◆ Repetition
M₃G₂			◆◆	◆	
	30 Time (min.)	30 Time (min.)	30 Time (min.)	30 Time (min.)	
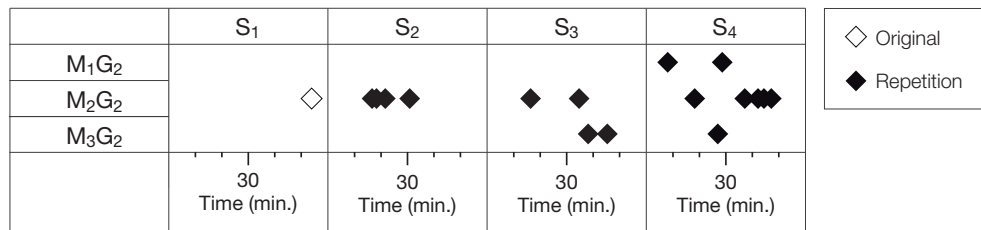
FIGURE 4.17: Timeline visualisation of mimicking one technique over time
(G2: S1-S4)
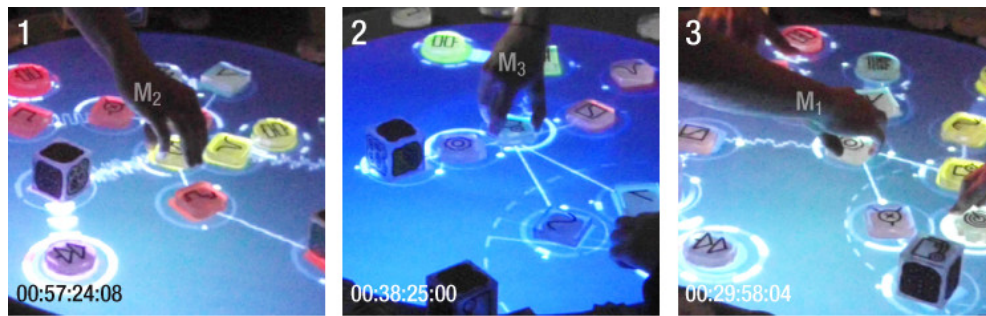
FIGURE 4.18: Mimicking (G2: S1, S3, S4)

Mimicking was widely used and constituted a form of implicit peer learning: it allowed musicians to share techniques and practices between them by active imitation. What specific techniques and practices were mimicked varied depending on the different group dynamics and situated contexts.

4.9.2 Explicit peer learning: verbal communication

Conversations happened mostly at the beginnings and endings of the sessions, only occasionally in the middle, and was used for comments, dialogues, sharing expressions of satisfaction, or learning. As another form of peer learning, verbal communication was primarily used to share knowledge explicitly. The most common events in which conversation and peer learning took place were error/repair situations, discoveries, question/answer situations, think alouds, or group discussions.

A representative *error/repair situation*, which happened in all groups, occurred during the use of the global volume object. This can be used to control the

global volume of the Reactable by rotating the object, which also has a slider, which defines the amount of reverb or echo. However, the reverb feature is not explained in the user manual. On several occasions, by just rotating the object, the reverb was increased unexpectedly to its maximum value. In all groups it took a certain time, either in the same session or in the following sessions, to understand the behaviour of this object. In general, this issue was resolved by an individual (e.g. by trial and error), who then shared the knowledge with the group through explicit communication. We here give one example:

> *Vignette 12: Peer learning from error/repair situation with global volume* (G1S3). The issue arose and was repaired in S3. M2 started the reverb situation, and immediately repaired it by moving the slider. M2 used it again intentionally. After the ending, M2 shared his discovery by having a group discussion.

There were also examples of verbal communication in particular group situations. As shown in Fig. 4.19, in G3S2, the group of three discovered how the radar trigger object worked as a metronome set to 4/4 (see video 4-10A). The group discovered the object's possibilities through discussion and experimentation. The group remembered and re-used this technique in the following sessions S3 and S4 (see videos 4-10B, 4-10C) where another participant joined them, using more objects and threads.

> *Vignette 13: Peer learning from showing "the sync" to a newcomer* (G3: S2-S4). In S2, there is a discovery of using the radar trigger object as a 4-step sequencer (frame 1). In S3, there is a reuse of the radar trigger technique as a 4-step sequencer (frame 2). In S4, again, there is a reuse of the radar trigger technique (frame 3). The technique was named "the sync" in S3 by M3 ("Let's try to do the sync"). The new member was instructed by doing, and with explicit explanations. For example, M3 explained to M4: "This is like a four-step sequencer, this is time 1, 2, 3, 4" pointing to the four quarters (see video 4-10B). In S4 the technique was repeated twice without the need of naming it, as part of the group repertoire: one at 00:15:42:23 (see video 4-10C), and the other at 00:42:37:21.
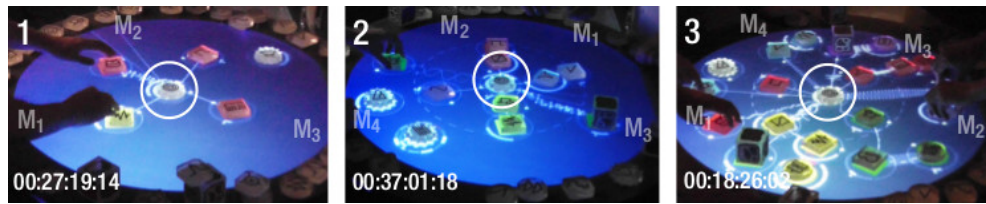
FIGURE 4.19: Explicit peer learning from showing "the sync" (G3: S2-S4)

The explicit transmission of knowledge by means of verbal communication varied according to the dynamics of each group.

4.10 Discussion

In this section, we first discuss the findings in the light of our research questions about group tabletop interaction and workspace awareness. We summarise the implications for design in terms of the lessons learned, and, finally, we highlight the study's limitations.

4.10.1 Reactable's interface

From exploratory behaviours to musical purpose with dynamic patching. Our findings suggested that the Reactable's dynamic patching of automatic connections promoted serendipity and creative discovery, a mechanism that seems useful for creative domains, especially during early exploratory phases. Such findings are in agreement with recent research on collaborative learning and interferences [Pontual Falcão & Price, 2010], where unintentional interferences are found to promote exploratory learning in tabletop talkative contexts. Furthermore, our findings indicated that groups controlled dynamic patching over time, and included it as part of their musical repertoire (e.g. intentional effects using dynamic patching). With this function the Reactable eases access for beginners and casual users, yet it is not possible to disable the mechanism for full manual control of the connections.

The latter would be useful for advanced users, who tend to prefer having full control of the interface [Blaine & Fels, 2003].

Personal, shared and storage areas. The observed division of the Reactable's interface into personal and shared spaces concurs with the territorial behaviours found by Scott & Carpendale [2010]; Scott et al. [2004] in conventional tabletop collaboration. Our finding that musicians tend to play with the objects nearest to themselves matches the authors' observation that the spatial location of resources influences the perceived ownership. The storage territory of the rim area had a shared use, sharing that happens, as described by Scott & Carpendale [2010], when the storage territory is located in the group territory. Accordingly, an open question is whether there would be more reservation (i.e. less exchange of objects) and organisation of resources with auxiliary storage territories.

Lack of divisions promotes dynamic self-organisation of space. Our data indicates that territorial constraints might be harmful for group dynamics (cf. Scott & Carpendale [2010]; Scott et al. [2004]) during tabletop free improvisation activities, as it might potentially interfere with the early exploratory behaviour that musicians utilised to understand the effects of different objects and manipulations. In addition, once the initial phase of exploration was completed, invasions of other musicians' supposed territory or threads tended to have a clear musical purpose and were not objected to by the other group members: indeed, they were seen as useful contributions. This concurs with Hornecker et al. [2008] suggestion that allowing interventions into others' spaces can promote a rich range of opportunities for collaborative work, group dynamics, and expertise development. Moreover, as Q. Wang et al. [2006] suggest, having no ownership markers might support participants in feeling part of a group.

Fluid transitions between individual and group work. Objects with different levels of thread influence, ranging from local to global, seem to encourage a multiplicity of interventions into others musicians' spaces and more fluid transitions between individual and shared work (e.g. complex invasions). Sharing the same interface and easily manipulating others' threads contrasts

with traditional ensembles where the musicians are usually placed some distance apart from other musicians and instruments, which in itself operates as a territorial indicator: each musician has an individual musical instrument and musicians rarely play others' musical instruments.

4.10.2 Tangible interaction

Modular TUI systems: how much complexity? Our observations indicated that tangible building blocks and a higher WA allowed users to build from basic to complex configurations. We assume that musicians expect to build more complex structures as they master the TTI, potentially using more tangible objects. Nonetheless, a large collection of building blocks that have a one-to-one effect on the system might become difficult to manage on a table-top surface due to its space limitation and increasing complexity. This reduces the ultimate level of complexity compared to other modular systems with fewer restrictions on space (e.g. physical modular synthesizers, LEGO bricks or virtual simulations of modular structures). An open question is how to manage this modular approach within the space restriction of a table-top surface, which also limits the potential level of complexity. For example, the Reactable's circular shape limits the space available, and only operates with two-dimensionally positioned objects. These limitations might begin to be noticed after several days of practice. Having said that, restrictive forms may benefit the creative process, as discussed by Magnusson [2010], thus unlimited systems may inhibit musical creativity because they can lead to overwhelming complexity.

Tangible interaction: simultaneous communication and manipulation gestures. We observed how the use of tangible objects supported deictic communication during a handover: a communicative gesture performed while manipulating an object. This characteristic of tangible interaction contrasts with purely touch-based interaction, which often suffers from the difficulty of distinguishing a communicative gesture from a manipulative gesture. As

the literature on the Reactable's development highlights [Jordà, 2008], tangible manipulation, when compared to touch-based interaction, can free musicians' visual attention to focus, for example, on other objects, or watch what their peers are doing on the table surface, while still retaining control over the tangible held.

***Synched hand and bodily performativity with lack of eye contact*.** Body and hand movements are characteristic of tangible and physical interaction. When using the Reactable, groups performed a range of bodily actions and hand gestures with generally little eye contact and intense focus on the tabletop surface. A similar lack of eye contact has also been found in other studies of group tabletop interaction [Hornecker et al., 2008], echoing early research on video conferencing systems, which revealed that visibility of the workspace is often more important for awareness and collaboration than that of a 'talking heads' [Nardi et al., 1995]. The bodily actions coordinated between users we observed (e.g. handovers or bobbing heads in synch) seem to be promoted by interface mechanisms such as synched multi-threading (i.e. using the same system time clock for all tangibles' connections); as well as by having a shared interface with real-time feedback. This synchronicity of input and output is suggested as a design feature in TUIs that enables users for exploring the effects of different interface actions [Zuckerman et al., 2005]. Our study found that the Reactable encouraged coordinated hand gestures and certain bodily interaction (e.g. head bobbing), facilitated by playing music while standing and facing the other musicians. By contrast, with touch-based mobile devices, there is less need to interact with the body, and interactions are on a smaller scale; thus the joint action is of a different character and appears much more introverted (cf. Swift [2012]). This chimes with Jordà [2008]'s description of one of the Reactable's design goals being to increase *performativity* in electronic music making, recalling how traditional musical instruments are performed with a greater use of the body.

Basic or complex mappings between the physical and the digital domains in TTIs? The reported physical explorations with the objects and their relations

(e.g. stacking or tossing objects) seem to be a vehicle to explore the mappings between the physical and the digital domains. A dialogue between physical explorations and real-time feedback supports discovering hidden connections between the physical and the digital. As suggested by Ishii et al. [2012], we are still in the infancy of TUIs, and many of the possible manipulations of physical objects have no corresponding output in the system. For example, the Reactable only registers two dimensions, and so stacking objects has no effect on the output. A number of challenges and limitations arise from these physical explorations. First, the Reactable's tangibles operate as controllers where the output is projected on them as visual feedback, but the tangibles lack dynamic embedded information, unlike other tangible systems such as Siftables [Merrill et al., 2007]. An open question is to what extent having computationally embedded tangibles would reduce this trial and error dialogue. Second, the Reactable limits the number of parameters, and the number of inputs and outputs. Yet, in the broader unit generator paradigm [Roads, 1996, pp. 787–788], a unit generator can include multiple parameters and inputs/outputs. A finer-grained representation of the unit generator paradigm on a TUI is a challenge. Newer Reactable's versions such as the Reactable Live or the Reactable Mobile app[3] include configuration settings and additional parameters that support more complex mappings. How to compromise between beginners and experts' expectations from constrained to complex systems is an open question.

4.10.3 Tabletop musical improvisation

The need to support workspace awareness of musical actions in early sessions: multimodal feedback? In the early sessions, while participants were still familiarising themselves with the interface, there was sometimes confusion as to who was doing what, including each participant's own actions. By contrast, the preview technique shows how an understanding of control of individual actions is developed over time. Blaine & Fels [2003] propose

---

[3]`www.reactable.com` (accessed 30 September 2014).

providing multimodal inputs and outputs in collaborative interfaces, to reinforce individual actions. Zuckerman et al. [2005] suggest using multimodal representations with TUIs to engage different senses (e.g. touch, vision, audition) and support different learning styles. Both approaches are based on the identification of actions through sense-based perception. Thus, a possible solution to cope with awareness issues in early sessions might be to provide multimodal feedback, which seems to support different individual and group dynamics [Zuckerman et al., 2005]. Yet, we need to be careful about how to best provide multimodal feedback, as inexperienced users can get confused. Our approach in Chapter 6 explores supporting workspace awareness with additional auditory feedback.

Interconnected, interdependent, dynamic and egalitarian musical roles with a shared interface. The tabletop interface, the free-form nature of musical improvisation and the characteristics of the Reactable's interface seem to promote egalitarian participation with dynamic role changes, in contrast with more hierarchical structures (e.g. soloist vs accompaniment). This can be related to the interdependence found in an *interconnected musical network*, a term coined by Weinberg [2005, p. 31] that was discussed in §2.2.2 used to describe networks "that allow players to influence, share, and shape each other's music in real-time" and that should be "interdependent, dynamic, and function as facilitator of social interactions". Although this definition focuses on networked systems, in our study we found interdependent relationships arising from the shared interface: the level of interdependency is high compared with traditional ensembles such as jazz combos because musicians play a shared instrument rather than individual, independent instruments. This suggests new perspectives and roles in the practice of improvisation, similar to those discussed in network music (§2.2.2). Thus, while in jazz ensembles musical improvisation tends to be a self-reflective process because musicians play individual instruments [Sawyer, 2003], the Reactable provides opportunities for musical knowledge to be transmitted in a *collective reflective process* through a shared interface with real-time visual feedback.

Sharing a centralised interface also has drawbacks: fewer independent individual voices are possible, and the musicians face one another instead of the audience.

Musical improvisation in TTIs: no need for predefined hierarchical roles. Our data showed how musical tabletop design promotes equal participation, and dynamic and versatile roles, in agreement with Hornecker et al. [2008]. This contrasts with the idea of a need to define roles in collaborative music, when sharing the same digital media [Brown, 2006]. A number of dialogues in our study were reminiscent of those occurring in traditional musical ensembles, in particular jazz ensembles, where there is a distinction between rhythmic accompaniment and melodic soloist roles [Monson, 1996]. However, variations in tempo were easily executed with the Reactable (e.g. the metronome object controls the global tempo), and changes of voices and musical roles tended to be fast-paced, features that seem to be particular to this musical tabletop interface.

Positive non-participation in TTIs. Non-participation is traditionally associated with passivity, whereas here it can constitute an active role. Non-participation may occur for several different reasons. First, the limited number of resources available, the size and shape of the tabletop surface, and the number of collaborators seem to influence the performance considerably. Second, with the Reactable it is not necessary to physically play all the time in order to produce sound. Third, silence can be also considered a contribution in music performance [Cage, 1961]. Thus, non-participation (or rather: bodily inactiveness) can be considered a positive aspect of both conventional and tabletop musical improvisation. Yet, in tabletop musical improvisation, non-participation does not necessarily mean not playing. This change of role of the performer was already envisioned by Chadabe with the first interactive systems for computer music in the 1980s: "The performer is simultaneously a participant and an overviewer who functions interactively within the system, devising strategies for using the unexpected to advantage, and also functions outside the system by supplying a more global perspective and guiding the

progress of the music as a whole" [Chadabe, 1984, pp. 26-27]. Active non-participation in musical tabletops contrasts with tabletop studies in which activity is measured according to activity plots [Nacenta et al., 2010; Scott & Carpendale, 2010].

4.10.4 Situated peer learning

Mimicking as a form of peer learning in TTIs. Mimicking is a common practice among improvising musicians. As Bailey [1980/1993] suggests, improvisation is based on imitation, repetition and exploration. Mimicking is greatly facilitated by TTIs in general, and the Reactable in particular, because participants are co-located and face-to-face with the same interface, with no disparity in the tangible objects available to them. Thus, interactive tabletops make collaborators' bodily actions clearly visible over the tabletop surface, and these can be seen and imitated immediately. Participants can easily reproduce existing configurations or gestures, and with the Reactable, this reproducibility is based on a clear mapping between building blocks and sounds.

Construction of shared meanings in the course of action. Our study groups developed ensemble skills such as solving problems in teams (e.g. error/repair situation) and sharing limited resources (a collection of tangible objects). This is in line with the findings of Harris et al. [2009]; Rick et al. [2011], who present interactive tabletops as suitable environments for collaborative learning. Our observations of group development and peer learning also relate to *social constructivism* and the role of peer interaction [Vigotsky, 1978], which views learning as a collaborative process in a meaningful social context: here peers construct new meanings together in particular contexts. For example, the technique "the sync" was co-invented and co-developed by group G3. The learning process seems to be similar to that of *situated learning*, where knowledge is shared and co-constructed within a context and *community of practice* (CoP), understood as a group that shares an activity [Lave & Wenger, 1991]. Yet this literature has typically focused on examples of beginners learning from experts. Our participants learned from each other, and, only in

the case of G3 was a newcomer instructed in later sessions on certain techniques. As with other TUIs in education, *situated peer learning* can happen through hands-on experience [Zuckerman et al., 2005]. Our findings show how group dynamics and tabletop design can influence situated peer learning. This learning happens, in agreement with Suchman [1987]'s findings, in a *situated action*: action and knowledge are interrelated, and collaborative learning happens in the course of action. Our findings also relate to those of Rick et al. [2011] who discuss the benefits of supporting different group dynamics when learning using interactive tabletops. Furthermore, theoretical accounts of collaborative learning by doing, such as Roschelle [1992]'s notion of *convergent conceptual change* are relevant here. In such learning, people gradually construct a shared, convergent meaning, which is situated [Suchman, 1987]—the construction of shared meanings depending on the actors, the context, and the technology used. Repairing the reverb situation when using the global volume exemplifies this: understanding the behaviour of this object was co-constructed in collaboration with peers, and even though it happened in all groups, each group solved it differently in their own meaningful context.

4.10.5 Implications for design

We derive the following set of design considerations from the previous discussion: Reactable interface (1–2), tangible interaction (3–5), tabletop musical improvisation (6–7), and situated peer learning (8). These design considerations are aimed at exploring how to better support long-term collaboration on a TTI for music performance, in particular musical improvisation:

1. *Allow self-regulation of space*. In this study, the lack of territories seemed to promote a self-regulation of space, which was beneficial for musical improvisation. The nature of this self-regulation appeared to be influenced by the available space, the number of group members, the available number of tangibles and their relations (e.g. from local

to global influence), and the musicians' increasing mastery of the interface.

2. ***Automated connection mechanisms can support creativity and learning in early stages***. Creativity can be computationally enhanced with automated system behaviours, yet we should rethink how to complement them with an optional mechanism to control them. An interactive tabletop with a mechanism that automatically connects the tangibles' inputs and outputs (e.g. Reactable's dynamic patching) seems to be useful to promote serendipitous actions and creative discovery. Yet it can also be a constraint when users want more control over connections. In this case, an additional mechanism to control them manually seems appropriate.

3. ***Consider a balance between a complex and a restrictive modular TUI***. Our findings show that a modular tangible interface can allow for both simplicity and combinatorial complexity, supporting multiple compositional possibilities. Yet a future challenge is the scalability of a modular set of tangibles on a tabletop interface that can only be built within a limited physical space. Another challenge is to decide the limit of complexity, given the potential cognitive load as systems increase in complexity.

4. ***Provide synchronicity in actions with objects and space for bodily actions***. The Reactable's synchronicity mechanisms (e.g. global tempo clock, real-time feedback) promoted coordinated bodily actions (e.g. bobbing heads, handovers) between users. The nature of these embodied gestures fits well with collaborative learning during hands-on activities, because bodily actions and gestures can be seen and mimicked, and practical knowledge can be transmitted by doing. These gestures require sufficient space to be performed.

5. ***Allow real-world, multi-dimensional interactions, which have correlations with the digital output***. We are still in the infancy of understanding mappings between the physical and the digital domains in

TUIs. The digital domain in TTIs, usually represented by real-time feedback, is often discovered via a trial and error exploration of the physical properties of the tangibles and their relations (e.g. stacking or tossing objects), which can then be easily reproduced by others. The kinds of actions that users perform in this physical exploration may inform future tabletop design by suggesting actions that could be digitally interpreted, going beyond the 2D mappings which are currently implemented with the Reactable.

6. ***Provide real-time multimodal feedback of synched musical actions***. Transferring traditional creative group activities such as improvisation onto TTIs presents new challenges to group collaboration (e.g. awareness issues). Providing multimodal feedback, related to touch, visual or audition can mitigate awareness issues and facilitate different learning styles, an area for future research.

7. ***Allow for flexible approaches to participation***. Permitting flexibility in roles and participation (e.g. active non-participation, egalitarian participation) without disrupting the activity seems relevant here, in contrast to more fixed and hierarchical roles in traditional improvisation. This is possible because tabletop actions are automatically in sync, and there is a shared interface in which instruments (tangible objects) are easily exchangeable.

8. ***Allow for a variety of group dynamics and development in hands-on situations***. A modular, flexible and shared environment based on constructive building blocks seems to be suitable to support varied group dynamics and situated peer learning. This approach promotes different group development processes and hands-on learning styles that can be mimicked and shared, including different problem solving styles and different levels of compositional complexity. It also places no constraints on the number of users and their positions; participants can join or leave the activity at any time.

The lessons listed above include a range of future directions and explorations for design. Our study further demonstrates the utility of longitudinal studies and shows that studying musical improvisation on interactive tabletops can inform tabletop research and provide insights on tabletop musical practice as well as group coordination and development over time.

4.10.6 Study limitations and future work

We below discuss study limitations and related future work beyond the potential areas of research discussed in the previous section (§4.10.5):

A lab study. A potential criticism of our study may be that we studied improvisation practices in a lab instead of in natural settings. But as Rooksby [2013] argues, social analysis does not necessitate fieldwork (e.g. Suchman [1987]'s famous study was conducted in a lab setting), and lab-based studies can provide an adequate setting for observing situated action. The musicians in our study relied largely on the same resources for establishing awareness, coordination, and for learning as they would do in other settings. Moreover, we attempted to achieve adequate ecological validity by: 1) using a casual and relaxed set-up; and 2) working with expert musicians, who are familiar with musical improvisation, and are used to playing (and engaging) for long periods in group work, even with strangers. As we discussed earlier, the musicians' behaviour (e.g. leaving the room to take phone calls, arriving late for sessions, turning the speakers' volume level up or down at will) indicates that we were successful in this. Furthermore, we studied musical improvisation, a free-form activity that includes known protocols (e.g. beginnings, endings, dialogues) that allow musicians to play with other musicians without necessarily knowing them. A similar in-the-wild study would be very difficult to conduct: it is unlikely that we would be able to access several bands having the same musical tabletop instrument in their studios. In addition, with growing expertise, learning and development processes slow down and become harder to observe. We would thus need to observe groups

at increasingly longer intervals. Here, we have therefore focused on the initial phase of learning to jam together on the Reactable.

Sample size. A second limitation could potentially be that we only studied four groups: for this reason, we have made a detailed qualitative analysis. However, we believe that the developments we have observed are typical of the user group (i.e. expert musicians) chosen for the study, with consistent behavioural patterns across all four groups. At the same time, the observed differences in group dynamics among the groups indicate the variety that might be expected if more groups were to be studied.

Users' gender. Finally, a third potential limitation of our study may be that groups had only male participants. We conducted this study in a predominantly male department in Music and Technology, and all participants volunteered as people interested in improvising electronic music on the Reactable for multiple sessions. As the study was very demanding, we were careful not to coerce people to participate.

Future work. Future work with female, mixed, and novice groups seems promising: female groups, because, with a few exceptions [Rodgers, 2010], there is little research on collaborative music between expert women in electronic music; mixed groups because musical tabletops are a promising platform for collaborative music making beyond traditional combos; and novice groups because musical tabletops offer an easy-to-use interface for them to create music. In the three cases, it would be interesting to see whether group tabletop interactions vary beyond the expected changes, from group dynamics observed in this study including the Reactable's interface use, tangible interaction, musical interaction, and social peer interaction. Additional open questions include how to best support workspace awareness in early sessions with a musical tabletop.

4.11   Summary

In this chapter, we investigated collaborative learning and group development of expertise during musical improvisation on the Reactable by groups of expert musicians. We examined the challenges and opportunities presented by the provision of a set of constructive building blocks on a tabletop tangible interface over time. We observed group tabletop interaction and workspace awareness informed by interface, tangible, musical and social peer interactions. Our findings suggest that, similarly to other TUIs based on constructive building blocks, the Reactable promotes hands-on collaboration and development of modular configurations, from basic to complex. Furthermore, tabletop interaction with the Reactable over time especially promoted group coordination and learning by mimicking others, in which groups co-constructed meaning differently. Multiple sessions with the Reactable also raised new challenges to group collaboration, concerning awareness of who was doing what and lack of control in early interactions, that were mastered over time. We also discovered more democratic participation roles than in traditional ensembles, with more dynamic exchanges of roles. We found that the Reactable's lack of territorial constraints and its automated connection mechanism promoted group exploration and creative discovery. This may positively motivate collaborative learning in creative group activities such as musical improvisation. In sum, this research suggests that the use of a TTI for music performance can promote peer learning within a democratic setting. This can potentially inform constructivist approaches to learning in other domains based on improvisational and egalitarian tasks (e.g. brainstorming, storytelling) using a computationally enhanced tabletop environment.

The following chapter presents a second exploratory study on the Reactable in a public science centre with visitors involved in one-off interactions rather than expert musicians in multiple sessions.

Chapter 5

Study 2: Naïve Social Interaction with the Reactable

This chapter reports on our second study which explores the experiences of groups of visitors playing with the Reactable in the INTECH Science Centre. Our focus was on the nature of the participants' social interaction and workspace awareness in a real setting. From the analysis of the study findings, we provide design recommendations of how to enable transitional spaces between groups, and how to improve hands-on learning of computer music using TTIs and TUIs in public settings.

5.1   Introduction

TTIs and TUIs in public settings tend to be casual, unstructured, walk-up-and-use, dynamic, and fast-paced interactions, as shown in many in-the-wild studies [Hinrichs & Carpendale, 2011; Hornecker & Stifter, 2006; Marshall et al., 2011]. Previous tabletop studies investigated social interaction using basic systems for casual users [Hinrichs & Carpendale, 2011; Hornecker, 2008; Marshall et al., 2011], in which group transitions have been uncovered as an issue [Marshall et al., 2011]. There is, however, a paucity of detailed research

on how groups share access to these technologies and manage the sociocultural notion of *liminal spaces* [Turner, 1964]. Liminal spaces are transitional spaces that exist between two areas or states [MacDonald, 2014], including transitions between physical spaces, physical-digital spaces, or sound field spaces, among others. Previous work has studied liminal spaces in technology-mediated performance [Broadhurst, 2006]. The concept of transitional spaces can, Broadhurst argues, help us to understand performance. Collaborative music creation has been researched using highly constrained systems for visitors (cf. [Blaine & Fels, 2003]'s survey on collaborative instruments for novices). Musical tabletops appear to be a promising platform for collaboration and social interaction among visitors, with the potential to promote rich and complex interaction. Through detailed observation of interactions with a musical tabletop, our approach provides in depth empirical research findings to inform the design of smooth transitions and collaborative hands-on learning. The motivation for this study is to explore what a complex musical tabletop can tell us about casual social interaction in public settings. Our research questions emerged from analysing video data, framed within our overarching research question about the nature of group tabletop interaction and manifestations of workspace awareness (§1.4).

We conducted a field study of casual social interaction using a complex musical tabletop, the Reactable, which is at the INTECH Science Centre with the purpose of demonstrating concepts of *computer music* (cf. §1.1.1) with a hands-on and collaborative interface (see Fig. 5.1). We were interested in the behaviour of visitors using the Reactable's tangible objects, in an open-ended music-making task in an in-the-wild environment. Our observations point to rich social interaction at different levels of interaction between groups, beyond the more common within-group interaction. Tabletop studies focus on the latter both in the lab and in the field. Our findings suggest that fluid between-group interaction is as important as within-group interaction, and is an overlooked aspect when designing tabletops and TUIs for public settings. Between-group interaction was both synchronous and asynchronous, with fluid transitions and overlaps in use between groups. We identified groups who spent a long time interacting with the Reactable, including visitors who

FIGURE 5.1: Visitors interacting with the Reactable at INTECH

learned by doing. By contrast, we also noted a number of groups who played and seemed to have fun, yet progressed no further than a basic exploration of the interface. Our findings indicate the importance of facilitating smooth transitions between groups in casual social interactions, and the potential role of musical tabletops for understanding concepts of computer music.

Oppenheimer funded the Exploratorium in San Francisco, pioneering the hands-on approach in a science and technology museum. He considered music and sound an important theme when planning the centre: "The section on hearing might be introduced with a collection of musical instruments." [Oppenheimer, 1968, p. 207]. We claim that tangible music on interactive tabletops in public settings is a promising area for visitors to experience music in casual social interaction using tangible objects. This research aims at contributing to this vision.

5.2 The system

The Reactable Experience is a tangible music interface with a rounded table especially designed for public settings. The interface is controlled by positioning acrylic tangible objects on the tabletop surface. There are more than 40 different tangible objects available. Objects can be combined to produce

different sound outcomes which are all synchronised. It is possible to change the sound parameters of an object either by rotating it, or by modifying the visually projected circular controls surrounding it. The rim of the surface is extended and divided into slots for placing the tangibles. As shown in Fig. 5.2, the slots are indicated with a replicated icon of each object and a text label.

At INTECH, the exhibits are organised spatially into sets of categories, one of which is *Computers in Music*, in which the Reactable is set up permanently in an open-plan space close to other interactive exhibits. The purpose of the Reactable in this context is to demonstrate concepts of computer music using tangible objects. Next to the Reactable, there is a TFT display at eye level, which shows a looping demo video of basic interactions with the system. The exhibit also has four pairs of headphones attached to the tabletop, as well as open audio via two loudspeakers embedded within the table. In addition, there is an artificial metal ceiling above of the tabletop for sound insulation purposes. In the exhibition space, there are several stools, that can be moved, for visitors to sit or children to stand on.

A session begins in standby mode with the message "Place the objects on the table border". To start a session, there needs to be at least one object on the tabletop surface. After several minutes of inactivity, the standby mode is activated, and objects must be removed to start again. Once the objects are removed, a view of the type of tangible objects that can be used is presented (see Fig. 5.3 right). The screensaver disappears when the first object is placed on the tabletop surface.

5.3   The study

The study investigates the nature of group tabletop interaction (GTI) among groups of naïve users on the Reactable in a public setting, and how workspace awareness (WA) is manifested during these group interactions.

FIGURE 5.2: The Reactable Experience at INTECH

5.3.1   Study design

The goal of the study was to better understand the nature of visitors' casual social interaction and manifestations of workspace awareness. Social interaction (cf. §1.3) refers here to casual collaborative experiences that include both between- and within-group interaction mediated by an interactive system. Our interest in investigating between-group interaction in terms of transitional spaces between groups emerged from data analysis. Workspace awareness (cf. §1.3) refers here to the up-to-the-moment understanding of another visitor's interaction within a shared workspace. Again, our interest in investigating within-group interaction in terms of how the Reactable was used for understanding concepts of computer music emerged from the data.

5.3.2   The setting

The INTECH Science Centre[1] is a hands-on science and technology centre containing a number of interactive exhibits, most of them developed in house, and a planetarium. The goal of the centre is to promote public knowledge and understanding in the STEM fields (i.e. science, technology, engineering and mathematics), among the younger generation. The ethos of the INTECH science centre is hands-on: at the entrance of the building there are signs with

---

[1]www.winchestersciencecentre.org (accessed 30 September 2014).

sentences such as "I do & I understand", "Doing is believing", or "The hand is the cutting edge of the mind".

The centre is managed by an independent educational charity, The Hampshire Technology Centre Trust Ltd., and is visited both by schools and the general public. It is generally attended by school visitors on weekdays and by the general public at weekends. It has an upper and a lower exhibition area both using an open floor plan.

5.3.3 Data collection

We conducted the study over two days in March 2011. Throughout the data collection period, a researcher adopted a non-interventionist, observational approach, observing the Reactable from a distance, whilst walking around the exhibition space. A poster and leaflets informing visitors of the study were placed both at the entrance to the museum and next to the Reactable; these included research details and a timetable of the video recordings. Data collection was anonymous. If visitors had any concerns about being recorded, they could ask the researcher to switch off the camera or delete the footage at any time.

We collected 8 hours of video data over one well-attended weekend (reported by the centre as 666 visitors in total), consisting of four 2-hour slots allocated in two mornings and two afternoons. For data collection we used two cameras set up non-intrusively: these were positioned to provide a general view of the Reactable and the bodily interactions of visitors, and a close-up view of the visitors' hand gestures on the tabletop surface (see Fig. 5.3). We also collected the Reactable's audio output with a handheld recorder replacing a headphone's output channel as additional data for potential future research. The researcher only approached the tabletop if a visitor wanted to ask a question, which only happened a few times. None of the visitors involved in the study requested their data be withdrawn.

FIGURE 5.3: Left: camera general view. Right: camera close-up view

5.3.4 Method

In order to produce an overview of the full dataset, we coded all of the videos according to the number of groups, the group members, and time of group interaction, using ELAN software. Since the data was collected in the wild, it was rich but unstructured. We used thematic analysis [Braun & Clarke, 2006] to identify relevant themes within the data, and thus to impose an analytical structure on it (the coding of all the video data was done by the author). This enabled us to qualitatively identify relevant themes related to our research questions, and to choose vignettes illustrating these themes. In addition to group discussions with the same team used in Study 1 (§4.3.5), group discussions were conducted with experts in the fields of museum installation design, multimodal interaction, and video analysis in the course of two workshops.[2,3] Discussing our data with experts in the field enabled us to gain further insight into the data, and to better understand the science centre environment.

5.4 Coding scheme

We developed a set of behavioural codes through iterative video analysis. Our analysis focused on groups' behaviours around the Reactable. We found

---

[2]`https://mobiquitous.cis.strath.ac.uk/old/Main/MuseumWorkshop` (accessed 30 September 2014).

[3]`http://mode.ioe.ac.uk/2012/08/12/data-analysis-workshop-analysing -gesture-body-posture-and-action-in-digital-learning-environments -2nd-october-2012` (accessed 30 September 2014).

interesting between-group interaction and within-group interaction. A *group* here refers to one or more people who approached and interacted with the Reactable. We recorded the *time spent* with the interactive tabletop starting from when the first object was positioned on the table, and ending when the last object was used before the group moved to a new exhibit. In the case of groups that approached the table more than once, we added up the different intervals. We also coded individuals as a group so we could consistently investigate social interaction between groups. A *session* refers to one or more participants interacting with the Reactable over a certain period of time. Group sessions tended to *one-off interactions* of approaching and interacting with the table only once. The coding scheme presented here focuses on between-group interaction: transitions and overlaps (§5.4.1), and continuist vs rupturist (§5.4.2).

5.4.1 Between-group interaction: transitions and overlaps coding scheme

Through the data analysis we identified a set of mutually exclusive styles of group interaction related to collaboration between groups, ranging from:

- *Transition*: process of change from a current group leaving the table (*fade out* or sequential moving out) and a new group starting to interact with it (*fade in* or sequential moving in). In such cases there could be a 'handover' of objects (for example, headphones, or tangibles).

- *Overlap*: a new group joins or 'merges' with a current group for a certain amount of time, and both groups interact, mediated by the table, including observation of the practice of the current group.

- *Transition and overlap*: a new group overlaps with an existing group (a new group joins an existing group) before or after there is a group transition from an existing group to the new group.

- *No interaction*: no transitions or overlaps with other groups.
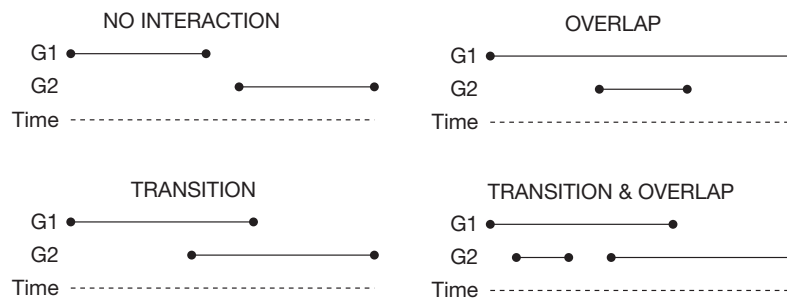
FIGURE 5.4: Between-group interaction: transitions and overlaps

We acknowledge that this categorisation simplifies the complex interactions that took place when an existing group fractured and a part of the group left and a part stayed, for example. However, this approach was useful as it helped us to understand general patterns of between-group interaction. The scheme used a separate interaction style to describe and classify *transition and overlap* in order to identify different levels of between-group collaborations. Visitors who interacted in this way generally showed more interest in the Reactable exhibit, since they were inclined to repeat the experience. Figure 5.4 illustrates the four approaches to between-group interaction.

5.4.2 Between-group interaction: continuist vs rupturist coding scheme

Another level of between-group interaction relates to how the tangible objects were left by a group and 'inherited' by the next group: a 'heritage' that could be sequentially continuous (e.g. during a transition) or sequentially spaced in time (e.g. without human interaction). Each group adopted either a *continuist* or a *rupturist* approach to their musical compositions (patches). The 'continuists' started or ended their sessions in such a way as to continue with others' patches or to allow others to take over. Whereas the 'rupturists' brought them to starting from scratch or ending with no objects. The two approaches are described next:
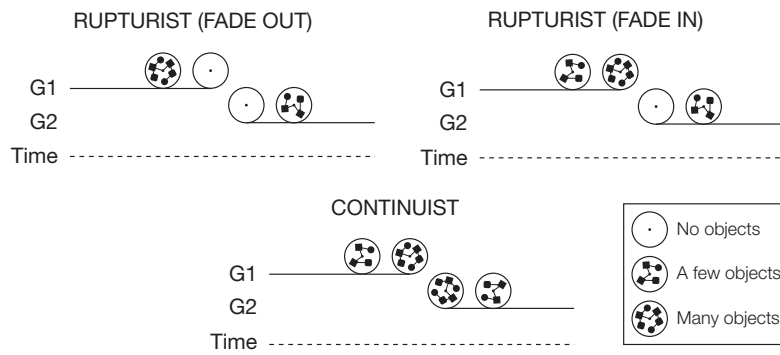
FIGURE 5.5: Between-group interaction: continuist vs rupturist

- *Continuist*: leaving a patch that produces sound on the table for the next group to continue (*fade out*), or inheriting an existing patch from a previous group and modifying it (*fade in*).

- *Rupturist*: removing or *tidying up* all the tangible objects from the table and leaving a 'blank page' before leaving the Reactable exhibit (*fade out*), or inheriting an existing patch and removing all the tangibles to start afresh (*fade in*). A system's standby mode is automatically activated after several minutes of inactivity, which can be also considered rupturist.

Figure 5.5 illustrates the three strategies of between-group musical interaction.

5.5 Findings: overview and summary of the vignettes

We collected no data about the identities of participants rather we counted the groups as they arrived at the Reactable exhibit. We use the following nomenclature from here on: G# for groups and P# for participants, in which b# refers to a boy, g# refers to a girl, M# refers to a man, and W# refers to a woman (e.g. G1 W1 refers to woman 1 of group 1). We observed 54 groups, five of which were, however, clearly subgroups of five bigger groups, such as families (G4 and G47) or school field trips (G35), that arrived together,

percentage of groups by group types

18.52% 11.11%

11.11%

40.74%

18.52%

● Alone   ● Dyad
● Triad   ● Quad
● Quad +
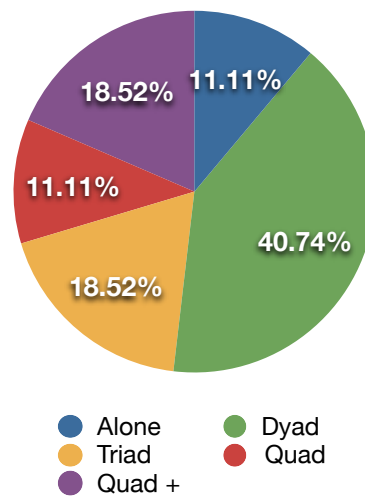
FIGURE 5.6: Pie chart: percentage of group sizes

resulting in a total of 49 groups. The letter A or B is appended (e.g. G4A and G4B) to subgroups.

The 54 groups comprised 6 individuals, 22 dyads, 10 triads, 6 quads, and 10 larger groups from five up to 14 members. Figure 5.6 shows the percentage of group sizes. We noted 48 out of 49 original groups as comprising: 36 families, 2 school field trips, 6 groups of colleagues, and 4 individuals. Classifying a group as 'a group of colleagues' was established through the features of group behaviour and age. Our observations include a total of 170 individuals: 74 adults (43 women, 31 men) and 96 children (52 boys, 44 girls). In addition, four participants were carrying toddlers. A few individuals spontaneously stopped for a few seconds in the exhibit area or slowed down when passing by the exhibit. These interactions did not provide relevant data, and were excluded from the analysis.

The time each group spent with the interactive tabletop was recorded. The median duration of use by groups was $4\,\text{min}\,14\,\text{sec}$ ($254.4\,\text{sec}$). Group interactions tended to be one-off—only eleven groups repeated the visit to the table, who tended to be those who spent longer time. The time spent with the interactive tabletop by group was not normally distributed: a Shapiro-Wilk test confirmed this ($p = 6.286e - 07$); a Kolmogorov-Smirnov test did not reject

TABLE 5.1: Time spent by groups

Condition	Duration (sec.msec)	Duration (min:sec)
Minimum time spent	2.1	00:02
Maximum time spent	1778.122	29:38
Median duration of use of groups (54 groups)	254.4	04:14
Median duration of use of groups (17 groups) < 2 min (peak 1)	61.0	01:01
Median duration of use of groups (14 groups) >= 2 min and < 6 min (peak 2)	331.6	05:31
Median duration of use of groups (18 groups) >= 6 min (long tail)	697.3	11:37

an exponential distribution. The histogram in Fig. 5.7 points to an exponential distribution, in which there is a large peak at $0\,min$–$2\,min$ ($0\,sec$–$120\,sec$), which includes 17 groups, and another peak at $4\,min\,30\,sec$–$6\,min$ ($270\,sec$–$360\,sec$), which includes 14 groups. There is an expected exponential fading out (cf. Hornecker & Stifter [2006]) with a number of short interactions and successively less longer ones.

Table 5.1 outlines the minimum, maximum, and median duration of groups. These results support the median duration found in museum studies, in which it is common to find up to 4 min spent in different installations, ranging from traditional object exhibits to computer-enhanced hands-on exhibits, and longer interactions up to 15 minutes [Hornecker & Stifter, 2006]. The two peaks are unusual, and indicate different levels of engagement with the Reactable.

A qualitative analysis with vignettes to complement these results is provided in the next section. We observed both synchronous and asynchronous interactions both within and between groups. We were especially interested in between-group interaction because it showed that tabletops promote a wide social collaboration in public settings beyond preformed groups (§5.6, 5.7). We were also interested in the different group approaches to collaboration mediated by the table (§5.8).

Table 5.2 sets out the vignettes discussed in this chapter.
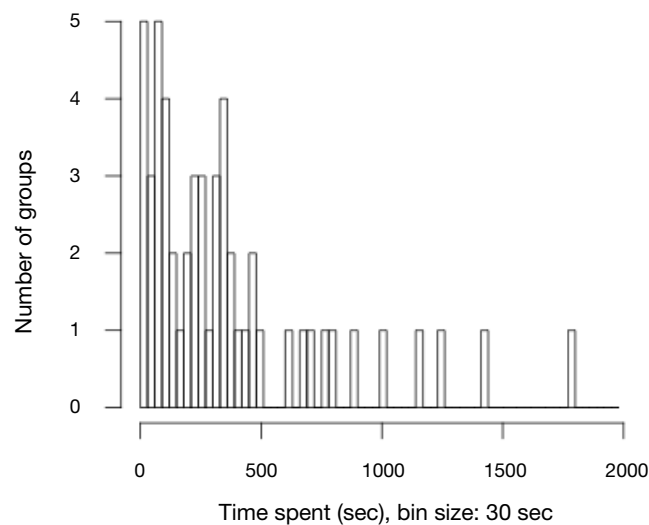
FIGURE 5.7: Histogram: distribution of time spent by groups

TABLE 5.2: Overview of vignettes

Section	Topic	Vignette	Figure
Social interaction (§5.6, §5.7)	Interactions between groups	1. Fluid transition	5.9
		2. Smooth overlap	5.10
		3. Asynchronous continuist	5.12
		4. Asynchronous rupturist	5.13
Workspace awareness (§5.8)	Physical-digital awareness	5. Experts and beginners	5.14
		6. Beginners	5.15
		7. Beginners to experts	5.16

5.6 Findings: fluid transitions and overlaps in between-group interaction

Thirty three groups, out of the 54, interacted with other groups (17 "transition only", 8 "overlap only", and 8 "transition and overlap"), whilst 21 did not. Figure 5.8 shows these results by percentage. It is worth mentioning that between-group interactions, and transitions in particular, were fewer, or absent, when the public venue was sparsely attended. In these circumstance, groups took turns. Similarly, no interaction took place when the public centre was sparsely crowded. When it was crowded, there were more instances of transitions and overlaps. The groups which did overlap and transition tended to first approach the interactive tabletop as observers, normally led by one member of the group; to arrive and later interact with the musical
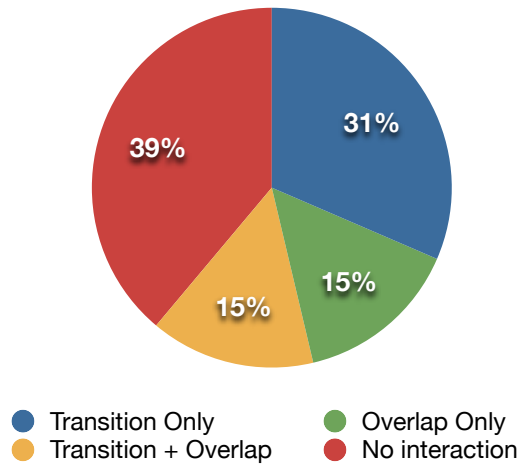
FIGURE 5.8: Pie chart: percentage of between-group social interaction styles

tabletop as a whole group. This self-regulation of turn-taking is interesting in a public venue context in which social interaction is recognised as an important factor in visitors' experience of museums and galleries and to influence their behaviour [Heath & vom Lehn, 2009]. Why this happened and how it can influence the design of exhibits is discussed in the next sections §5.6.1 and §5.6.2, in which we present the nature of these between-group interactions.

We next present two vignettes: *fluid transitions* (§5.6.1), and *smooth overlaps* (§5.6.2). We also describe the two combined (§5.6.3).

5.6.1   Fluid transitions

Between-group transitions are expected in public settings when there are a number of interactive exhibits available. Transitions were a frequent pattern of change of control from a current group to a new group. They were fluid and can be classified as being one of two types: 1) members of a current group ceasing to use the tangibles and moving to another location, with members of a new group approaching and starting to use the tangibles; and 2) members of a current group exchanging the headphones (as these were a limited resource) with newcomers, sometimes chatting about how the table works, before moving to another location. Vignette 1 exemplifies the latter pattern.
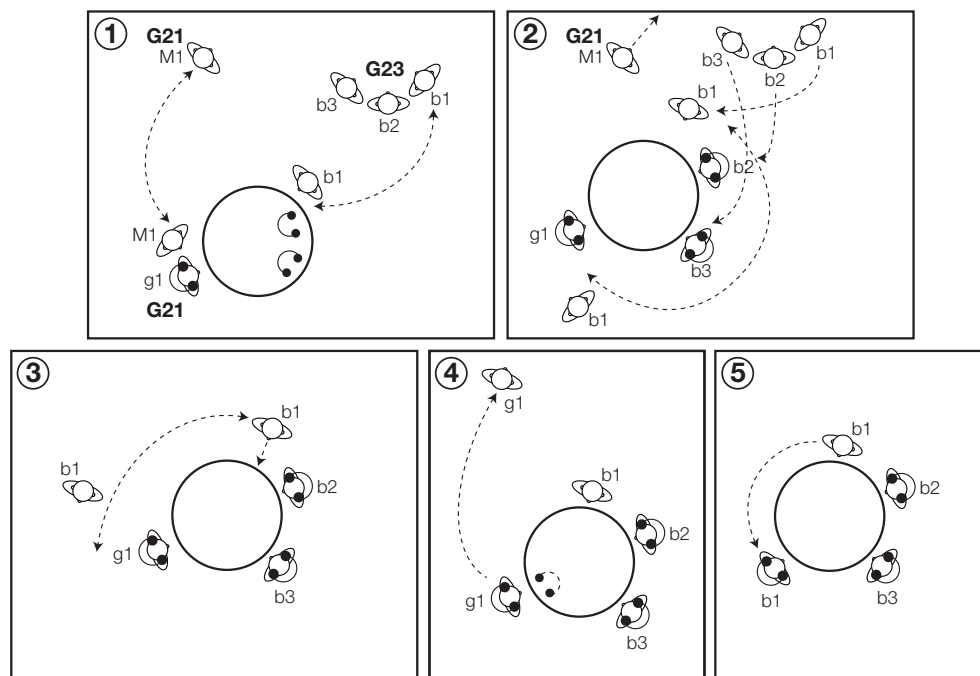
FIGURE 5.9: A fluid transition

Vignette 1: Fluid transition (G21 and G23). Duration: 2 min 56 sec. As shown in Fig. 5.9, a girl (g1-G21), who is wearing headphones, is manipulating tangible objects on the table (frame 1). The other two headphones are available at the opposite side of the table. A boy (b1-G23) approaches the table touching the surface with his hand and disappears. The girl's father (M1-G21) approaches the table and shows the girl a piece of paper from another exhibit and they both laugh. The father disappears. Group G23, a group of three boys (b1–b3), approach the table (frame 2). The boys b2 and b3 get the two available headphones and start interacting with tangibles (frame 3). The girl looks at them, looks back to the table, and continues playing. The boy b1, who has no headphones, walks around the table looking at the floor. He collects a tangible object from the floor and returns it to the rim area. Then he stops between the girl and the boy b2, and puts his arm on the table. He is not playing (frame 3). After a few seconds, the girl leaves her headphones on the table, along with a few active objects (frame 4). She searches out her father, moving towards him. The boy b2 points to the available headphones and touches b1's arm to let him know. The boy b1 moves to where the available headphones are, wears them, gets an

object from the rim area and starts playing with it (frame 5). The girl's objects are reconfigured over time.

This vignette illustrates a fluid transition between two groups. This is possible because group members can make their intentions explicit and visible to all. Group change is smoothly supported by the system because tangible objects play autonomously and they can be added or removed by dynamic patching (cf. §2.5.2). The girl noticed the arrival of a new group with an interest in playing with the table, along with the lack of sufficient headphones for them. Leaving the headphones on the table makes visible a change of turn to the next group. The new group smoothly reconfigures the objects that the girl has left. In other groups there was a similar handover of headphones although it was made more explicit than in this vignette. A handover of headphones reveals the mutual group awareness of an impending group change. Headphones are used here to make a group change visible as well as to communicate that they are to be used to experience the interactive exhibit. Predicting the number of visitors in public settings can be a challenge, thus designing interactive exhibits that promote on-demand fluid transitions can facilitate a smooth flow between exhibits.

5.6.2 Smooth overlaps

We also noted smooth overlaps between groups. Our observations suggested that overlaps were far less common than transitions in public settings. A number of groups overlapped (see Fig. 5.8). An overlap could be either to 1) closely observe a current's group session and then move to another location; or 2) interact with the current group. Groups that only overlapped approached the table and joined a current group's musical activity, as shown in Vignette 2.

> *Vignette 2: Smooth overlap* (G10 and G11). Duration: 38 sec. As shown in Fig. 5.10, two girls, g4 and g5, from G10 (a school field trip group) are playing with the table while wearing headphones (frame 1). They are standing on opposite sides of the table. There is a pair of headphones
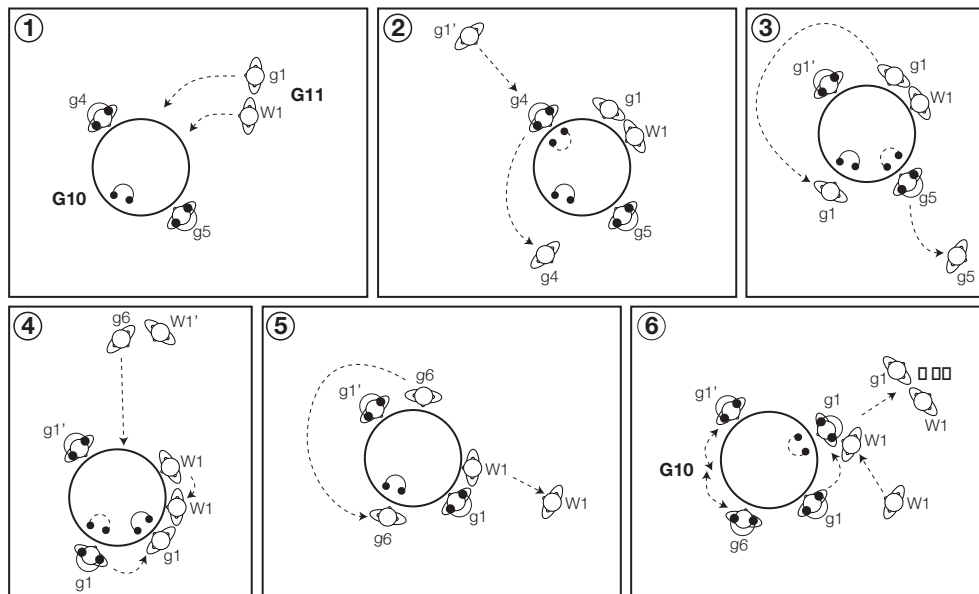
FIGURE 5.10: An overlap

available as well as free available space between the girls. The group G11, consisting of a girl (g1) and a woman (W1), are passing by holding hands. The girl g1 points to another exhibit, and the woman W1 says "here's the music darling" pointing to the Reactable. Both approach the table, get an object each from the table, and start interacting with them (frame 2). The girl g4 takes off her headphones and disappears. Another girl from G10, g1', who previously interacted with the table, moves towards the table, and puts on the pair of headphones left by g4. The girl g1 moves to where a third pair of headphones is available and puts them on, while the girl g5 takes off her headphones and disappears (frame 3). The girl g1' and woman W1 are manipulating objects, while the girl g1 replaces her headphones with the ones left by g5 (frame 4). The woman W1 moves closer to where g1 is, leaving available space between her and g1'. The girl g6 from G10, before approaching the Reactable, is next to another exhibit chatting with her group's chaperone (W1), she turns her head looking at the table, she approaches the table towards where the available space is, and then she moves to pick up the third pair of available headphones (at the moment g1' and g1 are wearing the other two pair of headphones) (frame 5). The girl g6 moves left to right between g1' and g1 (frame 6). The woman W1 steps back. The girl g1 moves to her right, where W1 was. The girl g6 moves to her right. Then

girl g1' swings from left to right. The woman W1 approaches g1, whispers something to her, g1 takes off her headphones, and they disappear together.

This vignette shows a smooth overlap between two groups. Three features of the Reactable make it possible for a group to join an existing group seamlessly:

1. Its obvious *visibility and audibility* when passing by that is associated with music.

2. The *rounded shape* of its surface allows visitors to reconfigure their positions dynamically, and to have different points of access.

3. Its *object-based interface* consisting of an extensive collection of objects available.

Furthermore, seeing others playing with tangible objects, along with an interface that has real-time audiovisual feedback, prompts newcomers to join the activity smoothly and explore the objects together with no need for the verbal communication. Although headphones do not seem to be necessary for engagement, they are a high value resource that is in demand. Encouraging collaborative group experiences with exhibits that orchestrate group interactions between strangers aligns with the museums' agenda of moving from object-centered to visitor-centered experiences [Vom Lehn, 2006].

5.6.3   Fluid transitions and overlaps

Other groups fluidly transitioned and overlapped. In these cases, a group, or some of the members of a group, overlapped with another group or groups, before or after formally interacting with the exhibit. Here formal interaction refers to a group positioned around the table and playing. We found examples of overlaps prior to, or after, transitions. A prior overlap followed by a transition, which was different from the simple overlap in vignette 2, tended

to be observational or with little interaction. A prior overlap was constituted as: 1) part of a or whole new group approaching the table as observers, or casual players, usually to get a sense of how the interactive exhibit works; and 2) a new whole group returning and transitioning with the current group. Overlaps after transitions happened less often, and tended to involve an earlier group, or part of an earlier group, returning to join the group that is now playing. For example, see vignette 7, explained in §5.8.3, in which a boy (b1-G12) returns to improve his interaction techniques.

5.7 Findings: continuist vs rupturist in between-group interaction

In collaborative music systems, *co-located collaboration* can be either synchronous or asynchronous (see §2.2.2, cf. Barbosa [2003]). We were particularly interested in understanding the continuist approach as an instance of co-located collaboration that includes both synchronous and asynchronous interaction. Continuist interaction with an interactive tabletop in a public setting can be seen as a form of collaboration between strangers, with a lapse of time in some cases.

Out of the 54 groups, 36 adopted a continuist approach (67% of the groups), of which 12 started a session from an inherited patch asynchronously from another group, whilst 24 did so synchronously. By contrast, only 13 groups adopted a rupturist approach (24% of the groups), of which 9 faded in (groups approached the table and removed all the tangible objects from the active surface), and 3 faded out ('tidying-up' before leaving). Figure 5.11 outlines these results. We discarded five groups (9%) because the video data did not enable the group's approach to be ascertained: they were already playing when we started the video recordings, or kept playing after we stopped recording.

In the following section we present two vignettes focusing on asynchronous interaction: *asynchronous continuist* (§5.7.1) and *asynchronous rupturist* (§5.7.2).
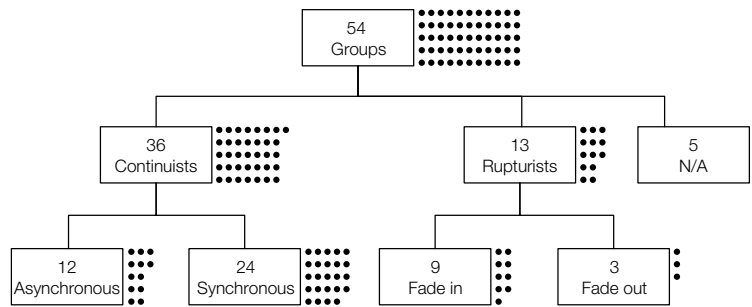
FIGURE 5.11: Tree structure of continuist vs rupturist approaches
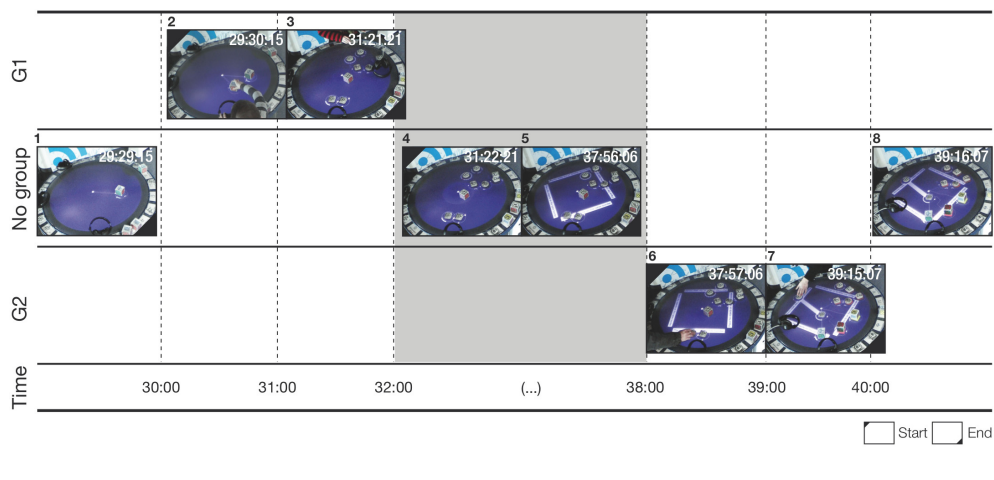
FIGURE 5.12: Continuist asynchronous interaction

We focus on asynchronous interaction because it can tell us more about the nature of transitional spaces than synchronous interaction, which was manifested during transitions already discussed in the previous section (§5.6).

5.7.1 Asynchronous continuist: *cadavre exquis*

Most of the groups adopted a continuist approach: approaching the table and continuing the patch left by the previous group. This can be seen as a form of asynchronous collaboration between strangers. Vignette 3 describes an example of this approach.

> *Vignette 3: Asynchronous continuist* (G1 and G2). Duration: 9 min 47 sec. As shown in Fig. 5.12, the group G1 (b1, g1, and W1) approaches the table (frame 2). The group is interacting with the table.

Each is wearing a pair of headphones. There are 19 active objects on the table. The boy b1 is manipulating two objects simultaneously. He leaves the two objects in the rim area. He moves two more objects from the active area into the rim area. He takes off his headphones, leaves them on the table, and moves to another exhibit saying "What's next?". The girl g1, who is standing on a stool, leaves the object that she is manipulating in the rim area, takes off her headphones, puts them on the table, and moves to where the boy b1 is. The woman W1 leaves the object she is manipulating in the rim area. She takes off her headphones. She sequentially moves four objects in the rim area, the sound stops (frame 3). The boy asks from another interactive exhibit: "Mummy what's this?". The woman W1 says: "Pardon?". She puts the pair of headphones used by g1 on the floor and moves to where b1 and g1 are. There are 9 objects on the active area of the table, with visual feedback, but no sound because there are no active sound generators (frame 4). Two minutes later, the standby screen mode appears, with 9 inactive objects with visual feedback on the table surface (frame 5). About five minutes later, the man M1 of G2 approaches the table. He takes one of the objects off the table and pushes it towards another object twice (frame 6). The woman W1 of G2 approaches the table and looks at the table surface for a few seconds. There is visual feedback but no sound output. The man takes a pair of headphones next to him and puts them on. He pushes twice another object towards the centre of the table. He takes off the headphones, returns them to the table surface, and leaves. The woman takes another pair of headphones, wears them momentarily, and puts them back on the table surface. She changes the position of a cube by dragging the object, she changes the side of another cube. There is visual feedback of the objects on the table along with the standby mode screen, but no sound output. She changes a few more sides of another cube, drags two more objects, and then leaves (frames 7 and 8).

This vignette illustrates a different type of collaboration from synchronous between-group interaction. The first group leaves a patch with a certain configuration of objects, and after a few minutes, another group approaches the table and modifies the patch by changing the configuration of the tangibles. Even though there is no sound, both the automatic visual standby mode

along with the visual feedback of objects seem to encourage visitors to approach the table and interact with the objects. Other groups worked with sound output instead. The benefit of inheriting a patch is that for some visitors it can be easier to start than from scratch. This interaction style recalls the game *cadavre exquis*, also known as *exquisite corpse*, in which participants create a written or graphical collaborative piece by contributing in sequence, sometimes only seeing part of the piece. An asynchronous continuist group only sees and hears the last patch left by the previous group, as well. It also recalls post-it based interfaces, which support a collaborative activity in which participants can leave messages as well as read others' messages. Promoting an asynchronous collaborative interaction style in public settings is a promising approach to supporting collaboration based on visitors' heritage. This approach aligns with Hindmarsh et al. [2005]'s suggestion of leaving an 'activity trace' for future visitors. Furthermore, Hindmarsh et al. provided *design sensitivities* for encouraging social interaction between visitors around exhibits and discussed the potential of collaborative work between companions and strangers in public settings whether co-located or remote, and the need of creating 'opportunities of interaction'. Our findings draws on this work.

5.7.2 Asynchronous rupturist

A few groups adopted a rupturist approach, which meant either 1) a current group moving all the tangible objects to the rim area before moving them to another location (fade out); or 2) a new group approaching the table with an existing musical patch and moving all the tangible objects to the rim area and starting their musical session with a clear table (fade in). Vignette 4 shows an example of a rupturist fade in.

> *Vignette 4: Asynchronous rupturist (fade in)* (G52). Duration: 35 sec. As shown in Fig. 5.13, the woman W1 approaches the table, on which there are already a number of tangible objects producing sound (frame 1). W1 starts to move each object from the active area to the rim area until the active surface is empty (frames 2–4) and the screensaver appears
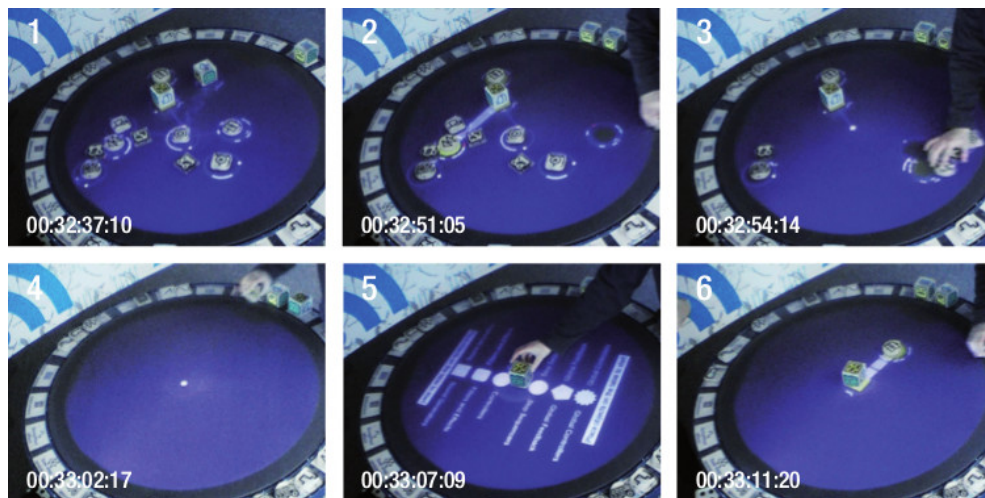
FIGURE 5.13: Asynchronous rupturist (fade in) interaction

(frame 5). She then positions a cube in the middle (frame 5) and a step sequencer (controller) that affects the cube (frame 6).

This vignette indicates a behaviour of tidying up the existing tangible objects and starting from an *empty canvas*. Some visitors seemed to prefer to remove all the existing objects when starting a session. There could be a number of reasons for this, including a preference for having more personal control of the patch. It could also relate to a controlled, progressive, step-by-step learning, which seems useful for hands-on learning.

5.8 Findings: physical-digital awareness in within-group interaction

Our observations suggest that in most of the groups (20 groups), group members approached the table together and started to interact at the same time; yet in some groups (4 groups) group members approached the table sequentially; in other groups (18 groups) group members interacted as subgroups; and in further groups (6 groups) members were divided into actors and observers. A few members only used headphones and observed the other team members' interactions. Active participation with the table was thus high.

FIGURE 5.14: Experts and beginners

Some of the participants interacted for long sessions, and developed expertise. We next present three vignettes that provide detail of different levels of physical-digital awareness with the Reactable: *experts and beginners* (§5.8.1), *beginners* (§5.8.2), and *experts* (§5.8.3).

5.8.1 Experts and beginners

On a few occasions we observed expert users who already knew how the musical tabletop works, sharply contrasting with the casual visitor, for whom this was the first interaction with a musical tabletop. Vignette 5 illustrates a knowledgeable visitor who, after a solo session, returned for a second session with a colleague and showed her how the system works.

> *Vignette 5: Experts and beginners* (G45A). Duration: 21 sec. As shown in Fig. 5.14, there are several objects on the table. A woman (W1) points to the melody generator object, which is in the opposite rim area of where the woman W1 and the man (M1) are positioned (frame 1). M1 passes the object on to her (frame 2). She positions the object in different places on the table, but the object does not seem to work (frame 3). He asks her for the object with a hand gesture (frame 4). M1 shows her how

the melody generator works by positioning the melody generator next
to the electric keyboard instrument (the melody generator emits a pre-
defined control sequence to the instrument) (frame 5). Once working,
he points to the melody generator, the electric keyboard instrument, and
then the dot in the middle (frames 6–7). She rotates the electric keyboard
instrument changing the frequency of the instrument (frames 7–8) while
M1 is looking at her action (frame 8). They both interact together: M1
rotates the melody generator, which changes the control sequence affect-
ing the instrument; whilst W1 rotates the instrument, which changes the
frequency of the instrument (frame 9).

This vignette illustrates that in a hands-on science museum, we may find not
only casual visitors, but also experts who can help beginners in understand-
ing how the system works. It also indicates that a hands-on exhibit can be a
suitable tool for exchanging practical knowledge between peers. In particu-
lar, tangible music permits hands-on peer learning among visitors based on
bodily interaction, with no need of verbal communication.

5.8.2 Beginners

We observed a number of examples of beginners interacting with the musical
tabletop for the first time. On many occasions, these visitors built on previous
work generated by other groups. Vignette 6 shows an example of a group of
five girls exploring the musical tabletop.

> *Vignette 6: Beginners* (G34B). Duration: 3 min 27 sec. As shown in Fig.
> 5.15, a group of five girls (g1–g5) approaches the table. The girl g3 takes
> the lead and says "take them out" and the group remove all the objects
> to the rim (frame 1). Three of them wear the three available headphones
> (g1, g2, g5), leaving g3 and g4 without. The girl g3 asks "can I have the
> headphones for a minute? (...) can I hear, just for a second please?", but
> none of the three with headphones comply. In a few seconds the table
> is crowded with objects (frame 2), the objects being rapidly banged, or
> dragged from left to right or top to bottom. The fiducial markers of some
> objects are facing up. After exploring the interface, the talkative girl g3
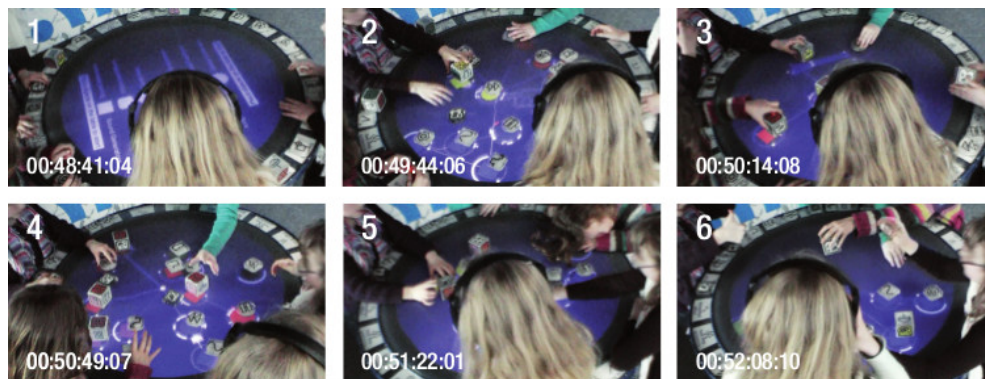
FIGURE 5.15: Beginners

suggests: "Take everything off and start with big blocks, everyone gets a big block". Apart from the girl g1, the rest remove all the objects and start again. Four of them hold an object and put it on the table, which rapidly becomes occupied again. The girls keep bashing or dragging the objects at a frenetic pace, with rhythmical sounds of collisions. A few seconds later, there are at least 15 objects on the table (frame 4). The three girls with headphones are dancing while operating an object each. The girl g1 swiftly drags an object from left to right. The girl g5 rapidly moves a cube up and down, and then draws circles with it. The girl g4 jumps to reach an object located at the opposite area of the rim from her. Objects quickly change position and side, top, down, left, right. The fiducial markers of some of them are facing up. The girl g4 (with no headphones) moves to another side of the table, and she puts her ear to the table after she puts a tangible down on the surface (frame 5). The girl g3 asks "can I get the headphones please now?". The girls put back all the objects in the rim area, but quickly put objects back onto the table. The girl g3 says "Ok we'll do one more and someone needs to change", but the other girls seem to not pay attention to the girl g3 and they do their own operating objects randomly. The girl g3 (with no headphones) raises an object and asks to the girl g5 (with headphones) "tell me when it's off" then puts back the object on the table and tries with another whilst the girl g5 thumbs up in reply to g3's query (frame 6).

This vignette illustrates how the group seems to enjoy the hands-on exhibit in terms of the physicality of the tangible objects and their potential behaviours

(e.g the sound of the impact of acrylic objects on the surface or against other objects; or physical actions such as banging, spinning, or throwing objects). While working collaboratively, amid the chaos rhythms emerged from exploring the tangibles randomly. Sometimes a leading voice emerged (for example, girl g3), but rapidly dissipated. The girls iterate from restarting (*tabula rasa*) and turntaking to generating complex configurations. There are sonic, haptic, visual, and social explorations. It appears unclear what the children are focusing on, whether acoustic, touch, or visual cues. Random actions seem to be part of this exploration. Moreover, headphones seem an important asset for experiencing the exhibit but are a scarce resource. For girl g3, the experience seems to be less immersive; she is continuously requesting headphones with no success. She seems less aware of what the objects do. The girls tried different scenarios in collaboration, producing dynamic changes in the position of objects, bodies, and sounds produced. There seems to be little workspace awareness of the objects, as indicated by a number of instances of turned over objects. This indicates some need for structured support in a learning process on concepts of computer music beyond exploration, which is the purpose of exhibits like the Reactable in public spaces.

5.8.3 Beginners to experts

As shown in Table 5.1, a subset of 18 groups performed long sessions (longer than 6 minutes) compared to the median duration found in other museum studies. Some of these groups repeated the interaction twice. The following vignette shows a group member (a boy) who started as a beginner and developed some skill in a short period of time, based on hands-on interaction along with observing others and mimicking actions from the video demo.

> *Vignette 7: Beginners to experts* (G12 b1). Duration: 18 min 7 sec. As shown in Fig. 5.16, the boy b1 is wearing headphones and interacting with the table by moving objects from the rim area to the surface, and back. Boy b3 approaches the table and puts on a pair of headphones. The boy b1 sees the monitor of the demo video and points it out to his colleague (frame 1). Both watch the video for a few seconds. Then, b1
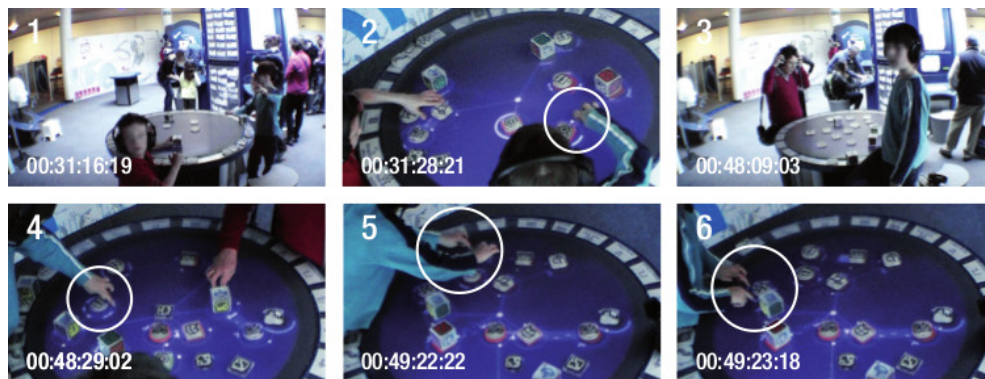
FIGURE 5.16: Beginners to experts

moves his fingertip on the projected slider of a ring modulator, an action that controls the amount of the effect (frame 2). Both leave the table (…), the boy b1 approaches the table again after a break of 15 minutes (…), the boy b1 puts on a pair of headphones, and watches the video demo for about 30 seconds (frame 3). Then, he turns on/off some steps of a square step sequencer with a fingertip (frame 4). Later, b1 repeatedly turns on/off some steps of a square step sequencer with two fingertips (with his left and right hands) (frame 5). Subsequently, he repeatedly turns on/off some steps of a saw step sequencer with two fingertips (left and right hand) (frame 6).

This vignette illustrates that some casual visitors are willing (and able) to develop skill within a hands-on science museum, in which the usual expectation is that visitors will engage in short one-off interactions. The Reactable seems to support both approaches simultaneously, which points to the public centre as a space for collaborative learning as well as one-off interactions. By watching the video demo, in which interaction techniques are shown in detail, the boy learns and progresses by imitating. In this example, the boy interprets the video first to learn how to move the projected slider of a sound effect, and then how to change rhythmic patterns of a step sequencer. He then makes connections between similar objects and applies to them the learned techniques, but with greater complexity (e.g. using two fingertips rather than one). This shows how moving from exploration to enquiry is possible within tangible computer music in a casual museum setting. This aligns with the

aims of setting up the Reactable in a museum—hands-on learning with the interface of a particular digital instrument.

5.9 Discussion

5.9.1 Casual social interaction

Our findings indicate the importance of enabling transitional spaces for casual social interaction to occur in collaborative interactive systems for public settings: spaces between groups (e.g. transitions, overlaps, or continuist collaboration), and spaces between exhibits (e.g. auditory transitions from experiencing sound with speakers to experiencing sound with headphones). These findings show that fluid transitions and smooth overlaps between groups can be orchestrated by a tabletop TUI and reflect the nature of casual interaction in a public setting. Furthermore, a system that facilitates asynchronous group interaction promotes more approaches to collaboration, which can enrich visitors' experience because there are more possibilities and levels of group interaction. Enabling transitions, overlaps and continuist collaboration recalls the notion of *multiple access points*, a characteristic of TUIs highlighted by Hornecker & Buur [2006]'s framework, features by distributing control among users and thus promoting a more democratic setting. In tabletop design, features such as enabling *joint work* and *tailorability* of parts of the system are two common and salient design goals [Dalsgaard & Halskov, 2014] that relate to within-group interaction. The richness of social interaction in the real world can be modelled with tabletop and TUIs if liminal spaces and transitory states are enabled with the interactive exhibit to support better experiences, this is discussed further in §5.9.3 as design recommendations.

As shown in vignette 3, the possibility of an asynchronous continuist approach across groups seems to support social interaction beyond immediate simultaneous physical interaction. For example, when a group leaves a musical configuration on the table, other visitors can hear it, and later a new group can develop it. This appears a particularity of musical TTIs, especially

one with a circular surface, which promotes *multiple access points* so that visitors can access it from any direction. In particular, the use of the auditory channel sound in tangible music seems to strengthen both synchronous and asynchronous co-located musical collaboration because visitors can hear others' work and modify it later in time, asynchronously. This approach fills a gap in Barbosa [2003]'s classification of computer-supported collaborative music, based on user's location and performance synchronicity (cf. §2.2.1). We expand the term *local inter-connected musical networks*, classified by Barbosa from just synchronous co-located collaboration in networked systems, to include also asynchronous co-located collaboration within the context of CSCM: a collaborative musical installation, in this case a particular configuration of tangible objects on the Reactable, can be modified over time, not only synchronously (generally within groups), but also asynchronously (generally between groups).

From our observations, we give design recommendations in Section 5.9.3 about how to support casual social interaction through smooth transitions and liminal spaces between groups considering: 1) the role of headphones; 2) the role of an autonomous system; 3) the role of a tabula rasa state; and 4) the role of physical tangible objects. In this context, a liminal space refers to one group leaving off and a new group taking over, or individuals passing from one state to another (for example, when one puts on headphones to when one takes off headphones).

5.9.2 Physical-digital awareness

As shown in vignettes 5 and 7, there were instances of groups making sense of the interactive exhibit via experimentation with the tangible objects. However, our data on short-term interactions suggested that it is generally difficult to assess whether groups are learning how to operate the Reactable and learning computer music by doing or not. As shown in vignette 6, leaving fiducial markers facing up is evidence of a lack of understanding of how to create

sound with the tangible objects and thus the connections between the physical and the digital domains. Tangible music in public centres seems a suitable tool for attracting a range of varied groups to play with digital music and potentially learn computer music in a casual manner. The complex process of scientific enquiry has been traditionally explored in hands-on centres such as Exploratorium. Rogers & Price [2004] divide the process of scientific enquiry into a set of strategies: exploration and discovery; describing and interpreting; hypothesizing and prediction-making; testing ideas and experimenting; making links; or giving explanations and concluding. We borrow this process to a more practice-based area, hands-on computer music and music making, which is similar to other computer-based practices such as programming, or electronics. From this angle, our observations indicate that most of the groups only developed exploration and discovery.

In one vignette, an expert user showed a beginner how two tangible objects worked together, and helped the beginner to manipulate the objects with greater control. In another, a beginner watched the demo video several times, and imitated certain interaction techniques from the video, which helped him to make progress with interface control. These examples suggest that practice-based instruction can help naïve users. The difficulty of progressing with the Reactable in a public setting arises from the fact that it is more challenging to learn about than some kind of push-button exhibit, and hence requires more concentration and systematic experiment to make progress with it.

The nature of experimentation that the adult visitors employed differed from that of children. As shown in vignette 5, groups of adults were more considered in the choices they made, making one at a time, and checking the feedback (exploration and discovery). As shown in vignette 6, groups of children, instead, experimented with more chaos, bodily interaction, and dynamic changes (exploration). In addition, the children engaged in the physical elements of the installation, such as bashing the objects. In both examples, the collaborative hands-on experience appears relevant to hands-on experience for learning.

From our observations, we offer recommendations for the design in §5.9.3 of musical tabletops that support workspace awareness and the learning of computer music as 1) guided activities in hands-on computer music; 2) short-term and longer-term one-off interactions; and 3) self-explanatory TUIs.

5.9.3 Design recommendations

Musical tabletops can encourage social interaction beyond the group itself: our results indicated that there were situations in which groups transitioned or overlapped fluidly (e.g. vignettes 1 and 2), or collaborated asynchronously (e.g. vignette 3). Yet we identified design aspects that could be improved for enhancing the collaborative experience for both between- and within-group interaction: for example, we also found moments of tension between visitors when the group was larger than the available number of headphones, as shown in vignette 6. We also found examples of learning progress (e.g. vignettes 5 and 7), as well as counterexamples of random actions (e.g. vignette 6). We list below a set of design improvements for the interactive experience, the aim being to enabling liminal spaces and smooth group changes along with supporting the learning experience in a democratic setting:

1. *The role of headphones*. Hearing the sound from the table speakers combined with seeing headphones allow for fluid transitions and smooth overlaps between groups because they indicate the type of activity and available spaces in which to join the activity. The purpose of the exhibit is clear to visitors at first sight, as seen in vignette 2 when G11 passes by, hearing loud music coming from the table and recognising headphones as a familiar object. Moreover, as seen in vignette 1 when G23 leaves the headphones on the table, the headphones facilitate a change of turn between groups, as they have to be exchanged. As soon as a visitor starts to wear headphones, there is a subtle auditory change from a less to a greater immersive experience, from a low volume to a louder volume of sound. Thus, this transition is less disruptive than if it would be from silence to sound. However, there is a

limit to the number of headphones that can be made available, and thus the number of visitors who can join to the activity as fully immersive (see vignette 6). It is a challenge to keep a balance between improving the auditory experience for those who are not wearing headphones and to provide a slightly more immersive experience. Furthermore, there seems to be several trade-offs between concentration and distraction in an open environment like a museum; between the isolation of groups engaged in an immersive experience, and the need to attract and interest other visitors; and between interest in the Reactable and interest in other exhibits.

2. ***The role of an autonomous system***. The Reactable operates in real time using audiovisual feedback. When objects are placed on the table surface, they have an autonomous behaviour with no need of human manipulation. However, objects can be manipulated as well. This twofold feature allows groups to approach the table, observe the system's behaviour, and intervene smoothly at any moment. These two features facilitate the ability of groups to notice the table even when there are no groups interacting with it. As shown in vignette 3, when G2 approached the table, our observations indicated that visual autonomous behaviours alone can attract the attention of visitors. However, it is unclear whether an autonomous visual standby mode with a lack of sound is useful in musical tabletops and audiovisual exhibits, as they can mislead visitors about how the system works. If the purpose of the interactive tabletop is to demonstrate concepts of computer music during casual interaction, that should be clear when approaching the table before beginning the interaction, during the transitional space from walking to starting the session.

3. ***The role of a 'tabula rasa' state***. As seen in vignette 4 when G52 starts a session, for some of the visitors, it was useful to remove all objects from the table, and start from the position of a tabula rasa. However, our observations suggested that an automatic tabula rasa state can have rupturist consequences within an asynchronous collaboration because

a configuration of objects (a patch) left by a previous group can stop playing with the objects on the active area of the table for a forthcoming group as seen in vignette 3. This can disrupt fluid transitions and collaboration between groups. A manual tabula rasa state seems useful, but it is unclear how an automatic system's standby mode after a few minutes of co-located inactivity can support social interaction.

4. ***The role of physical tangible objects***. As seen in vignettes 3 and 6 when visitors operate with objects, the fact that the Reactable's objects are constructed from acrylic allowed for bashing or dragging the objects. We observed that the availability of an extensive collection of objects and a large surface allows for a collective haptic exploration and experience, in which visitors can smoothly join or leave at any time. Moreover, seeing others operating the objects was also helpful for newcomers in relation to how to use them (e.g. during overlaps), as seen in vignette 2.

5. ***Guided activities in hands-on computer music***. As shown in vignette 6, most of the groups remained in a playful exploratory state with an unclear musical learning outcome. Designing guided activities appears important for helping groups of visitors to progress with the interface. These activities should, our findings suggest, be short and practice-based, maintaining the spirit of one-off social interactions in public venues, and of music performance as an open activity. Perhaps constructing an area of interrelated exhibits on sound and music computing, inspired by Oppenheimer [1968]'s vision, would reinforce an environment for learning computer music. The experience of hands-on centres such as the Exploratorium can help to build these musical activities. This centre investigates interactive exhibits that are both "'minds-on' as well as 'hands-on'"[Allen, 2004, p. S25] by eliciting *active prolonged engagement* (APE), experiences that involve both practical as well as cognitive knowledge. Visitor experiences with the Reactable would need to be designed so as to be *prolonged* (i.e. lasting for more than a few haphazard minutes) and *structured* in some way (so as to promote real exploration and learning). Computers can be a key element in this

approach to collaborative hands-on learning, and aligns with Papert [1980]'s proposals for solving problems by building artefacts (e.g. computer programs, in this case related to musical instruments). Guided activities for learning computer music in a hands-on and collaborative environment could include observing timbral differences between two musical instruments and reasoning why these occur; observing differences between a low pass filter and a high pass filter and reasoning why these differ; building simple computer music instruments; or telling stories using sounds (cf. the results about storytelling in §6.8.3).

6. ***Short-term and longer-term one-off interactions***. We found groups who interacted with the Reactable over short periods of time. However, there were other groups who interacted for longer periods and who started to develop skill, as seen in vignettes 5 and 7. This indicates that there are both casual and more interested visitors to hands-on science museum's exhibits, and so there is potential for a longer-term practice-based culture of "learning there" (e.g. in science museums).

7. ***Self-explanatory TUIs***. Arguably, the tangible interface itself needs to be rethought if we aim to better support the process of scientific enquiry in TTIs: not only exploring and discovering, but also interpreting, predicting, testing ideas or making links. For example, Reactable-like objects that have both symbolic markers for computer tracking and icons for human recognition, can confuse beginners as seen in vignette 6. Enhancing human-readability of the tangible objects can help naïve users to better understand the mappings between the physical and the digital. Explorations not only with graphic design, but also with haptic design, seem promising [Kaltenbrunner, O'Modhrain, & Costanza, 2004]. For example, objects could have symbolic markers that are both readable for computers and humans, so objects would be tracked more continuously; or categories of objects could have different haptic properties that could help in the understanding of the tangible objects and their role. A more self-explanatory TUI can thus promote a better understanding of the system over one-off interactions.

5.9.4   Study limitations and future work

We here discuss study limitations and potential future work from this research.

Generalisability. A better understanding of social interaction with tabletops and TUIs in public settings can strengthen the design of these systems for real-world environments. Musical tabletops combine the properties of tabletops and TUIs applied to the music performance domain, which can promote rich and complex interactions. In particular, a complex tabletop TUI for music performance such as the Reactable can tell us about casual social interaction in public settings and how to best support change and transitional spaces between groups in public environments.

A single exhibit. Furthermore, our study only investigated social interaction focusing in group interactions with a single exhibit, rather than conducting a comparative study with other exhibits. We argue that the Reactable can generate a rich stream of data from visitors' interactions. The Reactable is useful for the type of exploratory research we are conducting here. Future research can look at social interaction and workspace awareness from a more institution-led approach, considering whether and how visitors are involved in processes of enquiry related to either a set of exhibits, or the whole centre, as other scholars have proposed [Hornecker & Stifter, 2006]. This approach is inline with Walker's *visitor-constructed trails* using mobile digital technologies [Walker, 2010], an activity that supports processes for assessment of learning (*meaning making*) in and from museum's visits. Particularly research could focus on a section about science and music, in resonance with Oppenheimer [1968]'s vision of a section with musical instruments within the Exploratorium. The Reactable could be placed next to other interactive exhibits focused on showing the physics of computer music sounds, and thus enable visitors to make connections.

Available headphones. Another possible problem is that our study was conducted using three out of the four available headphones. However, the majority of groups were pairs, and so the limitation in the number of headphones

was rarely a constraint. It is noticeable that the built-in loudspeakers allowed groups to share sound without wearing headphones, as well as had a role as mediators to a more immersive experience.

Future work. For future work, an in-depth study of social interaction with TTIs among co-located asynchronous groups would be useful for understanding alternative approaches to co-located collaborative music. Analysing the sound output recorded could complement video data. Additional open questions include: how best to support workspace awareness in short-term one-off interactions with a musical tabletop; and whether a more immersive environment would complement the use of headphones, without disturbing nearby exhibits. These questions are explored in Chapter 6 of this thesis. Finally, we suggest four further possible studies as follow-on work to extend the findings of this study:

1. Other approaches to tangible music in public settings (e.g. portable devices) as a platform for democratic settings of collaboration and social interaction.

2. Design and assessment of tabletop and TUI features for enabling fluid transitions and smooth overlaps between groups.

3. Design and assessment of hands-on activities centred on promoting hands-on scientific enquiry of computer music using TTIs and TUIs.

4. Design and assessment of inclusive and accessible musical tabletops considering the above features.

5.10   Summary

This chapter described a study in which the interactional behaviour of casual visitors with the Reactable in the INTECH, a science centre, was observed and analysed. We were interested in exploring the nature of social interaction and how workspace awareness was manifested in a science museum. We videoed group interactions unobtrusively over a weekend, for later

data analysis. The findings confirmed that the hands-on exhibit attracted a range of varied visitors, including adults and young people. We found that there were groups who interacted for a longer period of time than the expected one-off interactions of up to 4 minutes (short-term) or 15 minutes long (longer-term) reported in the literature. We also found that the musical tabletop tended to promote social interaction not only within-group but also between-group. In between-group interaction, there were instances of synchronous and asynchronous collaboration between groups. We argued that designing tabletops and TUIs for public spaces includes consideration of how to support fluid between-group interaction by enabling liminal spaces and smooth group change. However, the findings also revealed that naïve groups remained in the exploratory phase of a potential scientific enquiry process, when operating computer-based interactive exhibits. Exceptions were expert users or visitors who developed some skill. One limitation of the study findings is the difficulty of measuring the extent to which naïve groups learned computer music concepts. Finally, we discussed the findings' implications for designing musical tabletops that can enable transitional spaces for casual social interaction and can improve hands-on learning of computer music in public settings. We also raised some questions about how a process of enquiry for hands-on collaborative learning can be best adapted to tangible music.

The next chapter investigates the effects of spatialisation auditory feedback on tabletop collaboration during one-off interactions, which builds on the findings in this and the previous chapter's studies.

Chapter 6

Study 3: Supporting Tabletop Collaboration and Awareness with Ambisonics Spatialisation for Heterogeneous Groups

This third and final study investigates the effects of providing auditory feedback using ambisonics spatialisation on tabletop collaboration, in particular workspace awareness and group interaction. In groups of four, participants with mixed musical backgrounds were asked to improvise collaboratively on *SoundXY4: The Art of Noise*, a TTI that includes sound samples inspired by Russolo's taxonomy of everyday noises. In this study, we discuss how using ambisonics with everyday sounds on a tabletop interface can deliver realistic, immersive musical experiences supporting both individual and group awareness, in an ecological setting.

6.1 Introduction

As a follow-up to Chapters 4 and 5, this chapter aims to explore the possibilities that auditory feedback, in combination with visual feedback, can

bring to teamwork processes in musical tabletop interactions, focusing on the overall group experience. In particular, this study investigates the affects of using ambisonics spatialisation on group collaboration and awareness during musical improvisation with everyday sounds on a TTI. The motivation is to support collaborative work in the early stages of manipulating tabletop interfaces, including both individual and group work. We adopt an unobtrusive, low-cost, and ecological approach that promotes an immersive group experience with beginners and experts, and with no need to use headphones. We provide a direct connection between the location of the object in the physical table space (represented by real-time visual feedback), and the location of the object in the physical space of the room (represented by a sound field using real-time auditory feedback). Our goal is to reduce awareness issues in the early stages of interaction with musical tabletops, as identified in Chapters 4 and 5. We hope this approach will provide a better understanding of why unobtrusively delivering information about the position, orientation and identity of objects in space matters.

We designed this study with co-located small groups using SoundXY4, a basic tabletop tangible prototype for music performance that implements *ambisonics spatialisation* for four loudspeakers. Ambisonics spatialisation is a surround sound technique that can be implemented with at least four loudspeakers [Gerzon, 1974a]; see §6.2.2 for more information about this technique. We compared two conditions: spatialisation (SP) vs no-spatialisation (NSP). Our expectation was that the SP condition, compared to the NSP condition, would result in a heightened awareness of the identity and location of objects in the tabletop's workspace; and novel uses of space within it. The sounds used for this study were based on Russolo's 1913 taxonomy of urban sounds [Russolo, 1913/2004]. We conducted the experiment with eight groups of four people. Figure 6.1 shows a group interacting with the tabletop system.

The results revealed that, using spatialisation, there was a clearer awareness of individual vs group actions, and a more musically immersive and realistic experience. Our aim is to support better usability design in tabletop research

FIGURE 6.1: Four participants interacting with *SoundXY4: The Art of Noise*

based on the potential uses of auditory feedback using ambisonics. In particular, we discuss the design implications of using this approach for tabletop design, interactive exhibit design, sonic design, and for musical education.

6.2   The system: SoundXY4

In this section, we provide a rationale for developing SoundXY4. First, we describe its implementation in terms of hardware and software, paying particular attention to the implementation of ambisonics spatialisation. Second, we explain the rationale for the categories of sounds used. Third, we outline the usage of the system.

6.2.1   Rationale

Studies 1 and 2 pointed to WA issues between individual and group work in early interactions with the Reactable. We wanted to address these issues using a non-intrusive approach, designed for both beginners and experts, and using a TTI based on constructive building blocks. Our approach was to work with everyday sounds and their familiarity from the real world. We were interested in the potential link between the spatial dimension of everyday sounds and the spatial position of the tangible objects on the tabletop surface.

The aim was to facilitate WA of individual and group work by providing non-intrusive auditory information about the position of the objects that could be easily understood by both beginners and experts.

SoundXY4 was designed to facilitate collaborative musical performance using a *surround sound system* with four loudspeakers, providing information on the location of tangible objects on a tabletop surface, with each performer receiving information from all of the speakers. We aimed to promote a group experience, in contrast to the more individualistic experience of using individual speakers or headphones for each of the listeners. This approach can be useful to support awareness of both individual and group actions on the tabletop surface using the same sound field for everybody, i.e. with no disruptions as would occur when using headphones. As seen in Chapter 5, using headphones limits the number of musicians that can interact together under similar conditions. It also physically isolates individuals from communicating with their collaborators, as well as isolating them from the sound of the environment. SoundXY4, however, offers a more ecological approach, permitting a free number of musicians within a small group to explore and discover together a set of sounds during a music performance, irrespective of their individual position next to the table. Also, using spatialised everyday sounds is aimed at supporting collaboration among people with different musical backgrounds.

SoundXY4: The Art of Noise was a celebration of the centenary of Russolo's 1913 *The Art of Noise* [Russolo, 1913/2004] and his taxonomy of sounds, which included urban noise and industrial sounds. We believe that Russolo's material, combined with suitable technologies, can be useful for music performance by users both experienced and inexperienced in music making. On the other hand, everyday sounds, such as the noises proposed by Russolo, are traditionally associated with *everyday listening* [Gaver, 1993], which implies paying greater attention to the spatial dimension than with traditional musical listening (see Westerkamp [2002] for an exploration of everyday listening in electroacoustic music). Russolo's ideas have influenced the NIME community, such as in the Croaker system [Serafin et al., 2006], or aesthetic

reflections on using the loudspeaker as a musical instrument [Mulder, 2010]. However, as far as we know, exploring Russolo's taxonomy of sounds on a tabletop system is novel.

6.2.2   Implementation

We developed a tabletop tangible interface for music performance using open source technologies. The SoundXY4 code is publicly available at Github,[1] and can be replicated, examined, extended, or adapted for other work. The design approach was based on the Reactable [Jordà et al., 2005] principles of: 1) a collaborative interface based on constructive building blocks modelled by physical objects on the tabletop that trigger or modify sounds; and 2) the lack of division of the interface into territories, as a mechanism for promoting the self-regulation of spaces, as shown in Chapter 4. However, SoundXY4 focuses on: 1) using only sound samples (i.e. no sound synthesis is used); and 2) delivering audio using ambisonics with four speakers, in order to support an overall group auditory experience that relates the auditory location of the sound in the sonic field to the position of the objects on the tabletop.

We developed SoundXY4 as a TTI for music performance. The software was written in SuperCollider, which allows the easy implementation of networks of sounds and effects in real time. We used SETO Quark,[2] which implements the TUIO protocol by Till Bovermann, and which allows SuperCollider to communicate with reacTIVision. Figure 6.2 provides an overview of the system.

SoundXY4's implementation was a follow-up of the tabletop waveform editor waveTable, a collaborative work with Gerard Roma [Roma & Xambó, 2008]. Both programs are based on a collection of classes that represent audio tools (*sound players* and *filters*), which also can be interconnected. SoundXY4 was constructed using the same *MVC (Model-View-Controller)* architecture

---

[1] http://github.com/axambo/soundxy4 (accessed 30 September 2014).
[2] http://sourceforge.net/p/quarks/code/HEAD/tree/seto (accessed 30 September 2014).
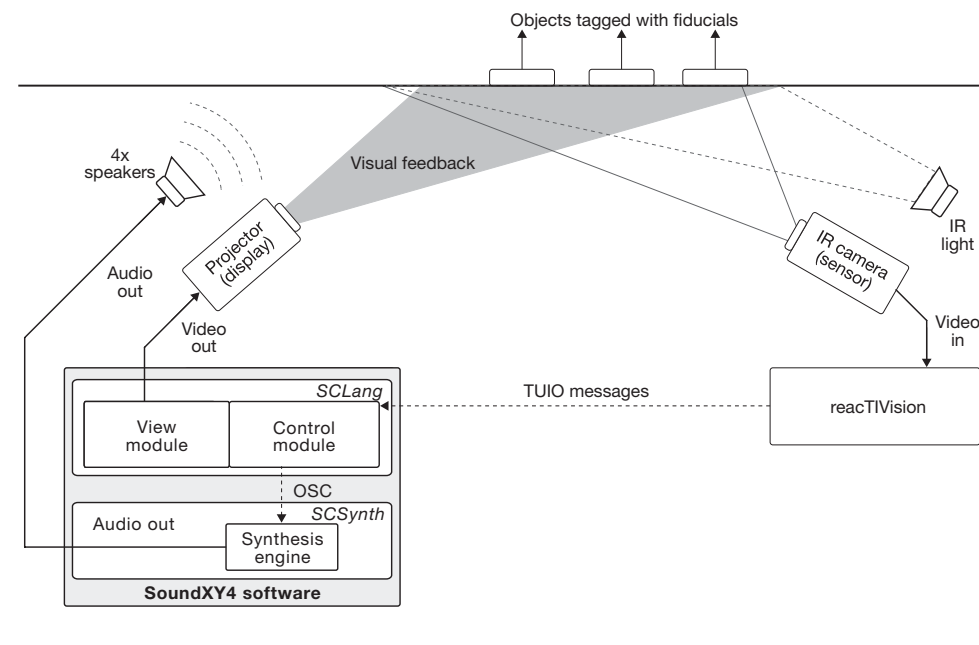
FIGURE 6.2: System overview of SoundXY4

as waveTable. The structure of the program was divided into three modules: Model, View and Controller, following the standard *MVC pattern*, and adapted to the SuperCollider architecture and its available classes.

In particular, the structure of the code consisted of four core classes: *XY4* (main and Model), *XY4Object* (Model), *XY4Controller* (Controller), and *XY4View* (View), as shown in Fig. 6.3. The Controller module *XY4Controller* classified the objects as either sound player objects or effects, and tracked their behaviour in terms of presence, location, or rotation. The Model modules *XY4* and *XY4Object* managed the *buffers* to play the sound samples; operated the relationship between the sound player objects and effects according to their vicinity; and controlled the order of execution of the *synth nodes* on the server including the synths for the sound samples, effects, and spatialisation. The View module *XY4View* managed the visualisation of the sound player objects and effects.

By invoking an instance of the class *XY4*, an instance for each of the other three classes was created. In order to distinguish between the sound players and filters, two classes were inherited from the *XY4Object* class:
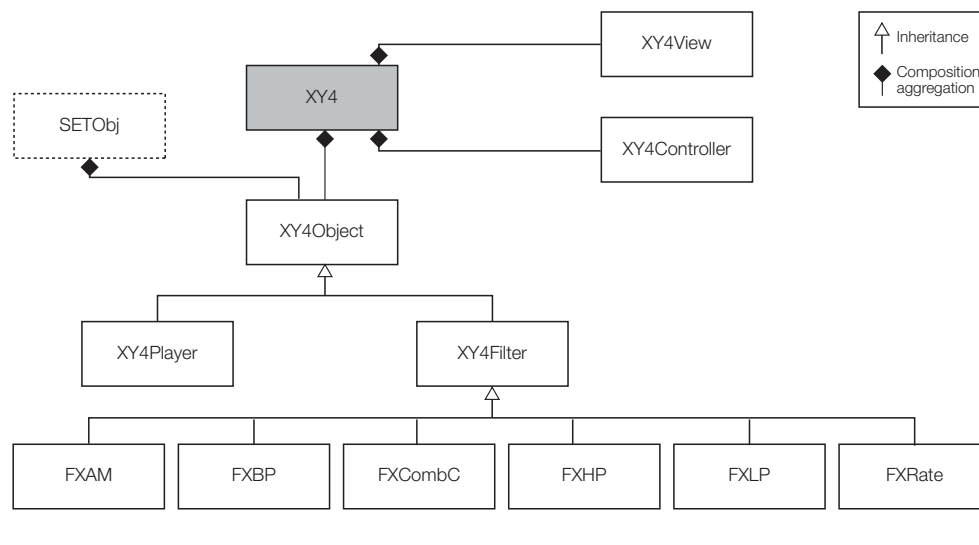
FIGURE 6.3: Class diagram of SoundXY4

XY4Player and *XY4Filter*. Finally, six more classes were inherited from the *XY4Filter* class: *XY4FXAM*, *XY4FXBP*, *XY4FXCombC*, *XY4FXHP*, *XY4FXLP*, and *XY4FXRate*. Each of these implements one filter. Figure 6.3 shows a class diagram of the application.

SoundXY4 tracks the identity, position and orientation of tangible objects tagged with reacTIVision fiducial markers [Kaltenbrunner & Bencina, 2007] and maps them to sound players or effects. In this study, there was a subset of 36 different fiducials that were mapped to 36 different sound samples. The sounds were grouped into six categories, inspired by Russolo's six families of noises (see §6.2.3). We used physical cubes with a unique marker on each side of the cube. Thus, each side represented a sound, and each cube represented a *sound category*. The sounds were played in a loop. We used another group of four different cubes, one for each team member, with six different filters repeated in each cube for modifying the sound of the sound players. The filters used were: a band pass filter (*BPF*), a resonant low pass filter (*RLPF*), a high pass filter (*HPF*), a comb delay (*CombC*), a pitch shifter (*PitchShift*), and an amplitude modulator (*AM*). Appendix D.3 gives further details about the filters implemented.

The application mapped the sounds and filters to a 2D plane using *ambisonics spatialisation* [Gerzon, 1974b] for four loudspeakers. Ambisonics is a surround sound system, providing sound from a full circle (360°) using sound localisation cues for enhancing the perception of sound, allowing performers to identify where the sound comes from, and over what distance. Appendix D.4 further explains this technique. We opted for a surround sound system because it offers an integrated group, as well as individual, musical experience, where otherwise headphones could disrupt the group experience. In particular, we chose the ambisonics technique, because, among low-cost systems, it is the most sophisticated in terms of realism compared to other systems such as quadraphonics or stereo panning [Gerzon, 1974b]. We implemented the technique in the basic horizontal surround set-up, using four speakers, because it remains an easy and cheap set-up, and maintains direct mapping with the tabletop interaction. The four loudspeakers were each positioned at 90° to the table, with each facing a corner of the table, as illustrated in Fig. 6.4. We used a SuperCollider implementation of ambisonics, divided into two unit generators: the 2D encoder *PanB2*, which analyses the sound field; and the 2D decoder *DecodeB2*, which synthesises the sound field.

6.2.3   Taxonomy of sounds

The choice of the sounds followed the six families originally proposed by Russolo. From the original text, it can be seen that the second level of the taxonomy is formed with characteristic sound types: "we have included in these six categories the most characteristic fundamental noises: the others are hardly more than combinations of them" [Russolo, 1913/2004, p. 10]. We gathered examples of these categories from Freesound.org [Akkermans et al., 2011], a free online database of sound recordings. For the precise selection of sounds, we consulted Gerard Roma, an expert on the online library. The main criterion was to find iconic sounds for each of the definitions in the text, but at the same time to avoid overlaps. Following this idea, we extended our search where necessary in order to ensure six sounds per category, trying to carefully find sounds that were clearly distinct and yet all representative of
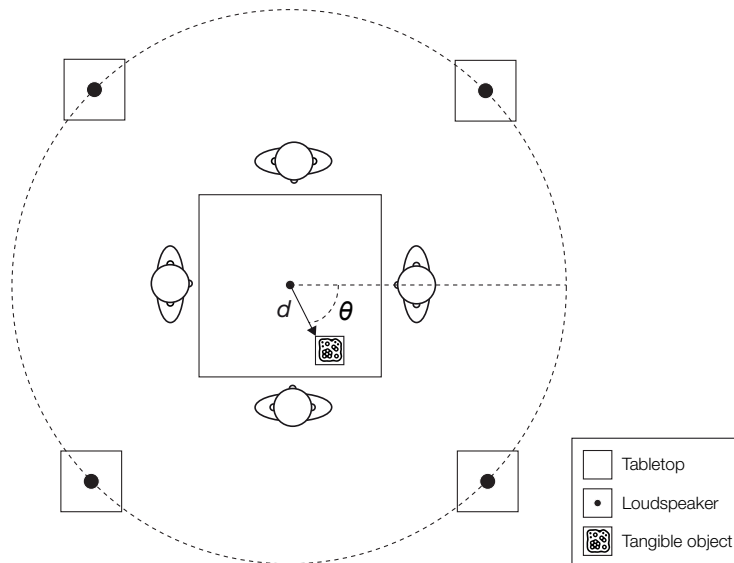
FIGURE 6.4: Top view of an imaginary circle representing ambisonics on a tabletop system. The sound output depends on the angular position (θ) and distance (d) of the tangible object with respect to the centre of the table, irrespective of the object's rotation

TABLE 6.1: Used sounds by category

Category	Used sounds
Explosions	Water splash, boom rumble, car crash, war boom explosion, thunderstorm, gun shot
Percussion	Tom drum, rocks smashing, wood impact, ceramic impact, falling metal, metallic impact
Screeches	Metallic-ceramic friction, fire crackle, plastic rustling, lamp buzz, howl, metallic whine
Voices	Tibetan chant, man laugh, child laugh, scream, cat howling, groan
Whispers	Wind mouth gurgle, wind mouth whisper, cat purring, stomach gurgling, group whispering, crowd murmur
Whistles	Leaking gas hiss, boat whistle, tunnel short whistle, water pipe hiss, air escaping hiss, long snort

the main category. A summary of the sounds used can be found in Table 6.1. Appendix D.1 shows a full list of the tracks used.[3]

---

[3]A selection of audio extracts mixing these sounds can be listened to here: `http://soundcloud.com/soundxy4` (accessed 30 September 2014).

6.2.4 Program usage

As shown in Fig. 6.5, sound players were represented by white cubes, and filters were represented by black cubes. Each white cube represented one of Russolo's categories. Each filter modified the sound of the nearest white cube. The fiducial markers worked as icons as well. Visual feedback informed users about the position of the objects, the sound category, and what sounds were affected by which filters. When a sound player cube was on the tabletop surface, it was highlighted by a coloured square. There was a different colour for each category of sound player, and the nearest filter was highlighted in the same colour to indicate the sound that it was affecting. The background colour of the tabletop was blue, which was used as a neutral colour. See Appendix D.5 for information about the colours used, and D.6 for further information on the physical design of the program.

It was possible to change the volume for each cube by rotating clockwise (volume up) or anticlockwise (volume down). There was a projected mark of the value in the range (0–100), and also a number for each of the six sounds (1–6). Rotation of the black cube modified a parameter of the filter, whose value was indicated with a number next to the object from 0 to 100 (from min to max). The rotation and position of the white cube on the table surface defined the position of the sound source in the room, which was related to the position of the loudspeakers. For example, if an object was positioned in the centre of the tabletop surface, the sound would come equally from all the loudspeakers; as it would be positioned at the same distance from all the four loudspeakers. If the sound was positioned in one of the corners of the tabletop surface instead, then the sound would emerge differently from the loudspeakers; becoming louder from the loudspeakers in the region closest to the object. In a similar way, the filters had the same spatialisation effect as the sound player object they were affecting.[4]

---

[4]See video 6-1A listed in Appendix B for a video demo of this version.
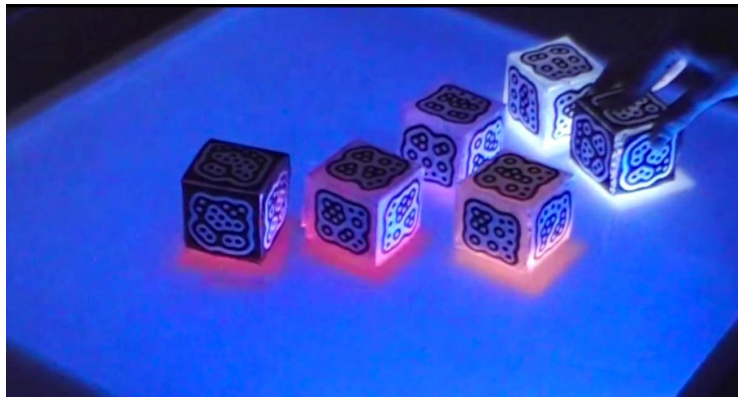
FIGURE 6.5: Closeup of SoundXY4

6.3 The study

This study investigates how the application of the ambisonics spatialisation technique influences the nature of group tabletop interactions (GTI) within music performance in its short-term use by heterogeneous groups; along with how workspace awareness (WA) is manifested in these group interactions.

6.3.1 Study design

The study comprised two conditions, the SP and NSP conditions:

1. **Spatialisation with ambisonics (SP)**: Collaborators worked on a table-top surface surrounded by four loudspeakers, and, depending on the position of the tangible objects on the tabletop surface, the sound output varied according to the sound localisation cues.

2. **No-spatialisation (NSP)**: Collaborators worked on a tabletop surface surrounded by four loudspeakers, where the position of the tangible objects on the tabletop surface did not affect the sound output, which was the same for all four loudspeakers.

The study used a *within-subjects design* (groups being 'subjects'), in which the groups took part in both conditions. The order was randomised; whereby, the

conditions were presented in a random order in each group, in order to control possible learning effects of performing better in the second condition. In this study, an advantage of a within-subjects design compared to a between-subjects design was that fewer participants were needed to test all the conditions, which in our case facilitated the challenge of recruiting participants to form groups of four.

6.3.2 Participants

As we were interested in observing the level of collaboration, we decided to conduct the experiments with groups of four people. Our scope was to work with small groups as explained in Chapter 1, and we decided that with a square table, and four loudspeakers, four was an appropriate group size. To our knowledge, most controlled experiments on collaboration with tabletop interfaces typically work with triads [Hornecker et al., 2008; Nacenta et al., 2010], including experiments with auditory feedback using audio icons [Hancock et al., 2005] or when using a sound surround system [Blaine & Perkis, 2000]. Thus, part of our contribution to this field is to conduct this study using auditory feedback with a larger small group than has been done typically.

Eight groups of four participants, 32 people in total, participated in the study as volunteers; made up of 13 females and 19 males, aged from 15 to 57 (M = 33.13 years old, SD = 11.91). All groups were gender mixed. Participants were also mixed in terms of years of musical training, including 8 beginners ("none"), 12 intermediates (8 with "1–2 years", 2 with "2–4 years", and 2 with "4–6 years"), and 12 experts ("more than 6 years"). All the groups were mixed level groups: one group with *beginners* (B), *intermediates* (I), and *experienced* (E); three groups with B and I; three groups with I and E; and one group with B and E. Some of them knew each other (7), others knew some of their group (8), and some didn't know each other at all (17). Participants were international (4 from Asia, 26 from Europe, 2 from North America—in all representing 16 different countries). Groups and participants were anonymised

to G# for groups and P# for participants (e.g. G1 P4 refers to participant 4 of group 1), which is the nomenclature used henceforth.

6.3.3 Set-up

We carried out the study in the Ambient Lab[5] at the Jennie Lee Building (Open University, Milton Keynes, UK), a multipurpose and large space ($7.6\,\mathrm{m} \times 16\,\mathrm{m} \times 2.47\,\mathrm{m}$) for controlled experiments. The lab has eight dome cameras, which tend to go unnoticed by participants. It also has tall display screens ($0.995\,\mathrm{m} \times 2.25\,\mathrm{m}$), which can be set up to partition the space into different configurations, such as for building smaller 'rooms'.

As shown in Fig. 6.6, we built a closed room with the display screens in the middle of the Ambient Lab, below four of the dome cameras, which could cover the centre of the closed area from the four different extremes of it. In order to improve the acoustics of the closed area, we positioned the screens to avoid a perfect square, and instead, we created acute and obtuse angles at the joints of the display screens. The tabletop system was positioned in the middle of the closed space, with four loudspeakers at a distance of $1.2\,\mathrm{m}$ from the table, each one facing one of the corners of the tabletop surface, at $90°$ from each other, and with two additional floor standing cameras on tripods. Figure 6.7 shows the set-up. These details are important in order to facilitate the replicability of the study. Next, we present the configuration of the six cameras, and the four loudspeakers.

We used four of the eight available dome cameras to get a top and wide perspective of each corner of the tabletop surface, and two more floor-standing cameras on tripods pointing to the two opposite sides of the tabletop for a closer perspective. The two tripod cameras were used to complement the top perspective, with the aim of getting a detailed capture of the group interactions. Later, in the video editing station we synchronised the six video data sources into a single file to facilitate the video analysis process. We added a

---

[5]`www.open.ac.uk/about/campus/jennie-lee-research-labs/about-the-labs/labs-and-rooms/ambient-technology-laboratory` (accessed 30 September 2014).
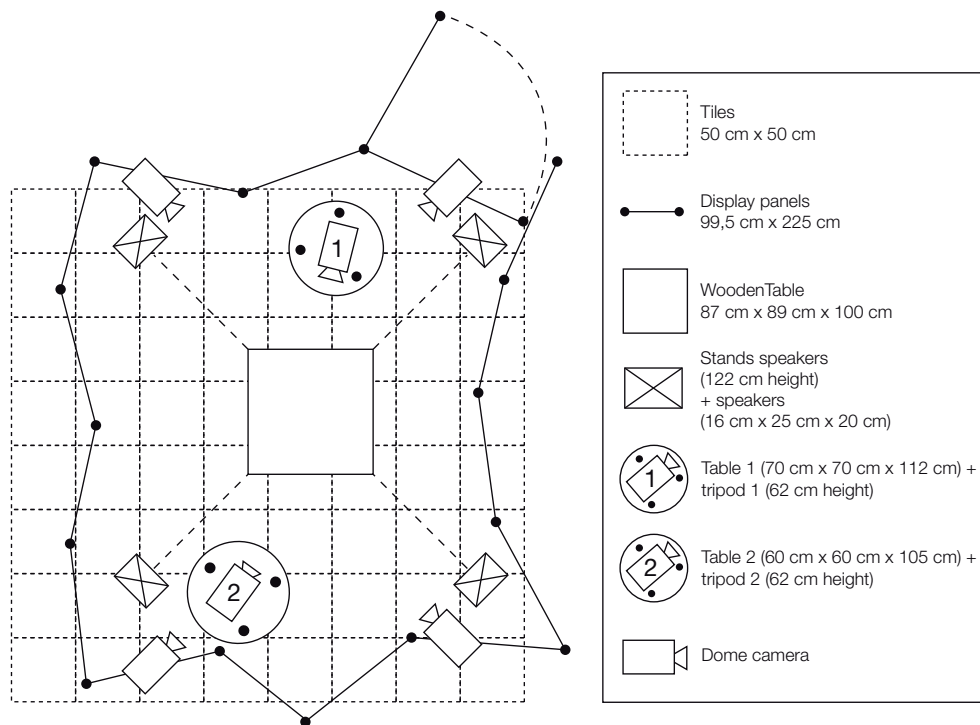
FIGURE 6.6: Top-view diagram of the SoundXY4 experimental set-up
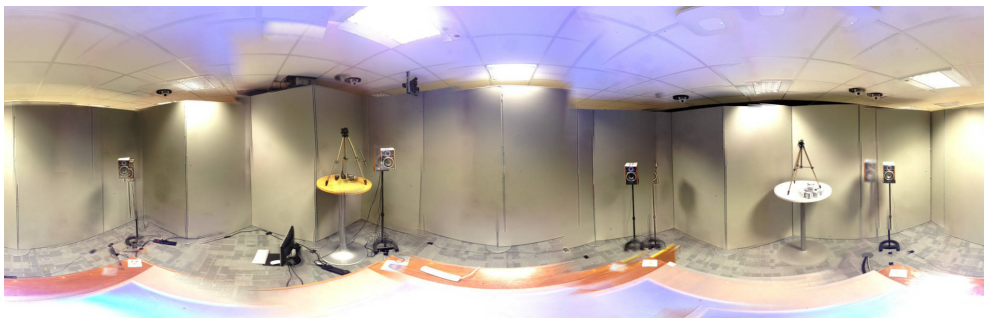
FIGURE 6.7: 360º panoramic photo of the SoundXY4 experimental set-up.
Photographer: © Dave Perry (2013)

number next to each corner of the table rim (1–4) to facilitate the video syn-
chronisation and analysis. Figure 6.8 shows a snapshot of the six different
cameras views, and Figure 6.9 shows a single view of the edited video file for
further analysis.

Loudspeakers were elevated at a height of $1.22\,\text{m}$, which we considered the
average height above the ground of a (standing) human's ear. The volume
was adjusted to comply with the Health and Safety Regulations of the UK
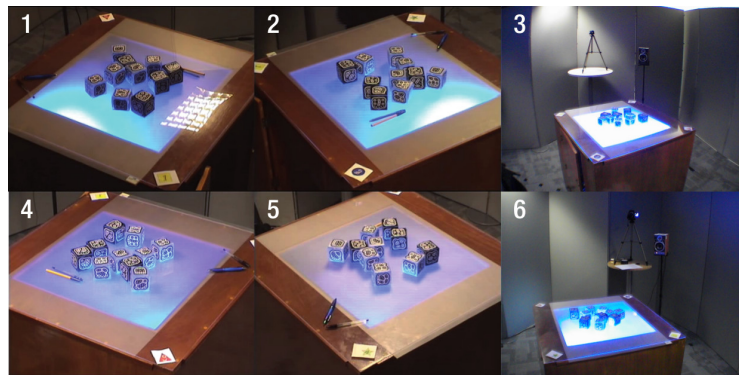
FIGURE 6.8: Views of the six cameras used in data collection: 1. Corner ceiling camera (#1 yellow square); 2. Corner ceiling camera (#2 blue circle); 3. Panoramic camera; 4. Corner ceiling camera (#3 red triangle); 5. Corner ceiling camera (#4 green star); 6. Standard camera
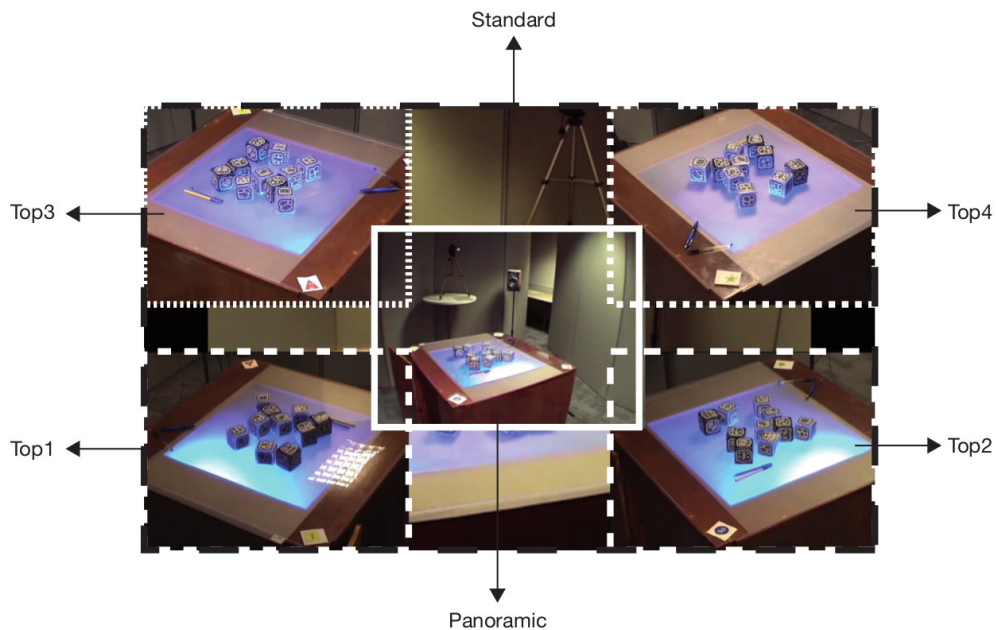
FIGURE 6.9: Video composition for the video analysis

government so the maximum overall sound level was $65\,\text{dB}$–$70\,\text{dB}$. We performed all the measurements with a sound level meter.
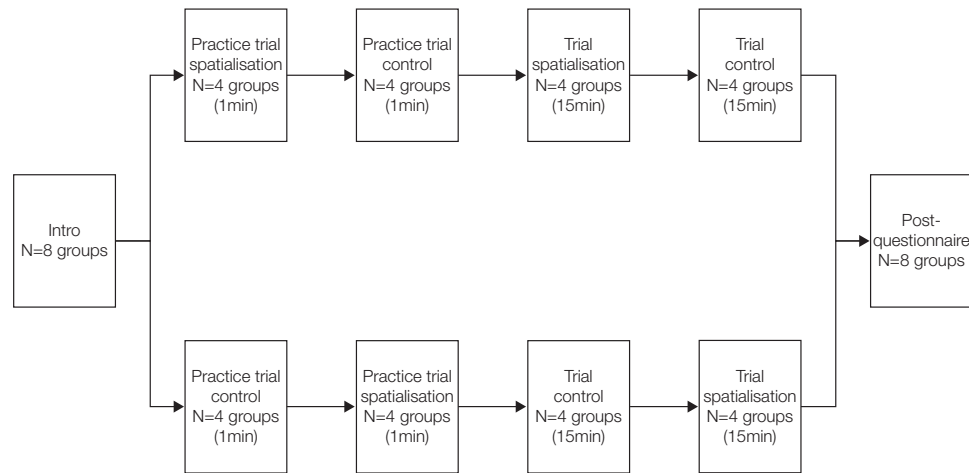
FIGURE 6.10: Experimental design flowchart of SoundXY4: within-subjects design in a randomised order across groups. The two branches represent the two conditions of the experiment

6.3.4   Experimental procedure

We conducted the experiment over five days in July 2013. We first asked participants to sign the consent form (see Appendix A.5.1). After a short introduction to the activity, participants were asked to improvise music with the tangible objects available, by playing twice with the sounds and filters as part of the group for 15 minutes. As explained in Chapter 3, musical improvisation is a useful mechanism for bringing together mixed groups of beginners and experts, who may be strangers, and getting them to interact with each other with no written rules, as happens in real contexts such as in public spaces or in certain musical performances. Before starting the trials, each group performed two rehearsals in the same randomised order as in the experiment, so that participants could familiarise themselves with the interface, in line with Hancock et al. [2005]'s study. After completing the experiment, the participants were asked to fill in a background questionnaire (see Appendix D.2). Figure 6.10 shows a flowchart of the whole procedure. Before this study took place, we did a pilot study with two groups to test the procedure.

After a short explanation of the activity outside of the panel room, the facilitator invited participants to enter the room. Participants were allowed to choose their positions at the table, but for a better experience, they were recommended to each stand in front of one side of the table (not at the corners). Once all were positioned, the facilitator brought in a tray with a collection of 10 objects for them to pick. As explained in Chapter 3, the facilitator tried to be as unobtrusive as possible, only entering the panel room for setting up and closing each of the rehearsals and trials. For this reason, during the introduction, participants were advised to call out for the facilitator in case of unexpected problems. In the course of the rehearsals and trials, the facilitator exited the panel room and moved out of the participants' visibility, but still remained in the lab to make sure that the experiment followed schedule and to control the situation in case of need.

We collected video recordings and interaction log files of the user interactions with the system during the performance of the task, which both demonstrate how the groups interact with the system.

6.3.5   Method

We analysed: 1) verbal communication (VC), from the videos, and 2) non-verbal communication (NVC), from the videos and log files. We quantified the proportion of verbal communication versus non-verbal communication for each group. We analysed the video data by focusing on the types of conversations relating to tabletop awareness, and the types of non-verbal behaviours relating to space use. We also analysed the log data relating to space use.

We thoroughly watched the videos (four hours of video data) and took notes about the emerging themes related to our research questions on workspace awareness and space use. To gauge workspace awareness, we transcribed the verbal communication, so that we could clearly identify when there was an understanding or lack of understanding of the activity. We aimed to identify emerging themes related to WA, such as the identification of sounds

or expressions of musical immersion, among others. From these transcriptions, which were substantial, we classified the different quotes into emerging themes and refined the themes as needed. We counted the occurrences by theme to obtain descriptive statistics, and chose representative examples (i.e. *vignettes*) to present each theme.

To gauge the space use relating to GTI, we transcribed the non-verbal communication from the videos and analysed the log files. For example, we quantified the number and type of transitions between work next to the individual and work in the centre of the table from the video data. We chose a few illustrative vignettes for providing a complementary qualitative insight. In terms of the log files, SoundXY4 logged *activity maps* of the sessions.

We used Elan software for organising and classifying the video material in the groups and in the two conditions. The transcriptions were done using a standard note-taking software, as this was the most straightforward approach. We analysed the interaction log files using R scripts and t-test statistical analyses (see Appendix A.6).

6.4   Coding scheme

We here present our coding scheme, which includes the terms most used in our video analysis. The terms are divided into verbal (§6.4.1) and non-verbal communication (§6.4.2). We also found differences between the groups in terms of how talkative they were:

Storytellers vs music performers. Data showed that groups adopted different approaches to musical improvisation. If we imagine a continuum, at one extreme there is the *storytelling approach*, which refers to telling stories based on the sounds; whereas, at the other extreme of the continuum, there is the *music performance approach*, which refers to a more non-talkative attitude to musical improvisation. An important aspect discussed in this chapter is group musical immersion, which is evidenced from the speech act occurrences. However, using video analysis rather than direct sight makes it difficult to tell whether

the non-talkative groups had a different level of musical immersion in the two conditions.

6.4.1 Workspace awareness coding scheme

From the verbal communication (VC) analysis of the video data, we identified a set of themes not mutually exclusive related to workspace awareness, and which are defined in more detail in Appendix D.7. Each relate to specific types of speech act that we observed during the experiment, including:

- *Identification of categories*: where, a particular group of sounds of a cube was explicitly identified using general concepts (e.g. [We hear the sound of a ceramic impact, a tom drum, a falling metal, a metallic impact] "These are all percussion sounds") or characteristic attributes (e.g. "Mine ['screeches' white cube/sound] is really nasty").

- *Impressions of sounds*: where, the first reactions of identifying the sound included feelings (e.g. "I don't like that"), and attributes (e.g. "That's a horror film now").

- *Identification of source*: identifying the sound source using general concepts (e.g. [We hear the sound of a cat howling] "A kitty! A kitty!") or characteristic attributes (e.g. [We hear the sound of a man laughing] "You've got the evil laugh").

- *Non-identification of person or object*: an explicit question or comment about who or what produces the sound (e.g. "Oh who's got that one?").

- *Non-identification of source*: an explicit question or comment about the sound source (e.g. "Yeah, I can't hear this wind at all" [while lifting up a white cube]), or when a sound source is clearly misunderstood (e.g. [We hear a sound of a purring cat] "Oh, that's the ... crickets!").

- *Identification of filters*: where, a particular effect or collection of effects (e.g. with the black cube) is noticed (e.g. "There is a difference!"),

identified (e.g. "There's an echo"), or described (e.g. "Uh that's very spooky").

- **Non-identification of filters**: an explicit question or comment about a particular effect or collection of effects (e.g. with the black cube) because of a lack of understanding about how it works (e.g. "What are these modifiers supposed to do?').

- **Stories, realism, and ambisonics**: where, a connection was made: 1) between sounds or associations (e.g. "sounds like a teapot"); 2) to storytelling (e.g. "A snoring man with a purring cat by the fire"); 3) to musical composition (e.g. "We've arrived, we are doing Stockhausen now"); 4) to realism (e.g. 'It is a bit more realistic the snoring..."); or 5) to spatialisation (or lack of it) in terms of associations with the position of the sounds and the position of the speakers (e.g. "I wonder whether it is interesting where we put these [tangible objects on the table] (...) we created a geographical representation as well as a sound one"). These five subtypes are not mutually exclusive. From all the identified themes on WA, *Stories, realism, and ambisonics* appeared to be the most substantial and nuanced. From the participants' quotes classified in this theme, we developed the following classification of subtypes for the comments, though they are not mutually exclusive, and are defined in more detail in D.3:

    - *Sound associations*: comments that connect or associate the sounds to real or imaginary characters (e.g. "ghosts"), sound sources (e.g. "evil laugh"), situations (e.g. "sounds like a supermarket"), or abstract concepts (e.g. "underground noises").

    - *Storytelling*: comments that narrate or build stories connected to the sounds beyond a mere description and qualification of the used sounds (e.g. "the snoring man with a purring cat by the fire"), including connections to memories or lived situations (e.g. "do you remember this song called popcorn?").

    - *Musical composition*: comments about music-related stories, in particular references to musical roles (e.g. "I am gonna give you guys

this rhythm), musical style (e.g. "we've arrived, we are doing Stockhausen now), or musical structures (e.g. "we can keep the ghosts but we can decrease the volume").

– *Realism*: comments that qualify the sounds, effects or the overall sound output as realistic (e.g. "a bit more realistic, it is a bit more realistic the snoring . . . ") or as setting a scene (e.g. "we created a scene that's all we wanted to do, wasn't it?").

– *Ambisonics*: comments that associate the speakers with the location of the objects on the surface (e.g. "I wonder whether it is interesting where we put these [tangible objects]"), or comments about the lack of spatialisation (e.g. "So the sound comes mainly from this speaker, yeah?").

- *Immersive musical experience*: explicit comments about enjoying the activity by characterising particular actions (e.g. "This is nice, it's a different way" [after pressing 'on' and 'off' a cube side]) or the overall sound output (e.g. "I like what we've got, it's good!"). For talkative groups, we assumed that musical immersion was indicated by both commenting more about producing enjoyable and 'real' music, as well as talking less, as individuals were concentrating more on the musical task. For this assumption, non-talkative groups were excluded, because of the lack of supporting evidence.

6.4.2   Space use and territoriality coding scheme

From the non-verbal communication (NVC) analysis of the video data, we identified a set of themes related to GTI, including:

- *Individual vs central work*: we observed that the use of territories was slightly similar to the social protocols established in the tabletop studies on territoriality [Scott & Carpendale, 2010] of personal (next to the person) and storage (rim area) territories. However, as the use of the tabletop space is also determined by the spatialisation effect, in which

the sound outcome varies depending on the position of the object, the use of the centre of the tabletop surface for group work is less obvious, as it can relate to individual exploration as well. We here analysed the use of the space next to the individual vs the use of the centre from this perspective.

- *Transitions*: we counted the number of transitions of the participants from an individual space to the central space, and the other way around. This measure was considered because, as we have seen in the literature review, supporting fluid transitions between individual and group work is important in collaborative interactive systems [Gaver, 1991]. Having said that, since the maximum use of spatialisation relies on the use of all the available tabletop space, we were aware that in this study, working in central areas did not necessarily indicate shared work. As there were four participants in each group, we defined three mutually exclusive states based on the tabletop territoriality literature [Scott & Carpendale, 2010] on the conventional distinction between individual space (the space in front or next to the person) and shared space (the space in the centre of the tabletop surface or the in-between spaces such as corners):

    - *Individual*: where, at least three participants are working in their own personal space.

    - *50% individual 50% central*: where, two participants are working in their own personal space, and the other two participants are working in the central or in-between spaces.

    - *Central*: where, at least three participants are working in central spaces (the central or in-between spaces).

TABLE 6.2: Summary of general group behaviours by condition

Group	Condition	Non-participation	Explorations	Concern markers
G1	NSP	-	-	-
	SP	●	●	-
G2	NSP	●	-	●
	SP	●	-	-
G3	NSP	-	●	-
	SP	-	-	-
G4	NSP	-	●	●
	SP	-	●	●
G5	NSP	●	●	●
	SP	●	●	-
G6	NSP	-	-	-
	SP	●	-	-
G7	NSP	-	●	-
	SP	-	-	●
G8	NSP	-	●	-
	SP	●	●	-

6.5 Results: overview

As in our previous studies (Studies 1 and 2), we found instances of non-participation, physical explorations of the tangible objects beyond the interactive system, and comments about the visual markers, irrespective of the condition. See Table 6.2 for a summary of these general behaviours by group and condition.

An incident occurred in three groups (G2–G4) towards the end of the last trial; whereby, after long use of the DIY tabletop system, the inside of the table reached high temperatures ($41°C$ and higher) causing the projector to turn off, and so the tabletop door needed to be opened for ventilation and the projector had to be turned on again. In all three cases where this occurred, the participants continued the experiment by continuing to play (as the sounds did not stop), and it was the facilitator who non-intrusively entered the 'room' and restarted the projector for the visual feedback. As participant P2 in G4 said "we are playing by memory". Thus, it had no serious consequences as the affected groups kept playing, and the problem was solved each time within a few minutes.
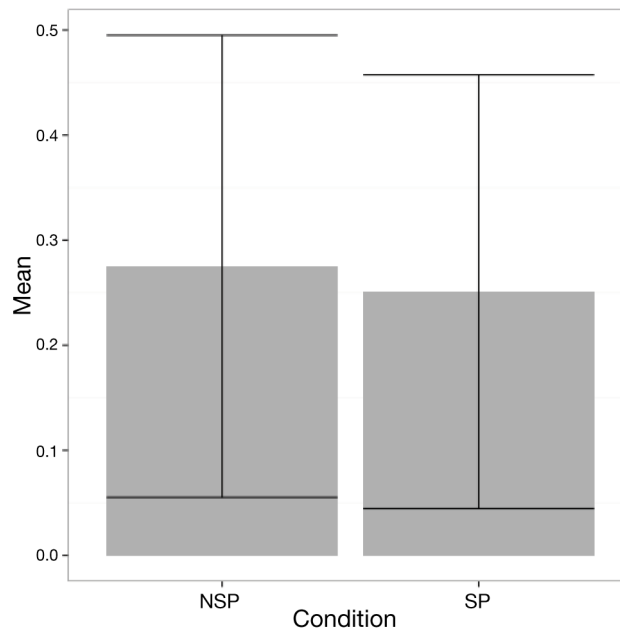
FIGURE 6.11: Bar plots, also showing the standard deviations, of the mean proportion of time spent on verbal communication across the two conditions

6.5.1   Verbal vs non-verbal communication

For each group and session, we counted the proportion of time spent on verbal communication (VC) vs non-verbal communication (NVC). As shown in Fig. 6.11, the mean proportion of time spent on verbal communication was slightly smaller in the SP condition ($M = 0.25$ proportion of time, $SD = 0.20$), compared to the NSP condition ($M = 0.28$ proportion of time, $SD = 0.22$). The standard deviation was similarly large in both conditions, which indicates that the variation from the average was large. Although there are no visible trends, the difference between the conditions suggests that with spatialisation there was less need to exchange words, which could indicate a slightly greater musical immersion. Having said that, the conversations in both conditions were related to the musical activity, so there seems to be a general interest and workspace awareness irrespective of the condition. This is further discussed in §6.6.
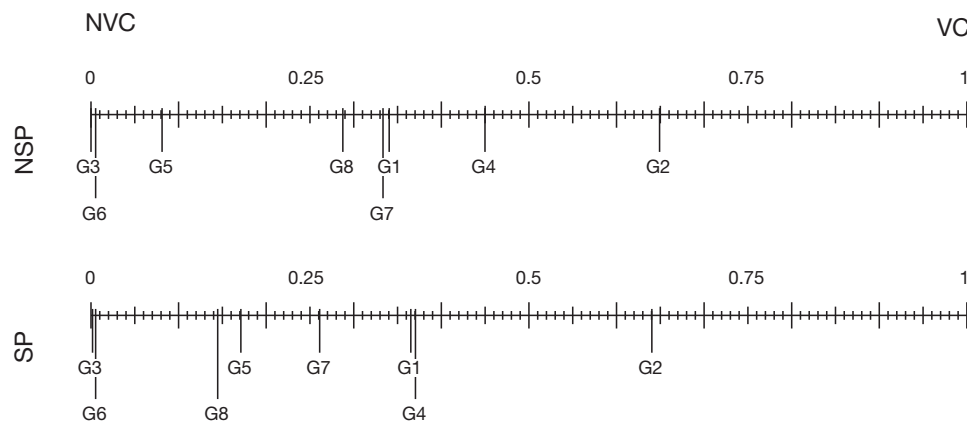
FIGURE 6.12: Continuum from music performance (NVC) to storytelling (VC)

We identified the qualitative differences between the groups irrespective of the conditions, which indicated the differences in respect to the group dynamics. For example, two out of the eight groups tended to be non-talkative. We related these variations to different group dynamics with a different understanding of musical improvisation using noises from the music performance to storytelling. As shown in Fig. 6.12, at one extreme of a continuum, G2 and G4 adopted a storytelling approach; whereas, at the other extreme of the continuum, G3 and G6 adopted a mostly non-talkative music performance approach. The other groups used both VC and NVC (mid-talkative groups), with no apparent predominance of either of the two approaches. Moreover, most of the groups seemed to enjoy the types of sounds, although two of the groups (G1 and G7) seemed to be more sceptical when evaluating the sound outcome. This suggests that the sample indeed involved a representative range of groups with different musical tastes and approaches to musical improvisation, as would be expected to be found in the real world.

6.5.2   Summary of the vignettes

In the next sections, we present the results in terms of workspace awareness and space use. Table 6.3 shows a summary of the vignettes discussed.

TABLE 6.3: Overview of vignettes

Section	Topic	Vignette
Workspace awareness	Sounds, categories, and filters	1. Identification of the sound of a purring cat in NSP
		2. Impression of the sound of a market in NSP
		3. Identification of the sound of a teapot in SP
		4. Identification of the sound of a crowd in SP
		5. Non-identification of the change in filters in NSP
		6. Non-identification the role of filters in SP
		7. Identification of the filters and the sounds in SP
		8. Identification of the filters vs the sounds in SP
		9. Non-identification of who controls the sound of a laughing man in NSP
		10. Non-identification of who owns the sound of long snort in SP
		11. Non-identification of who owns the sound of a car crash in NSP
		12. Non-identification of the sound of a cat howling in SP
		13. Non-identification of the sound of a cat purring in NSP
		14. Identification of the explosions category in SP
		15. Identification of the voices category in NSP
		16. Identification of an 'inconsistent' category in NSP
	Realistic and musical scenes	17. Creating a 'realistic scene' in SP
		18. Creating a 'soundtrack of a movie' in SP
		19. Creating a 'piece for the Olympic games' in SP
		20. Creating a 'song' in SP
		21. Creating 'actual fire' in SP
		22. Wondering about the positioning of the objects in SP
		23. Wondering about the positioning of the objects and loudspeakers in SP
		24. Wondering about from which loudspeakers the sound comes in from in NSP
Space use and territoriality	Space explored	25. Corners used for building a story in SP
	Transitions	26. Parallel individual and group work in SP

6.6 Results: workspace awareness

In this section, we use a series of vignettes to demonstrate the salient results related to workspace awareness under the two conditions. In the first part of this section, we give an overview of the verbal communication themes. In the second part of this section, we present the results on the identification of sounds, categories, and filters. Here, spatialisation just seems to better inform the participants about the identity of the sound source. In the third and fourth parts, we show how ambisonics spatialisation promoted more realistic scenes and musical scenes, as well as contributed to a more immersive musical experience. The combination of everyday sounds and ambisonics seemed to improve the quality of the collaborative musical experience.

6.6.1 Overview: verbal communication themes

We investigated VC from the video data. We examined the following aspects:

1. Awareness of ***sounds, categories, and filters*** (§6.6.2).

2. Awareness of ***realistic scenes and ambisonics*** (§6.6.3).

3. Awareness of ***musical immersion*** (§6.6.4).

Table 6.4 shows an overview of the number of occurrences of speech acts related to WA for each group and condition. Appendix D.8 includes the full transcripts. Figure 6.13 shows box plots for each theme comparing the two conditions.

In Fig. 6.13, we can see some general trends, such as little difference in identifying the sounds and filters between both conditions. The use of familiar everyday sounds, irrespective of the spatial cues, seems to support this. With only a few exceptions, there were fewer errors in identifying sound sources or filters under the spatialisation condition, which suggests that the positioning of the sound on the table can slightly better support the perception of the sound source. However, there was a similar number of errors between

TABLE 6.4: Overview of the number of speech acts related to each group, by condition

Category	G#	NSP	SP	Category	G#	NSP	SP
	G1	0	4		G1	8	9
	G2	1	0		G2	19	18
	G3	0	0		G3	0	1
Identif. of categories	G4	11	0	Impressions of sounds	G4	29	25
	G5	1	0		G5	15	13
	G6	0	0		G6	11	5
	G7	15	2		G7	22	22
	G8	3	1		G8	17	14
	G1	4	18		G1	2	4
	G2	14	36		G2	11	8
	G3	0	0		G3	0	0
Identif. of source	G4	20	5	Non-identif. of person/object	G4	21	8
	G5	3	6		G5	3	2
	G6	1	0		G6	1	1
	G7	15	10		G7	3	5
	G8	18	7		G8	4	2
	G1	2	1		G1	4	3
	G2	1	10		G2	5	2
	G3	0	0		G3	0	0
Non-identif. of source	G4	5	2	Identif. of filters	G4	15	6
	G5	0	0		G5	4	6
	G6	0	0		G6	5	3
	G7	4	3		G7	13	9
	G8	6	1		G8	8	7
	G1	1	5		G1	2	13
	G2	6	1		G2	8	29
	G3	0	0		G3	0	0
Non-identif. of filters	G4	4	2	Stor., realism & ambisonics	G4	14	14
	G5	0	1		G5	8	11
	G6	2	0		G6	2	3
	G7	6	4		G7	11	7
	G8	4	8		G8	6	5
	G1	1	0				
	G2	0	1				
	G3	0	1				
Immers. music. exp.	G4	2	3				
	G5	0	10				
	G6	1	1				
	G7	0	0				
	G8	1	2				

the conditions when identifying the object or the person that was causing a particular sound, which indicates that the positioning of the sound on the table was perceived with similar difficulties in both conditions. Therefore, the sound source seems to have been more clearly identified when using ambisonics, because the sound quality is arguably better; although, the physical

location of the sound is less apparent for users, and may require certain listening training.

The most prominent differences between the conditions were that the use of spatialisation in the SP condition seemed to promote slightly more instances of realistic stories and nuanced associations to ambisonics. In particular, as we can see in the plot of *stories, realism, and ambisonics* in Fig. 6.13, the median value in the SP condition is at the same level as the third quartile in the NSP condition. We also found more instances relating to an immersive musical experience in the SP condition; where, as we can see in the plot of the *immersive musical experience* of Fig. 6.13, the median in SP is also positioned at the same level as the third quartile in the NSP condition. This indicates that the talkative groups had a greater number of realistic and immersive experiences in the SP condition. However, this trend excludes the non-talkative groups.

By contrast, identification of the categories is more general in the no-spatialisation condition. It seems that the lack of localisation information of the sounds promoted a focus on the relationship between sounds and on the more categorical relationships.

These general trends are further developed in the next sections, with representative examples for each theme on WA grouped into *sounds, categories, and filters* (§6.6.2); *realistic scenes* (§6.6.3); and *musical immersion* (§6.6.4).

6.6.2 Sounds, categories, and filters

The use of familiar everyday sounds seems to facilitate the identification of sounds and filters, irrespective of the condition. However, we found fewer errors in the identification of the sound source in the spatialisation condition. The use of localised sounds seems to support a greater precision in identifying the sound source. The identification of the categories took place as a group activity, generally in the no-spatialisation condition, where the use of flat sounds seemed to promote more classificatory tasks, such as identifying the categories of the white cubes.
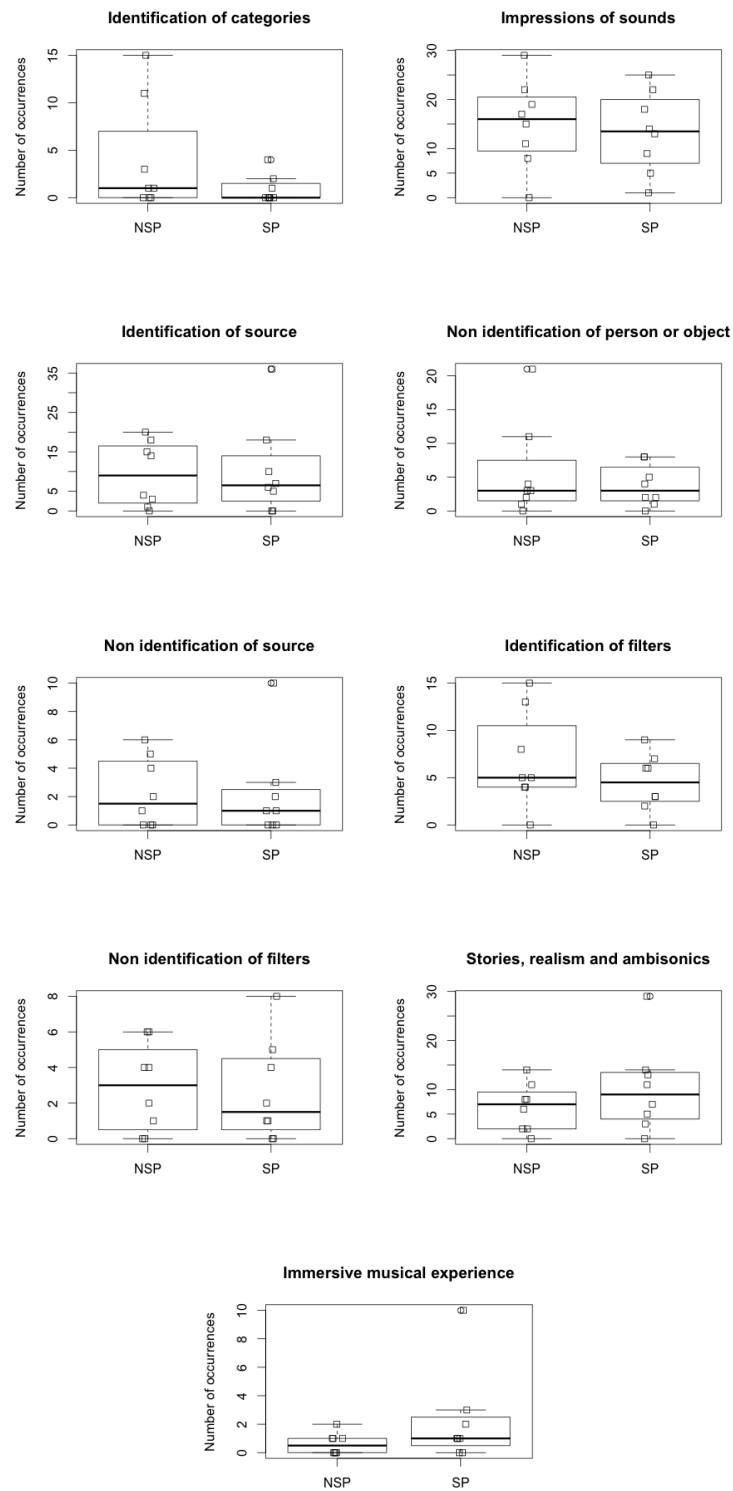
FIGURE 6.13: Box plots of emerging themes by condition

In all the groups, we found early reactions to sounds (e.g. expressions, exclamations, feelings), and also descriptions of the sound source, irrespective of the condition. Yet spatialisation appears to promote more connections to realistic and musical scenes, as we develop further in §6.6.3.

Vignette 1: Identification of the sound of a purring cat in NSP (G4 NSP): [We hear a sound of a purring cat] P3 "A purring cat".

Vignette 2: Impression of the sound of a market in NSP (G1 NSP): [We hear a sound of a group whispering] P4 "The market".

Vignette 3: Identification of the sound of a teapot in SP (G2 SP): [We hear a sound of air escaping hiss] P1 "Sounds like a teapot [laughs]".

Vignette 4: Identification of the sound of a crowd in SP (G8 SP): P2 [We hear a sound of a crowd murmur] "Ok that's the crowd noise".

As shown in 6.13, there was a similar number of comments about the effects of the filters in both conditions, where G4 and G7 commented even more about these effects in the no-spatialisation condition. Four groups (G2, G4, G6, and G7) had fewer awareness issues in identifying filters in the spatialisation condition, whilst three groups (G1, G5, and G8) had more awareness issues in the spatialisation condition. However, in the latter set, G8 had a group discussion in which the group members identified the difference of using or not using a filter.

Vignette 5: Non-identification of the change in filters in NSP (G2 NSP): [As P2 is manipulating a black cube next to a white cube with a sound of a fire crackle], P4 asks "What's the change?". P2 replies "I am not too sure about these enhancing things [the black cubes], what they do …they must do something or just increase the volume but …it's quite subtle if they are".

Vignette 6: Non-identification of the role of filters in SP (G5 SP): P2 "What do these black ones do?" P4 "Filters" P1 "It's supposed to subtly alter sound right?" P2 "They don't seem to add much" P1 "Not very much, does it?".

Vignette 7: Identification of the filters and the sounds in SP (G7 SP): [We hear a sound of a crowd murmur with four black cubes affecting it] P4 "I reckon it's a crowd (...) together".

Vignette 8: Identification of the filters vs the sounds in SP (G8 SP): P3 proposes checking whether changing the side of the "modifier" [filter] matters. P2 replies "No, it is very subtle". P1 asks "Are all the blacks [filters] the same?", and P3 replies "No". P2 adds "They modify the sounds slightly differently" [from previous experience in the NSP condition]. P1 suggests "This black is the same than this black one?", and P3 comments "Shall we try it?" [the group tries a couple of black filters]. P2 affirms "That's different!", and P3 confirms "That's better". P4 seconds "There is difference" while lifting up and down a filter next to a sound player.

Vignette 8 illustrates how a group explicitly commented about the subtlety of the filters, but in the SP condition, where they noticed greater changes in the sounds modified by the different filters. This points to a slightly greater awareness of the filters for some of the groups, at least in the SP condition.

With the exception of G3, all the groups asked questions about what person or object was producing a certain sound. There were also misunderstandings about the sound source in groups G1, G2, G4, G7, and G8, generally with fewer errors in the SP condition, except for G2.

Vignette 9: Non-identification of who controls the sound of a laughing man in NSP (G2 NSP): P3 changes the side of the cube [we hear a man laughing]. P2 asks "Oh, who's got that one?", and P2 replies "It's me".

Vignette 10: Non-identification of who owns the sound of a long snort in SP (G4 SP): [We hear the sound of a long snort] P2 asks "Who's got this noise? I have this noise? That doesn't go with the [inaudible]".

Vignette 11: Non-identification of who owns the sound of a car crash in NSP (G6 NSP): P2 moves a cube he was playing with from the tabletop surface to the rim area, and adds another cube [the tabletop system is momentarily not tracking it]. The other team members are manipulating

cubes as well. P2 changes the side of the cube [we hear a very loud car crash]. P2 lifts the cube up and down again [we again hear a very loud car crash]. P2 asks "That's me?", and P1 replies "Yes".

Vignette 12: Non-identification of the sound of a cat howling in SP (G2 SP): [We hear the sound of a cat howling] P2 asks "Is that a baby?", and P4 replies "Not sure".

Vignette 13: Non-identification of the sound of a cat purring in NSP (G4 NSP): [We hear the sound of a cat purring] P4 exclaims "Oh that's the ...crickets!".

As shown in the above vignettes, in most groups we found explicit questions about awareness issues of individual and of others' actions related to identifying who was manipulating a particular sound, and also related to identifying their own sounds. Identifying their own voice as well as others' voices is relevant here as an aspect of musical tabletop collaboration. Awareness issues around identifying the different voices happened less in the SP condition, as ambisonics seems to support more precision in identifying the sound source. However, here ambisonics does not seem to improve the identification of who or what is the object producing a particular sound. The size of the tabletop surface and the distance of the speakers could be factors in these results.

With some exceptions (i.e. G3 and G6), six of the groups had a passage about classifying the cubes into categories: G1 in the spatialisation condition; G2, G4, G5, G7, and G8 in the no-spatialisation condition.

Vignette 14: Identification of the explosions category in SP (G1 SP): The group is trying to identify the category of a cube [explosions category]. P4 manipulates the cube and says "So red is ..." [we hear a loop of a water splash] and P2 continues "Gat sounds", and P4 seconds "Gat sounds and gunshots and stuff". P3 contributes "That's red, danger ... [we now hear a car crash] ...warning".

Vignette 15: Identification of the voices category in NSP (G8 NSP): The group decides to try cube by cube to get a clearer idea of the categories.

P2 chooses a cube [voices] and starts trying out the different sounds it makes [we hear a cat howling, a Tibetan chant, a man groaning, a child laughing, a cat howling again, a man laughing]. P2 says "Okay, so quite a lot of animal noises on that one", and P1 replies "Animal noises . . . the baby crying? [laughs]"

Vignette 16: Identification of an 'inconsistent' category in NSP (G4 NSP): P2 is operating a cube [screeches category] with no visual feedback [the tabletop system is momentarily not tracking it]. P2 turns the side of the cube and drags it from left to right [we hear a subtle metallic-ceramic friction sound, while the others' cubes are heard louder], and says "Mine is very inconsistent".

Using cubes that represent categories of sounds seems useful for both individual and group awareness under the two conditions. Here, individual awareness is connected to the relationship between the identity of the tangible object and personal ownership; whilst, group awareness is connected to the relationship between the tangible object and the collective experience of discovering its identity, a situation supported by using four speakers where everybody can listen to the sounds, as listeners share the same sound field. The use of a flat sound in the no-spatialisation condition seems to promote more classificatory tasks.

6.6.3   Ambisonics spatialisation: realistic scenes

Table 6.5 shows an overview of the number of occurrences of the *stories, realism, and ambisonics* theme; which has five emerging subtypes: *sound associations*, *storytelling*, *musical composition*, *realism* and *spatialisation*. Appendix D.8 gives the full transcripts. Figure 6.14 shows box plots of each subtype by condition.

As a general trend, as seen in Fig. 6.14, we can see that in the *sound associations* plot, the median and variation are larger in the SP condition. Spatialisation promotes more conversations about sound associations with real or imaginary entities. By contrast, there were more instances of *storytelling*

TABLE 6.5: Overview of the number of speech acts of the *Stories, realism, and ambisonics* theme related to each group, by condition

Category	G#	NSP	SP	Category	G#	NSP	SP
Sound associations	G1	2	11	Storytelling	G1	0	0
	G2	6	12		G2	3	10
	G3	0	0		G3	0	0
	G4	9	7		G4	4	3
	G5	6	8		G5	2	1
	G6	1	1		G6	0	0
	G7	8	9		G7	6	1
	G8	4	1		G8	2	2
Musical composition	G1	0	0	Realism	G1	0	0
	G2	0	3		G2	1	7
	G3	0	0		G3	0	0
	G4	7	6		G4	0	1
	G5	3	7		G5	0	1
	G6	3	3		G6	1	1
	G7	1	1		G7	0	0
	G8	2	3		G8	0	0
Spatialisation	G1	0	2				
	G2	0	2				
	G3	0	0				
	G4	0	1				
	G5	1	0				
	G6	0	0				
	G7	1	0				
	G8	0	0				

in the NSP condition, probably promoted by the type of sounds used, combined with the lack of additional sound localisation information. However, the most salient differences between the two conditions were found in the subtypes: *musical composition*, *realism*, and *spatialisation*. In *musical composition*, we found a greater number of quotes (the median of the SP condition is larger than the third quarter of the NSP condition) referring to the musical aspects. Comments that characterised the sound output as realistic were mostly found in the SP condition. *Spatialisation* was generally unnoticed, although a few groups had conversations about the relationship between the position of the objects and the position of the speakers. These differences are presented next.

Irrespective of the condition, groups tended to associate the music to films, particularly horror films (G2, G4, and G7). In the SP condition, P2 in G4 said "it's like the soundtrack of some movie", and P2 in G7 described it as
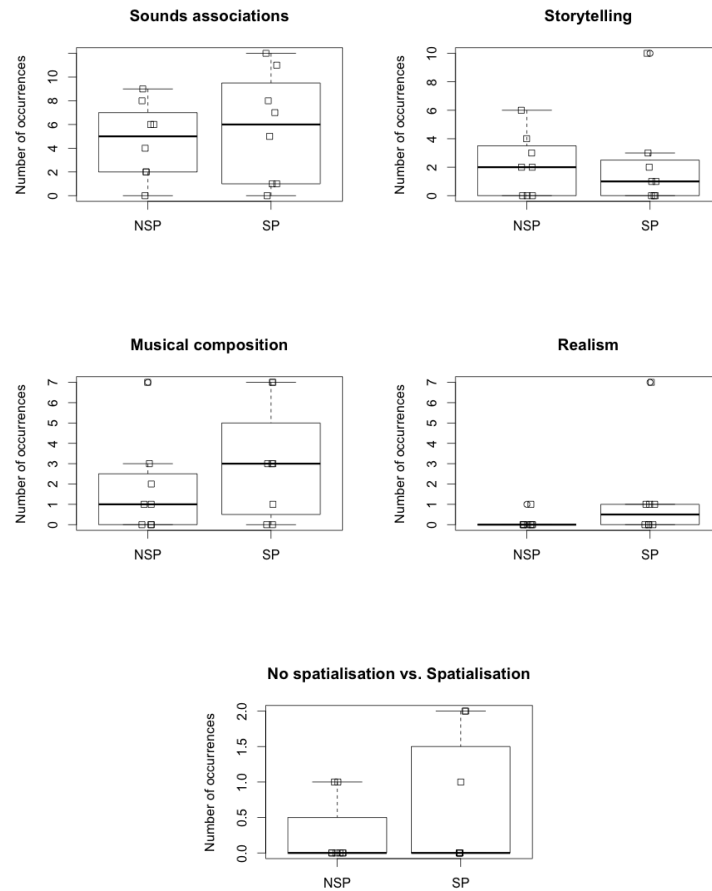
FIGURE 6.14: Box plots of emerging subtypes within the theme *Stories, realism, and ambisonics* by condition

"noises from films"; in NSP, P1 in G2 said "that's a horror film now", while P2 in G7 commented "let's make a horror movie", and P2 in G4 stated "we can make an audio movie". In all the groups and in all the conditions, we found descriptions of or associations to situations, such as "that sounds like a supermarket (...) a cashier of a supermarket" (P2 in G4, SP), or "that's horrible (...), like the computer humming back in the background" (P2 in G2, NSP).

However, we discovered that in the SP condition, the two groups who adopted a more storytelling approach (G2 and G4), attributed the musical output as more realistic, and associated the location of the sound player objects on the tabletop surface to the location of the speakers. In particular, we

heard comments about the realism of the musical events and its connection to scenes, which shows a greater awareness of the group experience as being realistic and immersive, as illustrated in the following example vignettes.

> *Vignette 17: Creating a 'realistic scene' in SP* (G2 SP): [We hear the sound of a thunderstorm, a tom drum, a fire crackle, a child laugh, and a purring cat] P2 says "This sound is awesome". P4 comments "We can make some noises". P2 adds "We created a scene that's all we wanted to do, wasn't it?". P1 replies "That's the purpose, we have to combine them". P2 says "We have our urban noises we've got a purring cat by a fire, and a baby, and some snoring, and a thunder, and a banging door; it seems realistic at least in terms of urban noises".

> *Vignette 18: Creating a 'soundtrack of a movie' in SP* (G4 SP): All the participants are manipulating tangible cubes, when, P2 explains "When we manipulate this, all I can think of is scenes of places, and things happening, to me it's like the soundtrack of some movie or something".

> *Vignette 19: Creating a 'piece for the Olympic games' in SP* (G5 SP): P1 says "It was the piece for a real scene really, wasn't it? It was a mix of the London Olympic game ceremony" [laughs].

> *Vignette 20: Creating a 'song' in SP* (G6 SP): P3 comments "We should record a song like this [with] 2.6 or ...".

In the above examples, we can see an association between the musical activity and the creation of scenes, which only appears in the SP condition. This involves connecting musical sequences to places and situations, which points to a greater connection to real settings for the talkative groups. We also found evidence in the SP condition for groups G2 and G4, of associating the sounds to realistic 'scenes', and of associating the position of the sound sources (cubes) to where the sound was located in the sound field, as illustrated in the following vignettes.

> *Vignette 21: 'Creating actual fire' in SP* (G2 SP): Previously, the group positioned the sound of a fire [fire crackles] to the corner, referring to it

as "a campfire in the corner". A few minutes later [we hear the sound of a thunderstorm], P1 turns his head to the speakers, then he turns back to the tabletop surface, points to his back left speaker, and says "The fire sounds really nice, it sounds as it is actually fire".

Vignette 22: Wondering about the positioning of the objects in SP (G2 SP): The group is listening to a "scene" they have created by looking at the objects [the sound of a fire crackles in one corner of the tabletop surface, the sound of a thunderstorm in another corner]. P2 says "I wonder whether it is interesting where we put these [tangible cubes], because we wanted the fire in the corner, if it had some significance, we created a geographical representation as well as the sound one". P1 adds "It could be a suggestion [to be added to the software]... geographical representations that change something, now it doesn't". P2 replies "I don't think it does".

Vignette 23: Wondering about the positioning of the objects and loudspeakers in SP (G4 SP): P1 says "I am curious if position affects which speaker sound comes out but it doesn't". P3 adds "No, [I don't think] there is no surround sound [looking at the four speakers]". P4 comments "It could be stereo sound; there are four speakers [looking at the four speakers]". P1 continues "It could be four channel, you could actually orient the proximity of the cube to the speakers, which channel to use, but... [looking at his right back and front speakers]".

The above vignettes show how G2 and G4 had conversations that suggest a perception of ambisonics effects, although non-explicit, for example, their mapping the location of the tangible objects on the tabletop surface to the location of speakers. This type of conversation indicates workspace awareness of how the sound was delivered. By contrast, perception of how the sound was delivered in the no-spatialisation condition was scarce. This suggests a less immersive group experience, as shown in the following example:

Vignette 24: Wondering about from which loudspeakers the sound comes in from in NSP (G7 NSP): P2 "So the sound comes mainly from this speaker, yeah? [pointing to his back right speaker]". P1 turns his head to his back

> left speaker, and P4 turns her head to her back right speaker and then to
> her back left speaker. There are no comments.

This example shows how the no-spatialisation condition is perceived as poor and less immersive, which contrasts with the previous examples in the SP condition of more connections, even though this was mostly posed in the form of questions about the position of the tangible objects on the tabletop surface and the speakers.

6.6.4   Musical immersion

For the non-talkative groups such as G3 and G6, it was difficult to tell from the video data whether their group experience was an immersive experience. In contrast, we found evidence of a greater engagement with the musical activity as an immersive group experience for the most talkative groups when using ambisonics spatialisation. Furthermore, we observed that groups tended to talk slightly less in the SP condition (see §6.5.1). As the above plot of the *immersive musical experience* in Fig. 6.13 shows (see Section 6.6.1), we found a greater number of occurrences of speech related to musical immersion in the SP condition in terms of comments about enjoying the activity: the median in the SP condition is greater than the third quarter in the NSP condition, and the range is larger, with a few groups reaching a considerable amount of occurrences on musical engagement. Except for G7, all the groups made explicit comments.

We found differences in describing the ongoing musical outcome, pointing to the greater association of producing enjoyable and immersive music in the SP condition. In the NSP condition, there were a few comments (G1, G4, G6, and G8) about the musical outcome, such as "this is nice" (G1 P1), "for a moment we had something nice" (G4 P2), "we are very close to a preview performance, aren't we?" (G4 P3), or "we started making some music" (G8 P2). In contrast, in the SP condition, we identified more comments in more groups (all groups except G1 and G7) about engaging with the group musical output, such as "this sound is awesome" (G2 P2), "I think we have reached

some musical plateau" (G4 P2), "we should record a song like this" (G6 P3), "it's music now" (G8 P3), or "we've arrived, we are doing Stockhausen now, we've arrived, that's definitely Stockhausen, take it back down, stepping in, boom" (G5 P1). Using localised sounds with a four-speakers surround sound system seems to support a more musical immersive experience.

6.7   Results: space use and territoriality

We also studied non-verbal communication, partly because there were non-talkative groups involved, but also because we were interested in understanding whether there were differences in the group tabletop interactions depending on the condition. Although the results of this second part were less substantial, there are some relevant aspects to our research questions worth highlighting here. In the first part of this section, we give an overview of the non-verbal communication themes. In the second part of this section, we present activity plots of the space explored. We observed different patterns, ranging from no differences between the conditions, to a greater use of the central area in the no-spatialisation condition. In the third part of this section, we present the transitions between the individual and central spaces. We also describe how we observed different patterns, with, on average, slightly more fluid transitions in the SP condition between individual and central work. In the fourth part of this section, we present the amount of activity, with no apparent differences. These results point to a similar amount of activity, irrespective of the condition, but also a trend towards using more space territory in the SP condition.

6.7.1   Overview: non-verbal communication themes

Our analysis of NVC complemented our above investigation on verbal communication (§6.6). In particular, we examined three aspects:

- The *space explored* (§6.7.2).

- The *transitions* between individual space vs shared spaces (§6.7.3).

- The *amount of activity* (§6.7.4).

In summary, we found differences in half of the groups in terms of *space use*, where participants used more the space next to themselves and the in-between spaces in the SP condition. In other words, half of the groups consistently used more the central area in the NSP condition. This points to a greater exploration of spaces in the SP condition for half of the groups, and a preference for individual voices or for manipulation of the sounds that could be heard from the nearest loudspeakers to the individual. We also discovered a greater number of transitions between working in the spaces next to individuals and working in the central area. This indicates that there were more *fluid transitions* in the SP condition. However, we found a similar amount of activity across both conditions. Therefore, spatialisation seems to equally promote activity as in the no-spatialisation condition. These general trends are further developed in this section with illustrative examples for each aspect.

6.7.2   Space explored

Figure 6.15 shows the activity maps of the sessions. When comparing the two conditions, the activity maps demonstrate three main patterns depending on the group: 1) no difference (G1, G5, G6, and G8); 2) more individual work in SP and central work in NSP (G3 and G4); and 3) more work in the spaces between participants, including corners, in SP and central work in NSP (G2 and G7). Here we give an example vignette that illustrates the third pattern.

Vignette 25: Corners used for building a story in SP (G2 SP): [24:11-39:11]
The four members constructed a story together using discrete gestures positioning on-off sound player objects on different parts of the tabletop surface, including in the individual spaces, the centre, and mostly the corners, while explaining a collective story using turn-taking. In contrast, in NSP, the objects are positioned mostly in the centre.
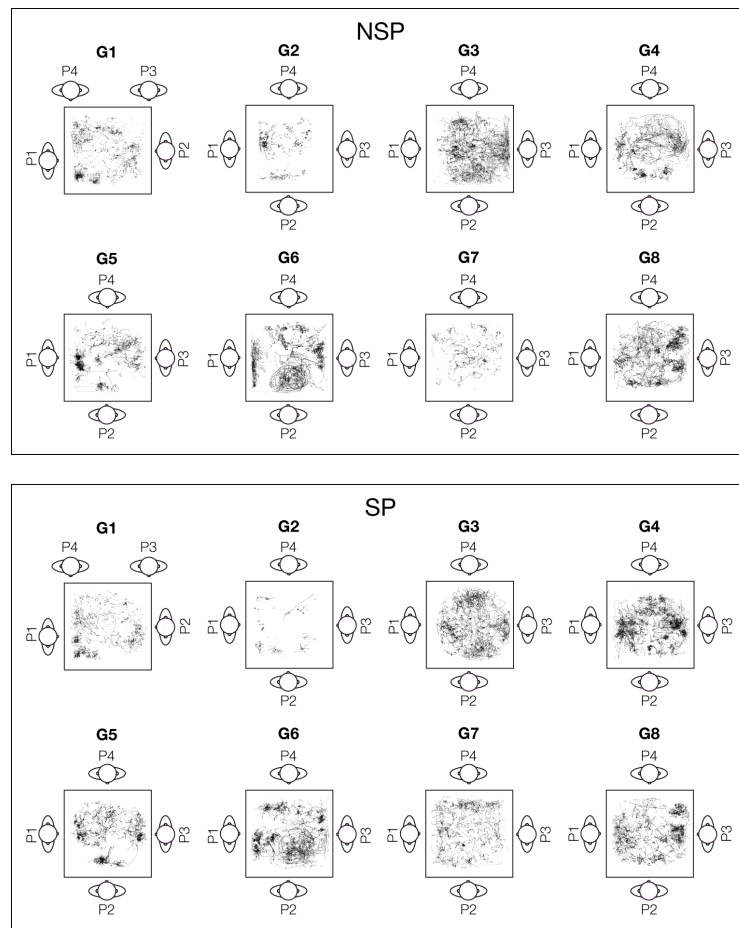
FIGURE 6.15: Activity maps for each group in the NSP and SP condition

The above vignette reveals that, for G2, ambisonics spatialisation helped the participants discover certain areas of the tabletop surface. In particular, the SP condition encouraged participants to use not only their individual spaces and the central shared space, but also the in-between spaces, which indicates that the approach supports fluid transitions between the personal voices and the group voices in this group.

We found no significant differences between the conditions in the area of the space used: i.e. in the mean of coordinates x and y; the standard deviation of coordinates x and y; the range of coordinates x and y; and more broadly, in the overall area of the distance travelled by the tangible objects. This points to a similar space being used in both conditions; although, as we discussed above, the areas explored vary depending on the group and the condition. For some

TABLE 6.6: Occurrences of transitions across the groups and conditions

Group	Overall		Individual		50% Individual 50% Central		Central	
	NSP	SP	NSP	SP	NSP	SP	NSP	SP
G1	39	28	16	13	18	11	5	4
G2	4	4	1	1	2	1	1	2
G3	65	67	20	25	30	33	15	9
G4	31	41	9	13	16	19	6	9
G5	42	43	15	19	20	20	7	4
G6	37	38	13	15	18	17	6	6
G7	12	50	4	16	5	23	3	11
G8	48	43	18	18	21	19	9	6

groups, spatialisation promoted more exploration of the individual's space next to themselves, and the spaces in-between participants. For other groups, there was a similar exploration of the individual space and central spaces, irrespective of the condition.

6.7.3 Transitions

We analysed the video data according to the categories on transitions (Table 6.6). As shown in Fig. 6.16, the mean number of transitions was smaller in the NSP condition ($M = 34.75$ transitions, $SD = 19.43$) compared to the SP condition ($M = 39.25$ transitions, $SD = 18.08$), which indicates that spatialisation promoted a greater number of transitions. We observed three main patterns: 1) more transitions in NSP (G1 and G8); 2) an equal number of transitions (G2) or slightly more transitions in SP (G3, G5, and G6); and 3) more transitions in SP (G4 and G7).

Figure 6.17 shows these results at a finer level, and where we can see the number of transitions by category and condition. Whilst the *individual* and *50% individual - 50% central* values were greater in the SP condition, the amount of work occurring in *central* areas was slightly greater in the NSP condition. These results suggest that the SP condition, apart from promoting more transitions between individual and central work areas, also promoted more work in the space next to the individuals, as well as hybrid configurations of some
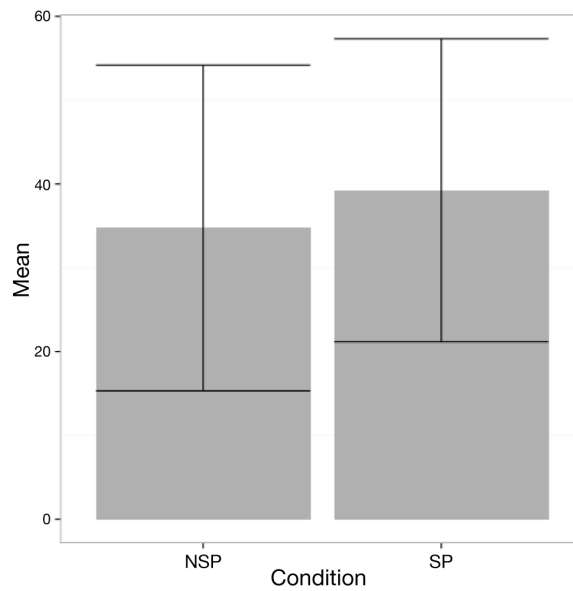
FIGURE 6.16: Bar plots with the standard deviations of the number of transitions across the two conditions

participants working together and others working individually. Spatialisation seems to promote a more flexible approach to tabletop collaboration.

Next we provide an example that illustrates this workflow between individual and group work.

> *Vignette 26: Parallel individual and group work in SP* (G4 SP): The group starts with all objects in the centre of the tabletop surface, then they choose cubes for themselves to play with. P1 and P2 keep manipulating cubes next to themselves, whilst P3 and P4 build in collaboration by positioning the cubes in the centre.

This vignette illustrates how ambisonics spatialisation supported more versatility in the use of territories. For G4, the SP condition supported versatile collaboration between two musicians in the centre, while the other two musicians were manipulating tangible objects in their individual spaces.
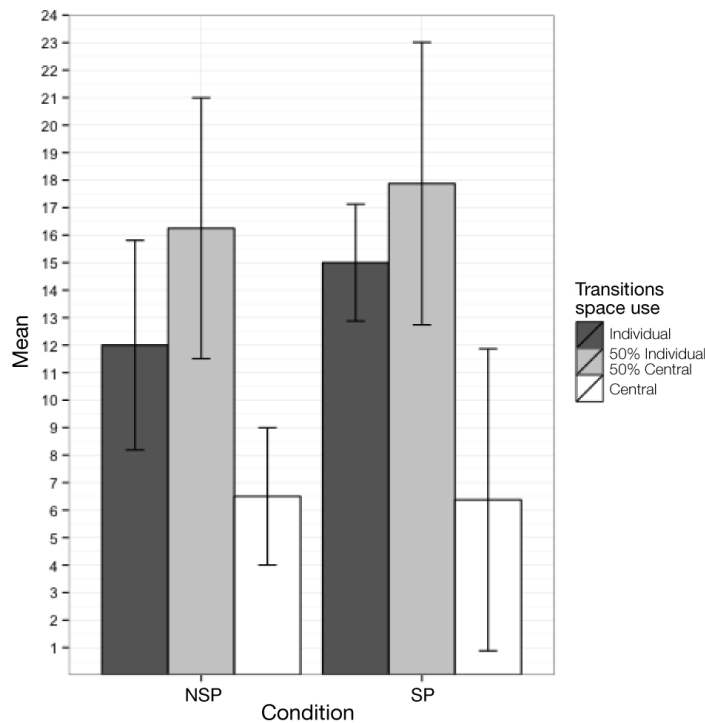
FIGURE 6.17: Bar plots with the standard deviations of the number of transitions by the type of space use across the two conditions

6.7.4 Amount of activity

We found no significant differences between the two conditions in the amount of activity carried out, in terms of number of object *sets*, *updates*, and *deletes* of tangible objects on the tabletop surface. We found no significant difference in the average of objects utilised per second either. On average, groups contributed similarly in terms of the amount of activity carried out, irrespective of the condition.

We plotted the tabletop surface as an imaginary matrix of 4×4 and counted the number of occurrences of an action taking place in each of the 16 cells of the matrix for each group and condition. The aim was to reveal the patterns of use related to an individual's area (from top to bottom and left to right cells 2, 3, 5, 8, 9, 12, 14, and 15), the central area (cells 6, 7, 10, and 11), and the in-between spaces (cells 1, 4, 13, and 16). We then averaged the data across all groups and used the matrix to visually compare the number of occurrences of

NSP

0.03 (0.04)	0.06 (0.04)	0.06 (0.03)	0.02 (0.02)
0.08 (0.05)	0.1 (0.03)	0.12 (0.04)	0.06 (0.03)
0.04 (0.04)	0.08 (0.06)	0.09 (0.06)	0.06 (0.04)
0.03 (0.04)	0.09 (0.04)	0.06 (0.04)	0.01 (0.01)

SP

0.04 (0.05)	0.06 (0.04)	0.07 (0.04)	0.03 (0.04)
0.08 (0.06)	0.07 (0.04)	0.08 (0.05)	0.06 (0.04)
0.07 (0.03)	0.07 (0.04)	0.08 (0.03)	0.07 (0.02)
0.04 (0.06)	0.07 (0.03)	0.08 (0.03)	0.02 (0.02)
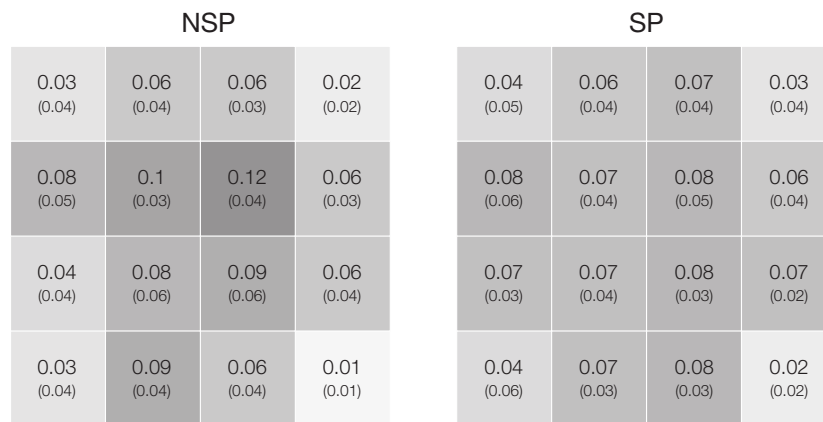
FIGURE 6.18: 4 x 4 matrix visualising the proportions of occurrences of objects across the groups (mean and standard deviation)

objects between conditions. Figure 6.18 shows the results. Here, NSP shows a slightly more patchy pattern with a major concentration in the centre of the tabletop surface; while SP shows a slightly more consistent pattern, where the individual's own area and the centre of the tabletop surface are both visited more uniformly. Even though the differences between the conditions, in terms of visited matrix cells, are subtle, these averaged results agree with the results on the transitions of a greater distributed use of space.

Overall, we found that ambisonics spatialisation affects the space use and leads to more fluid transitions occurring between the individual and central spaces. However, the amount of activity measured as sets, updates, and deletes of tangible objects seems unaffected by the condition. Next, we discuss the presented results.

6.8 Discussion

In this section, we first discuss the general behaviours and group dynamics, and then we present the effects of spatialisation on workspace awareness and space use. We found that the spatialisation condition tended to promote a more immersive and realistic group experience, with higher workspace awareness in the collaborative activity and a more fluid use of the individual

spaces and central spaces, but with differences depending on the group dynamics. We also found general tabletop behaviour with differences in group dynamics that resonate with our previous findings.

6.8.1   Tabletop general behaviour

In this study, we found general behaviours similar to our previous studies, irrespective of the condition, which suggests a set of common characteristics and issues:

- **Non-participation** seems to be promoted by interactive tabletops for music performance. This is different to inactivity, as sounds in these TTIs are looped and hence there is no need to continuously manipulate the objects to get them to play. We also found instances of non-participation in our study with expert musicians over time (Study 1) and in our study with visitors in a science centre (Study 2). In particular, in Study 1, we argued that non-participation can be seen as a positive action.

- **Physical explorations** suggest that the physical domain is explored independently from the digital domain, and in this case, independently of whether it has a musical effect. We found similar instances in our previous work in Studies 1 and 2. These exploratory interactions can shed light on the future mapping between the physical domain and the digital domain. For example, stacking objects could have a 3D ambisonic spatialisation effect.

- **The role of icons** in an interface facilitates identification of the elements using graphic representations. In this study, the lack of representative images, and the use instead of markers designed to be understood by computer vision software, led to difficulty in associating the images to sounds for users. The association of sounds to their respective fiducial markers seemed to be difficult, due to the graphical similarity between the markers, thus making the visual-to-sound association problematic.

In this study, an approach was designed to promote a greater use of the auditory channel, and the real-time visual feedback was only used to indicate whether the tangible object was active on the tabletop surface. The results suggest that including human-readable icons or visual references that refer to the content of the tangible objects is important for a better user experience, at least in terms of a lower cognitive load for associating the icons to sounds.

- **The role of auditory feedback** and its relevance in the TTIs was evidenced by the incident in which real-time visual feedback was suspended for a few minutes in three groups, yet the groups affected maintained the musical activity. This suggests that, even though visual feedback is important, it can be seen as secondary, compared to auditory feedback in the TTIs, which includes not only the computer-generated audio, but also the physical sound of the objects. For example, our Study 3 on providing stereo spatialisation for two users had no real-time visual feedback, and yet the groups could still improvise the music. Although, having said that, real-time visual feedback seems to improve the overall interactive experience, as shown in our Studies 1 and 2.

6.8.2 Group dynamics

Our results show how the use of ambisonic spatialisation in tabletop collaboration is far from a 'one-size-fits-all' technique — group dynamics are different, and so it is explored differently in each group. In the present study, we have shown how mixed groups engaged with the music improvisation process using different working styles, from storytelling to music performance. Rick et al. [2011] recognised these differences in group dynamics and highlighted the importance of supporting them on interactive tabletops for promoting collaborative learning, for example. Here we found instances of groups creating soundscapes together with everyday sounds, which seems to support this notion of learning in a group.

Our approach seems promising for investigating further how best to deliver an immersive musical experience by using auditory feedback with tabletop systems that can inform users at both the individual and group levels in a democratic setting beyond just using everyday sounds. We observed that it is possible to adopt a range of musical improvisation approaches, from storytelling to music performance. For those groups where ambisonics spatialisation had a greater impact on space use, it allowed them to discover and better use the individual spaces next to themselves vs the other spaces distributed in the centre and corners of the tabletop surface. This points to the role of space being relevant when using spatialisation, and the use of the central space as not necessarily being the shared space that we generally find in tabletop literature. However, we still need to investigate further how to best support a wide range of groups. Next, we discuss in more detail our results in terms of workspace awareness and tabletop space use.

6.8.3 Workspace awareness

In response to our research question on workspace awareness, we found that ambisonics can improve WA. This improvement can be described as:

- A better perception of the sound source.

- More associations with realistic scenes and musical scenes.

- Deeper musical immersion.

A better perception of the sound source: why? In our study, sound sources were better perceived using ambisonics, which suggests that the quality of the perception of everyday sounds could be improved using ambisonics: the implementation of the spatialisation technique using four speakers seems to support this. However, there was not much difference between the two studied conditions in the identification of who or what object was producing a particular sound. We think that the experimental setting could have affected the listeners' lack of perception of the location of the sound, information only

present in the spatialisation condition, where the size of the tabletop surface, and the distance of the speakers from the table might be factors accounting for this lack of perception.

More realistic and musical scenes: why? Ambisonics impacted on two of the groups, who took a storytelling approach to musical improvisation. The position of the objects on the tabletop surface and their relation to the position of the speakers influenced the narration of realistic scenes. Here, it seems that the familiarity of the everyday sounds helped the participants to create these scenes, and the position of the sounds in the sound field also seemed to influence the story. The less talkative groups, although still with a certain amount of conversation, linked the sound output to more realistic musical scenes. Spatialisation replicates our acoustic perception in real environments, and hence our awareness with spatialisation is expected to be similar to our awareness in everyday listening [Gaver, 1993], particularly as we are using everyday sounds that listeners can connect to familiar situations. Furthermore, the invention of ambisonics aimed to create a more 'realistic' surround sound system, compared to flat quadraphonics, or in the words of Gerzon "'realistic' ambient concert hall recordings to be made with improved realism" [Gerzon, 1974b, p. 56] when compared to quadraphonics. Ambisonics combined with everyday sounds seems to facilitate workspace awareness of both individual and group activity, irrespective of the musical background.

More musical immersive experience: why? For the majority of groups, ambisonics spatialisation seemed to help them enjoy a deeper group musical immersion experience, irrespective of the musical background of the participants. In contrast with Blaine & Perkis [2000] and their approach of combining individual speakers in a surround sound system, here, we only used a surround sound system to support the overall group musical experience with no detriment to the individual musical experience.

6.8.4 Space use and territoriality

In response to our research question on group tabletop interaction, we found that ambisonics can influence GTI. This influence can be described as:

- More use of individual spaces and the in-between spaces.

- Greater number of transitions.

However, we did not find any influence in the amount of activity carried out, which suggests:

- A similar amount of activity is carried out irrespective of the condition.

More use of individual spaces and the in-between spaces: why? As shown with the activity maps of the two conditions (Fig. 6.15), for one half of the groups, the exploration of space in the SP condition can be related to the positioning of the objects in the sound field, in which there are differences between positioning the sound in the individual space, in the in-between spaces, or in the centre of the tabletop. In the no-spatialisation condition, instead, there is no change when positioning the sounds in the different available spaces, which explains why there is a less clear pattern and a greater focus on the central areas as a preferred area of group work [Scott & Carpendale, 2010; Scott et al., 2004]. For the other half of the groups there was no difference as a group, irrespective of the condition, indicating that group dynamics predominated over the condition. This observation indicates that spatialisation does not necessarily affect all the groups in terms of space use using this protocol and setting. Other studies in literature that focused on individual actions with colours for each participant traces found differences between the conditions [Tuddenham & Robinson, 2009], or between interaction techniques [Nacenta et al., 2010]. Thus it is an open question whether tracing individual actions with a colour for each participant would tell us more about these groups.

Greater number of transitions: why? We found a slightly greater average
number of transitions from individual work to central work in the SP con-
dition. As highlighted in the literature, transitioning between individual
work and group work shows workspace awareness [Gutwin & Greenberg,
2002; Tuddenham & Robinson, 2009]. A greater number of transitions in SP
points to a more fluid collaboration in this condition for some groups, as it
requires a greater negotiation of the available space; more space exploration;
and a greater use (and distinction) between individual and group spaces. Our
study shows, in contrast to Hancock et al. [2005]'s study, that a simultaneous
support of awareness using auditory feedback for both individual and central
actions, which potentially could be group actions, is also possible. Generally
there were more transitions in SP, but for one group there was no difference,
and two other groups did more transitions in NSP. Thus, spatialisation does
not necessarily affect all the groups, at least as measured by the number of
transitions using this protocol and setting. In particular, the space use of two
mid-talkative groups, who did more transitions in NSP, was not affected by
spatialisation either. This result leaves open questions about the conditions
affecting transitions and space use, which should be investigated further.

A similar amount of activity is carried out irrespective of the condition:
why? Contrary to our expectations, the range and area of the space used were
similar in both conditions. Ambisonics does not seem to support a greater
exploration of the tabletop space. Similarly, the amount of activity carried
out was similar in both conditions, only varying in the distribution of the
activity on the tabletop surface, i.e. both individual and central areas were
similarly utilised when using ambisonics. These no-difference results indi-
cate that measuring the engagement only using a quantitative approach of
counting the number of sets and removing the objects may be too reductive,
especially in an open activity such as musical improvisation.

6.8.5 Implications for design

Our results show why audio delivery matters for workspace awareness and tabletop interaction. A four-speaker ambisonics system may therefore have the following implications for the design of tabletop systems:

1. ***Central and individual work***. Studies of co-located tabletop systems provide evidence on how such systems support the coupling between individual and group work, compared with remote tabletop systems [Tuddenham & Robinson, 2009]. However, when using auditory feedback, and in contrast with Hancock et al. [2005]'s study of the context of auditory icons, our study shows that simultaneous support for the awareness of both individual and central actions, in which some of them are group actions, is possible in co-located settings. This approach could support collaborative work on tabletop systems using sound or music.

2. ***Ecological and democratic approach***. This study also showed how mixed groups engaged with the music improvisational process using different styles from storytelling to music performance. Blaine & Perkis [2000] experimented with audio delivery to mixed groups with different musical backgrounds, and found that the best choice of sound delivery for individual and group awareness was to combine individual speakers with a surround sound system. Here, we highlight a more ecological and democratic approach, using only a surround sound system to support the overall group musical experience.

3. ***Interactive, realistic, and immersive installations***. Using ambisonics spatialisation, we found evidence of a greater engagement with the musical activity as an immersive experience, including the associations to realistic situations. This approach thus seems promising for further investigating how to best deliver a realistic and immersive musical experience using auditory feedback with tabletop systems. This approach contributes towards using surround sound systems to improve group

awareness and collaboration in museums and public spaces that work with this type of exhibit.

4. ***Beginners and experts***. Using ambisonics spatialisation with everyday sounds on tabletops seems to be a novel approach towards democratic music making for mixed groups. The previous work on collaborative music with mixed groups tended to promote experts as the leading voices, and novices as the accompanying voices, as exemplified by Blaine & Perkis [2000]. A salient challenge of democratic music making in heterogeneous groups is that everyone participates in a similar way, irrespective of their musical background. The use of everyday sounds using ambisonics seems to offer a familiar environment to everyone, thus supporting democracy in mixed groups.

5. ***Sonic design***. Ambisonics spatialisation on a tabletop system could provide a suitable tool for sonic design that could be furthermore explored in collaborative sound design contexts, such as in participatory film *soundtrack design*. Another area of exploration is participatory *soundscape composition*, in contrast to Westerkamp [2002] and the traditional notion of soundscape composition by a single composer. Ambisonics and interactive tabletops used in a wide range of contexts from composition to performance, to artistic installations for collaborative sonic design are promising.

6. ***Musical education***. Combining ambisonics with interactive tabletops is an interesting approach for musical education. It could be used, for example, as a platform for improving environmental listening awareness, for improving the understanding and appreciation of our soundscape environment (see Westerkamp [2002]).

7. ***Controlled experiments***. We still know little about how our everyday listening and ecological acoustics work [Gaver, 1993]. Our approach of using ambisonics with everyday sounds could support controlled experiments in the psychology of everyday listening studies. Furthermore, there are cognitive psychological studies, such as navigational

tests, that need to create immersive realistic experiences, which could benefit from using ambisonics auditory feedback with a tabletop interface.

6.8.6   Study limitations and future work

Below we discuss study limitations and future work from this research:

Tabletop size, sample size, and sounds used. Our study included a set of aspects that need to be considered as they frame our results, including the size of the tabletop system, the sample size, and the type of sounds used. First, the size of the tabletop system that we used was a medium size suitable for the four users in each group. For the purpose of our study, the tabletop size was sufficient to distinguish differences between individual and group work, and indeed, participants did not comment about a lack of space. Second, the number of groups was small, but the number of group members was consistent with, and even large for, other tabletop studies. Third, Russolo's taxonomy and musical improvisation is not to everybody's taste, and so the risks of this approach are that some groups might not enjoy the experience so much. However, for this study, the musicians improvised music by creating soundscapes with everyday sounds, which were suitable for teamworking with heterogeneous groups.

Ecological validity. The use of a particular set of everyday sounds based on Russolo's taxonomy is an aesthetic choice, and one which we might expect could be enjoyed differently depending on the individuals and the group dynamics. Even though almost all the groups seemed to engage with the collection of sounds, there were two groups that expressed less pleasure. We think that this aspect strengthens the ecology of the study, as this situation can happen in real life where people like different musical genres, or where collaborative music making tends to be a negotiation between personal musical preferences and a group music style.

Future work. In the future, we hope to study the use of other collections of everyday sounds, and to develop alternative metrics to assess tabletop space

use in music performances. Additional open questions include the affects on group collaboration and awareness of using ambisonics implemented with more than four loudspeakers in: 1) the horizontal surround set-up; and 2) the 3D surround set-up.

6.9   Summary

In this chapter, we have seen how ambisonics spatialisation can support tabletop collaboration and awareness with mixed groups using a collection of everyday sounds inspired by Russolo's 1913 taxonomy of urban noise sounds. The results suggest that this ecological approach can support both individual and central work, as well as immersive musical experiences that, with the use of collections of organised everyday sounds, support collaborative and democratic work within mixed groups. We finally summarised the implications of these results for HCI, education, and music technology. This study has helped us to outline a novel approach to tabletop research using auditory feedback, incorporating ecological aspects, for small groups.

In the next and concluding chapter, we discuss the results of our three studies in the context of a broader picture of this dissertation's research questions.

Chapter 7

Conclusions

In this chapter, we discuss the challenges and opportunities raised by the empirical work of this thesis to tabletop TUIs for music performance. We summarise the theoretical, methodological, and practical implications of this work for the areas of NIME, tabletop research, and HCI research, and close with proposals for future work and some final remarks on tangible tabletop music research.

7.1   Introduction

A new generation of tangible tabletop interfaces (TTIs) has emerged in recent years, providing shared tabletop interfaces for real-time collaborative music making and thus new opportunities for music performance.

The range of research into musical collaboration in the NIME community (i.e. CSCM systems), and tabletop collaboration in the HCI community, is broad. However, there is a paucity of research on how best to design and evaluate TTIs for supporting collaboration in music performance. Moreover, little is known about how both beginners' and experts' use of TTIs can be equally supported, because most systems and studies focus on understanding either one or the other, or both together, by designing interfaces that are configured to treat beginners and experts as completely distinct groups, interacting with

the system in quite different ways. Our study contributes to work on group interactions on musical tabletops, by unveiling:

1. The potential for an egalitarian approach to music performance irrespective of the musical background of the group members.

2. The nature of complex and highly-demanding group interactions and practices found in short- and long-term uses of a musical tabletop.

These two areas are informative for the NIME community and the HCI community respectively.

7.2 Research questions revisited

The thesis sought to answer the overarching research question:

> *What are the challenges and opportunities provided by TTIs for music performance among beginners and experts?*

In particular, the thesis sought to answer the research questions:

1. *What is the nature of group tabletop interaction (GTI) on TTIs during music performance among beginners and experts?*

2. *In TTIs for music performance, how is workspace awareness (WA) manifested among beginners and experts?*

We have studied both subsidiary research questions in two distinct sets of circumstances:

1. *Long-term use* (e.g. music rehearsals).

2. *Short-term use* (e.g. museums, casual set-ups).

The key findings were summarised in the three empirical studies:

1. ***Expert peer learning with the Reactable*** (Chapter 4).

2. ***Naïve social interaction with the Reactable*** (Chapter 5).

3. ***Supporting tabletop collaboration and awareness with ambisonics spatialisation for heterogeneous groups*** (Chapter 6).

We now summarise the answers to these detailed research questions by collecting together the relevant empirical findings from Chapters 4, 5, and 6.

1a. What is the nature of group tabletop interaction on TTIs during music performance among experts in the long-term use?

We provided evidence of:

- ***Peer learning***: mimicking (§4.9.1), and explicit peer learning (§4.9.2).

- ***Egalitarian/flexible roles***: fast-paced exchange of roles with no need of verbal communication (§4.8.2), and positive non-participation (§4.8.1).

- ***Situated long-term collaborations***: development of group dynamics over time (§4.5), in particular from 'not knowing each other and little about the technology' to 'knowing each other and mastering the technology' (§4.6, §4.7, §4.8, and §4.9).

1b. What is the nature of group tabletop interaction on TTIs during music performance among beginners and experts in the short-term use?

- ***Social interaction***: multi-directional group interactions, including within-group and between-group (§5.6), along with synchronous and asynchronous collaborations in public settings (§5.7).

- ***Egalitarian/flexible roles***: fast-paced exchange of roles, sometimes supported by verbal communication (§6.6.3 and §6.6.4).

- *Situated short-term collaborations*: fast-paced exchange of groups (§5.5), wide range of group dynamics (§6.5), and suitability for collaboration with strangers (§5.6 and §5.7).

2a. In TTIs for music performance, how is workspace awareness (WA) manifested among experts in long-term use?

We provided evidence of:

- *Little WA control in early sessions*: expert musicians, who were also knowledgeable of the Reactable, initially had little control over WA (§4.6.1), but that this was overcome over time by practice (§4.6, §4.7, and §4.8), and by sharing the problem with others (§4.9.2).

- *Greater control of WA in late sessions*: expert musicians mastered WA over time by developing workarounds and practical knowledge (§4.7.1).

2b. In TTIs for music performance, how is workspace awareness (WA) manifested among beginners and experts in short-term use?

We provided evidence of:

- *Little WA control in early sessions*: with a few exceptions (§5.8.1 and §5.8.3), visitors focusing on group experiences, sometimes with little control over the interface (§5.8.2).

- *Auditory feedback and immersive experiences*: the importance of immersive experiences both in professional music and museum contexts (e.g. auditory feedback using a surround sound system in professional music (§6.6.4), headphones in public settings (§5.6.1 and §5.8.2)).

- *Engagement with physical objects irrespective of their digital responsiveness with the system*: visitors engaging with the tangible objects irrespective of whether the objects were tracked by the system or not (§5.8.2), an unclear feature of the tabletop system that can be mitigated by using tangible cubes that trigger sounds from all their sides (§6.6.2).

7.3 Theoretical, methodological, and practical implications

We now present the theoretical, methodological, and practical implications of our research.

7.3.1 Theoretical implications

Theoretical implications for NIME research include the following:

- Blaine & Fels [2003] suggest that highly constrained musical instruments should be designed for beginners, because these would prioritise group experience. Our results show that a complex interface that allows for both summative and multiplicative contributions, can offer an engaging experience for beginners (§5.8.2), experts (§4.6), as well as enabling beginners and experts to work together productively (§5.8.1), and can attract both beginners and experts together (§6.8.5).

- As we argued in the literature review (§2.5), there has been considerable attention paid to visual feedback in TTIs, but much less attention to the design of auditory feedback, a shortcoming that should be rectified (§5.9.3 and §6.8.5).

- Barbosa [2003] suggests that no examples of co-located asynchronous interaction in the music domain have appeared in the literature. We have shown that such a collaborative approach exists with musical tabletops in public settings, and our work provides the first investigation of its nature (§5.9.3).

Theoretical implications for tabletop research include the following:

- Scott et al. [2004]; Tang et al. [2006] suggest tabletop designs that focus on within-group interaction irrespective of location (e.g. workplace,

museums, etc), for example, collaborative coupling styles between individual and group work [Tang et al., 2006], or territoriality [Scott et al., 2004]. We have argued that the set of design considerations needs to be expanded to encompass between-group interactions, in particular to enable fluid interactions and smooth overlaps between groups in public settings (§5.9.3).

- Morris et al. [2004] suggest adding constraints and roles to tabletop group collaboration. However, we have argued that such interface constraints are not needed, as they tend to prevent self-regulation of spaces within collaborations (§4.10.5).

- When working with musical tabletops, workspace awareness issues are manifested in early interactions irrespective of the expertise of the group, as shown in §4.10.3, §5.9.2, and §6.8.3. Moreover, WA can be mastered over time, as shown in §4.10.3 and §5.8.3; and auditory feedback can facilitate the identification of the individual voice as well as the other participants' voices (§6.8.3). Our results extend the CSCW's workspace awareness definitions presented in §2.4.1: workspace awareness can be learned over time, it is thus a situated aspect of collaboration within TTIs; and the audible can also emphasise the intentionality of individuals, as the visible does.

- Scott & Carpendale [2010]; Scott et al. [2004] suggest clear divisions between tabletop surface areas in TTIs: individual, shared and storage. These categories need to be reconsidered for systems that provide auditory feedback using spatialisation, whereby the boundaries between individual and shared spaces are blurred (§6.8.5).

- Gaver [1991] highlight the use of audio in collaborative systems, but there has been little follow-up on this question in recent years. We have shown that providing auditory feedback with everyday sounds, can reinforce workspace awareness of individual and group work in practice-based activities (§6.8.5).

Theoretical implications for tangible interaction and HCI research include the following:

- Klemmer et al. [2006] suggest that practical knowledge and physical affordances from the real world can inform TUI design. As a follow-up, our research indicates that practical knowledge of musical tabletops can inform TUI design as well, because TUIs need to be considered as part of the real world. For example, we observed peer learning by doing and mimicking (§4.10.5), fluid transitions and overlaps between groups in a public setting (§5.9.3), and everyday listening driven by physical action (§6.8.5).

- Seamless coupling is important in TUIs for controlling and understanding a TUI system, as indicated in §2.3. We provided evidence that physical explorations are needed for discovering the mappings between the physical and the digital, which are not always straightforward (§4.10.2, §5.9.2). We have shown that the material of the tangible objects is relevant to these physical explorations with the TUI, in particular the sound produced from their interactions, and the level of human-readability of the tangible objects (§5.9.3).

- Zuckerman et al. [2005] show how modular TUIs facilitate hands-on learning experiences. We provided evidence that a TUI based on constructive building blocks proves to be suitable for both beginners and experts (§4.10.5, §5.9.3), supports varied group dynamics (§4.10.4, §6.8.2), and promotes both short-term and longer-term interactions (§5.9.3, §4.10.4).

7.3.2 Methodological implications

Methodological implications for NIME research include:

- Booth & Gurevich [2012]; Pugliese et al. [2012]; Swift [2012] studied group collaboration using ethnography, interviews, questionnaires, and

video as an elicitation tool. Our approach is novel in bringing from the social sciences the use of video for carefully analysing interaction phenomena, both verbal and non-verbal, such as peer learning or between-group interaction (Studies 1, 2, and 3).

Methodological implications for tabletop and HCI research include:

- Suchman [1987]'s study exemplifies a fruitful approach to observing natural interactions in the lab (cf. [Rooksby, 2013]). This thesis provides evidence that Suchman's approach can provide rich data, even from the constrained setting of the lab (Studies 1 and 3).

7.3.3 Practical outcomes

Our research has had the following *practical outcomes for NIME research*; we have:

- Developed software for ambisonics spatialisation in tabletop settings (Study 3).

- Produced a set of coding schemes for video annotation, with the aim of understanding group interactions on musical tabletops within two contexts: music performance (Studies 1 and 3) and public settings (Study 2).

- Produced a set of design considerations for the design of musical tabletop systems that will support:

    - Both beginners and experts (Studies 2 and 3).

    - An ecological approach to the musical experience (Studies 2 and 3).

    - Practice-based group collaboration as an open activity (Studies 1 and 2).

- Developed a methodology suitable for understanding group interactions in musical tabletops (Studies 1, 2, and 3).

Our research has had the following ***practical outcomes for tabletop and HCI research***; we have:

- Produced a set of design recommendations for the design of tabletop and TUI systems that can support:

    - Long-term use (Studies 1 and 2).

    - Fluid between-group interaction in public settings (Study 2).

    - A constructivist and hands-on learning (Studies 1 and 2).

    - A multimodal immersive experience (Studies 1, 2, and 3).

- Produced a set of coding schemes for video annotation that can be useful not only in NIME research, but also in tabletop and HCI research for understanding non-verbal practice-based interaction in tabletop settings (Studies 1, 2, and 3).

7.4   Future work

Musical tabletops are a promising platform for collaborative music making but they are still in their infancy. To fully understand the challenges and opportunities that these systems can offer, further studies at different levels of the design and evaluation of TTIs for music performance are required. Areas of further research that have not been covered yet, but should be investigated in the future, are:

Tangible user interfaces (TUIs). It is argued in some current work that the development of smart user interfaces based on transformable materials have rendered first generation of TUIs obsolete [Minuto et al., 2012]. However, one future direction suggested by Ishii et al. [2012] is to join both worlds. As our results have shown, musical tabletops have potential for group collaboration.

A next generation of tabletop systems can explore alternative designs for tangible objects including material-driven objects as tangible input, or objects with embedded information for hands-on activities. Tangible 3D tabletops have been recently emerged [Dalsgaard & Halskov, 2014], which use 3D representations of the objects, and which together with an ambisonic 3D set-up could promote 3D navigation. Further work on self-explanatory TUIs also seems promising: a way of integrating symbolic markers that are both readable for computers and humans. Finally, designing multi-purpose collections of tangible objects, as proposed by Zuckerman et al. [2005], seems useful: digital MiMs are a versatile and modular collection that allows for building different simulations (e.g. probability distributions or dynamic behaviours). Progress in the design of TUIs would benefit WA with TTIs in early interactions.

Tabletop Tangible Interfaces (TTIs). This research focused on collaborative work using public spaces only i.e. players lack private spaces in which to work their musical contributions in privacy from the other group members. Further research can include private spaces for individuals to make their musical contributions, in line with Fencott & Bryan-Kinns [2010]'s research. A combination of a tabletop with the group members' tablets or mobile devices seems promising for this (as already highlighted in tabletops for CSCW [Friess et al., 2012; Morris, 2006]), as they are common everyday devices that would not require additional hardware development. Furthermore, in public settings, it could support learning activities, in line with Walker [2010]'s research into the use of mobile devices to support learning in public settings. The use of audio icons linked to everyday listening was initiated by Gaver [1993]; further research in this direction supported by a tabletop interface seems promising both for supporting tabletop interactions, but also for learning about everyday listening in a suitable environment. Moreover, further research can develop new forms of co-located asynchronous collaboration in public settings, in line with Hindmarsh et al. [2005]'s work on social interaction between strangers in public settings. Lastly, one current issue with musical tabletops is their bulky size, which generally means that situated interactions can only take place in fixed settings. Future research on building a

more portable tabletop interface under the same principles seems interesting for casual on-the-go interaction. For example, a tablet with finger-size objects could be used. These future directions point to the increasing integration of TTIs into the environment, and social contexts.

CSCM systems. This study points to systems that support both beginners and experts under democratic principles. More research also needs to be done, to explore alternative interfaces to the Reactable and SoundXY4. Furthermore, the results of this study showed that the role of musical tabletops in public settings and education seems promising for demonstrating concepts ranging from computer music, to sonic design, and to everyday listening. However, more work needs to be done into understanding how best to deliver this musical practical knowledge. Finally, our long-term study was rich and time-consuming. Future research could continue the study of real uses of musical tabletops in the long term with different group profiles, such as beginners or mixed groups beyond a PhD timeframe. Musical tabletops are promising for the the CSCM area, yet there is still a scarcity of studies on the musical practice that they afford.

More immersive experiences in real contexts. Further research on how to best support auditory feedback in different contexts, from music venues to public venues, is needed, including the role of headphones vs surround sound systems. Research on providing multimodal feedback in TTIs (e.g. auditory, visual, haptic) appears interesting to enhance the group experience and to better support WA with these systems. Lastly, more research in the music production market using TTIs and spatialisation could be conducted, including delivery of more immersive experiences in music performances for both performers and audience, and more immersive sonic design activities using TTIs as an audio design tool.

7.5 Closing remarks

A new generation of tabletop systems for music performance has emerged, but we are only beginning to understand its potential for supporting group

collaboration in professional and casual contexts. This thesis contributes to that understanding. Focusing on the social and interactional aspects of tabletop music performance provided fruitful insights into the nature of group tabletop interaction and group awareness in collaborative musical activity on a TTI. Observing groups of collaborating music makers through this lens allowed us to develop and analyse tabletop systems that enable beginners and experts to work together in a social, situated, practice-based, and democratic way. Our empirical work has provided valuable data for future tabletop and CSCM design.

This thesis also aimed at contributing to a modernisation of music education and innovation in music creation. Tabletop systems in music education can encourage learners to explore and discover the potential of performing computer music in hands-on collaboration. This approach can facilitate their understanding of contemporary representations of music and provide them with tools and techniques for collaborative music creation, which could impact on society in the long term by applying these concepts in other domains. We hope that the ideas presented in this thesis might also inspire new musical instruments that will support new forms of collaboration in music creation, and produce new artistic results.

References

Akkermans, V., Font, F., Funollet, J., De Jong, B., Roma, G., Togias, S., & Serra, X. (2011, October). Freesound 2: An improved platform for sharing audio clips. In A. Klapuri & C. Leider (Eds.), *ISMIR 2011: Proceedings of the 12th International Society for Music Information Retrieval Conference, 24–28 October 2011, Miami, FL.* Miami, FL: University of Miami.

Allen, S. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education*, *88*(Suppl. 1), S17–S33.

Bailey, D. (1980/1993). *Improvisation: Its nature and practice in music*. New York: Da Capo Press. (Original work published 1980)

Bakker, S., van den Hoven, E., & Antle, A. N. (2011, January). MoSo tangibles: Evaluating embodied learning. In M. D. Gross, N. J. Nunes, E. Y. Do, S. A. Brewster, & I. Oakley (Eds.), *TEI '11: Proceedings of the 5th International Conference on Tangible and Embedded Interaction, 22–26 January 2011, Funchal, Madeira, Portugal* (pp. 85–92). New York: ACM.

Barbosa, A. (2003). Displaced soundscapes: A survey of network systems for music and sonic art creation. *Leonardo Music Journal*, *13*, 53–59.

Barbosa, A. (2006). *Displaced soundscapes: CSCW for music applications*. PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain.

Berry, R., Makino, M., Hikawa, N., & Suzuki, M. (2006, October). The augmented composer project: The Music Table. In *ISMAR '03: Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality,*

7–10 October 2003, Tokyo (pp. 338–339). Los Alamitos, CA: IEEE Computer Society.

Bischof, M., Conradi, B., Lachenmaier, P., Linde, K., Meier, M., Pötzl, P., & André, E. (2008, February). Xenakis: Combining tangible interaction with probability-based musical composition. In A. Schmidt, H. Gellersen, E. van den Hoven, A. Mazalek, P. Holleis, & N. Villar (Eds.), *TEI '08: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, 18–20 February 2008, Bonn, Germany* (pp. 121–124). New York: ACM.

Bischoff, J., Gold, R., & Horton, J. (1978). Music for an interactive network of microcomputers. *Computer Music Journal*, *2*(3), 24–29.

Blaine, T., & Fels, S. (2003). Collaborative musical experiences for novices. *Journal of New Music Research*, *32*(4), 411–428.

Blaine, T., & Perkis, T. (2000, August). The Jam-O-Drum interactive music system: A study in interaction design. In D. Boyarski & W. A. Kellogg (Eds.), *DIS '00: Proceedings of the 3rd Conference on Designing Interactive Systems, 17–19 August 2000, New York* (pp. 165–173). New York: ACM.

Bødker, S. (2006, October). When second wave HCI meets third wave challenges. In A. Mørch, K. Morgan, T. Bratteteig, G. Ghosh, & D. Svanaes (Eds.), *NordiCHI '06: Proceedings of the 4th Nordic Conference on Human-Computer Interaction, 14–18 October 2006, Oslo, Norway* (pp. 1–8). New York: ACM.

Bongers, B. (1998). An interview with Sensorband. *Computer Music Journal*, *22*(1), 13–24.

Booth, G., & Gurevich, M. (2012, May). Collaborative composition and socially constructed instruments: Ensemble laptop performance through the lens of ethnography. In G. Essl, B. Gillespie, M. Gurevich, & S. O'Modhrain (Eds.), *NIME '12: Proceedings of the 12th International Conference on New Interfaces for Musical Expression, 21–23 May 2012, Ann Arbor, MI* (pp. 61–65). Ann Arbor, MI: University of Michigan.

Bouwer, A., Holland, S., & Dalgleish, M. (2013). Song walker harmony space: Embodied interaction design for complex musical skills. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 207–221). London: Springer-Verlag.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Broadhurst, S. (2006). Digital practices: An aesthetic and neuroesthetic approach to virtuality and embodiment. *Performance Research*, *11*(4), 137–147.

Brown, A. R. (2006). Code jamming. *M/C Journal*, *9*(6). Retrieved 30 September 2014, from `http://journal.media-culture.org.au/0612/03-brown.php`.

Bryan-Kinns, N. (2013). Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, *71*(1), 76–90.

Bryan-Kinns, N., & Hamilton, F. (2009). Identifying mutual engagement. *Behaviour & Information Technology*, 1–25.

Burns, C. (2002). Realizing Lucier and Stockhausen: Case studies in the performance practice of electroacoustic music. *Journal of New Music Research*, *31*(1), 59–69.

Buxton, B. (1997). Artists and the art of the luthier. *Computer Graphics: The SIGGRAPH Quarterly*, *31*(1), 10–11.

Buxton, B. (2007). *Sketching user experiences: Getting the design right and the right design*. Burlington, MA: Morgan Kaufmann.

Cage, J. (1961). *Silence: Lectures and writings*. Middletown, CT: Wesleyan University Press.

Chadabe, J. (1984). Interactive composing: An overview. *Computer Music Journal*, *8*(1), 22–27.

Collins, N. (2006). *Towards Autonomous Agents for Live Computer Music: Real-time Machine Listening and Interactive Music Systems*. PhD thesis, University of Cambridge, Cambridge, UK.

Collins, N., McLean, A., Rohrhuber, J., & Ward, A. (2007). Live coding techniques for laptop performance. *Organised Sound*, *8*(3), 321–330.

Costanza, E., Shelley, S. B., & Robinson, J. (2003, September). Introducing Audio d-touch: A tangible user interface for music composition and performance. In M. Davies (Ed.), *DAFX-03: Proceedings of the 6th International Conference on Digital Audio Effects, 8–11 September 2003, Queen Mary University of London, London* (pp. 1–5). London: Dept. of Electronic Engineering, Queen Mary University of London.

Cox, C., & Warner, D. (2004). *Audio culture: Readings in modern music*. New York: Continuum.

Crawley, M. J. (2005). *Statistics: An introduction using R*. Chichester, UK: John Wiley & Sons.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: Experiencing flow in work and play*. San Francisco: Jossey-Bass.

Dalsgaard, P., & Halskov, K. (2014). Tangible 3D tabletops. *Interactions*, *21*(5), 42–47.

de Campo, A. (2014). Republic: Collaborative live coding 2003–2013. In A. Blackwell, A. McLean, J. Noble, & J. Rohrhuber (Eds.), *Collaboration and learning through live coding (Dagstuhl Seminar 13382)* (pp. 152–153). Dagstuhl, Germany.

Dietz, P., & Leigh, D. (2001, November). DiamondTouch: A multi-user touch technology. In J. Marks & E. Mynatt (Eds.), *UIST '01: Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology, 11–14 November 2001, Orlando, FL* (pp. 219–226).

Dillenbourg, P. (1999). What do you mean by 'collaborative learning'? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (Vol. 1, pp. 1–19). Oxford, UK: Elsevier.

Dourish, P. (2001). *Where the action is: Foundations of embodied interaction.* Cambridge, MA: The MIT Press.

Dourish, P., & Bellotti, V. (1992, November). Awareness and coordination in shared workspaces. In M. Mantel & R. Baecker (Eds.), *CSCW '92: Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work, 31 October – 4 November 1992, Toronto, ON* (pp. 107–114). New York: ACM.

Easterbrook, S., Singer, J., Storey, M.-A., & Damian, D. (2007). Selecting empirical methods for software engineering research. In F. Shull & J. Singer (Eds.), *Guide to advanced empirical software engineering* (pp. 285–311). London: Springer-Verlag.

Essl, G. (2010, June). The Mobile phone ensemble as classroom. In *ICMC 2010: Proceedings of the 2010 International Computer Music Conference, 1–5 June 2010, Stony Brook University, New York* (pp. 506–509). New York: International Computer Music Association.

Fels, S. (2004). Designing for Intimacy: Creating new Interfaces for Musical Expression. *Proceedings of the IEEE, 92*(4), 672–685.

Fencott, R., & Bryan-Kinns, N. (2010, June). Hey man, you're invading my personal space! Privacy and awareness in collaborative music. In K. Beilharz, A. Johnston, S. Ferguson, & A. Y.-C. Chen (Eds.), *NIME '10: Proceedings of the 10th International Conference on New Interfaces for Musical Expression, 15–18 June 2010, Sydney, Australia* (pp. 198–203). Sydney, Australia: University of Technology Sydney.

Fencott, R., & Bryan-Kinns, N. (2012). Audio delivery and territoriality in collaborative digital musical interaction. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers (BCS HCI 2012)* (pp. 69–78). Birmingham, UK.

Fernaeus, Y., & Tholander, J. (2006, April). Finding design qualities in a tangible programming space. In R. E. Grinter, T. Rodden, P. M. Aoki, E. Cutrell, R. Jeffries, & G. M. Olson (Eds.), *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 22–27 April 2006, Montréal, QC* (pp. 447–456). New York: ACM.

Fernaeus, Y., Tholander, J., & Jonsson, M. (2008, February). Towards a new set of ideals: Consequences of the practice turn in tangible interaction. In A. Schmidt, H. Gellersen, E. van den Hoven, A. Mazalek, P. Holleis, & N. Villar (Eds.), *TEI '08: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, 18–20 February 2008, Bonn, Germany* (pp. 223–230). New York: ACM.

Fiebrink, R., Morris, D., & Morris, M. R. (2009, April). Dynamic mapping of physical controls for tabletop groupware. In D. R. Olsen Jr., R. B. Arthur, K. Hinckley, M. R. Morris, S. E. Hudson, & S. Greenberg (Eds.), *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4–9 April 2009, Boston* (pp. 471–480). New York: ACM.

Fitzmaurice, G. W. (1996). *Graspable User Interfaces*. PhD thesis, University of Toronto, Toronto, ON.

Forsyth, D. R. (2010). *Group dynamics* (5th ed.). Belmont, CA: Wadsworth.

Franceschini, A., Laney, R., & Dobbyn, C. (2014, September). Learning musical contour on a tabletop. In A. Georgaki & G. Kouroupetroglou (Eds.), *ICMC/SMC 2014: Proceedings of the 40th International Computer Music Conference joint with the 11th Sound and Music Computing Conference, 14–20 September 2014, National and Kapodistrian University of Athens, Athens, Greece* (pp. 1755–1762). San Francisco: International Computer Music Association.

Freeman, J., & Van Troyer, A. (2011). Collaborative textual improvisation in a laptop ensemble. *Computer Music Journal*, *35*(2), 8–21.

Friess, M. R., Groh, G., Klügel, N., & Kleinhans, M. (2012). A tabletop application environment for generic creativity techniques. *International Journal of Computer Information Systems and Industrial Management Applications*, *4*, 55–65.

Fröbel, F. (1885). *The education of man*. New York: A. Lovell & Company.

Gallardo, D., Julià, C. F., & Jordà, S. (2008, October). TurTan: A tangible programming language for creative exploration. In *TABLETOP 2008: 3rd IEEE*

International Workshop on Horizontal Interactive Human-Computer Systems, 1–3 October 2008, Amsterdam (pp. 89–92). Los Alamitos, CA: IEEE Computer Society.

Gaver, W. W. (1991, September). Sound support for collaboration. In L. Bannon, M. Robinson, & K. Schmidt (Eds.), *ECSCW'91: Proceedings of the 2nd European Conference on Computer-Supported Cooperative Work, 24–27 September 1991, Amsterdam* (pp. 293–308). Dordrecht/Boston: Kluwer Academic Publishers.

Gaver, W. W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, *5*(1), 1–29.

Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures* (pp. 3–30). New York: Basic Books.

Gerzon, M. (1974a). Surround-sound psychoacoustics. *Wireless World*, *80*(1468), 483–486.

Gerzon, M. (1974b). What's wrong with quadraphonics? *Studio Sound*, *16*(5), 51–56.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gresham-Lancaster, S. (1998). The aesthetics and history of the Hub: The effects of changing technology on network computer music. *Leonardo Music Journal*, *8*, 39–44.

Gurevich, M., & Cavan Fyans, A. (2011). Digital musical interactions: Performer-system relationships and their perception by spectators. *Organised Sound*, *16*(02), 166–175.

Gurevich, M., Marquez-Borbon, A., & Stapleton, P. (2012). Playing with constraints: Stylistic variation with a simple electronic instrument. *Computer Music Journal*, *36*(1), 23–41.

Gutwin, C., & Greenberg, S. (2002). A Descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, *11*(3), 411–446.

Hammersley, M. (2013). *What is qualitative research?* London: Bloomsbury.

Han, J. Y. (2005, October). Low-cost multi-touch sensing through frustrated total internal reflection. In P. Baudisch, M. Czerwinski, & D. R. Olsen (Eds.), *UIST '05: Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, 23–26 October 2005, Seattle, WA* (pp. 115–118). New York: ACM.

Hancock, M. S., Shen, C., Forlines, C., & Ryall, K. (2005, May). Exploring non-speech auditory feedback at an interactive multi-user tabletop. In K. Inkpen & M. Panne (Eds.), *GI '05: Proceedings of Graphics Interface 2005, 9–11 May 2005, Victoria, BC* (pp. 41–50). Waterloo, ON: Canadian Human-Computer Communications Society School of Computer Science, University of Waterloo.

Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009, June). Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *CSCL'09: Proceedings of the 9th International Conference on Computer Supported Collaborative Learning: Volume 1, 8–13 June 2009, University of the Aegean, Rhodes, Greece* (pp. 335–344). International Society of the Learning Sciences.

Harrison, S., Tatar, D., & Sengers, P. (2007, April–May). The three paradigms of HCI. In M. B. Rosson & D. J. Gilmore (Eds.), *Alt.Chi Session. CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 28 April – 3 May 2007, San Jose, CA* (pp. 1–18). New York: ACM.

Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in qualitative research.* London: SAGE Publications.

Heath, C., & vom Lehn, D. (2009). Interactivity and collaboration: New forms of participation in museums, galleries and science centres. In R. Parry (Ed.), *Museums in a digital age* (pp. 266–280). Oxford, UK: Routledge.

Hindmarsh, J., Heath, C., vom Lehn, D., & Cleverly, J. (2005). Creating assemblies in public environments: Social interaction, interactive exhibits and CSCW. *Computer Supported Cooperative Work (CSCW)*, *14*(1), 1–41.

Hinrichs, U., & Carpendale, S. (2011, May). Gestures in the wild: Studying multi-touch gesture sequences on interactive tabletop exhibits. In D. S. Tan, S. Amershi, B. Begole, W. A. Kellogg, & M. Tungare (Eds.), *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 7–12 May 2011, Vancouver, BC* (pp. 3023–3032). New York: ACM.

Holland, S., Marshall, P., Bird, J., Dalton, S., Morris, R., Pantidi, N., . . . Clark, A. (2009, February). Running up Blueberry Hill: Prototyping whole body interaction in harmony space. In N. Villar, S. Izadi, M. Fraser, & S. Benford (Eds.), *TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction, 16–18 February 2009, Cambridge, UK* (pp. 93–98). New York: ACM.

Holland, S., Wilkie, K., Mulholland, P., & Seago, A. (Eds.). (2013). *Music and Human-Computer Interaction*. London: Springer-Verlag.

Hornecker, E. (2008, October). "I don't understand it either, but it is cool" Visitor interactions with a multi-touch table in a museum. In *TABLETOP 2008: 3rd IEEE International Workshop on Horizontal Interactive Human-Computer Systems, 1–3 October 2008, Amsterdam* (pp. 121–128). Los Alamitos, CA: IEEE Computer Society.

Hornecker, E., & Buur, J. (2006, April). Getting a grip on tangible interaction: A framework on physical space and social interaction. In R. E. Grinter, T. Rodden, P. M. Aoki, E. Cutrell, R. Jeffries, & G. M. Olson (Eds.), *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 22–27 April 2006, Montréal, QC* (pp. 437–446). New York: ACM.

Hornecker, E., Marshall, P., Dalton, N. S., & Rogers, Y. (2008, November). Collaboration and interference: Awareness with mice or touch input. In B. Begole & D. W. McDonald (Eds.), *CSCW '08: Proceedings of the 2008 ACM Conference on Computer-Supported Cooperative Work, 8–12 November 2008, San Diego, CA* (pp. 167–176). New York: ACM.

Hornecker, E., & Stifter, M. (2006, November). Learning from interactive museum installations about interaction design for public settings. In J. Kjeldskov & J. Paay (Eds.), *OZCHI '06: Proceedings of the 18th Australia Conference on Computer-Human Interaction, 20–24 November 2006, Sydney, Australia* (pp. 135–142). New York: ACM.

Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: The MIT Press.

Ishii, H., Lakatos, D., Bonanni, L., & Labrune, J.-B. (2012). Radical atoms: Beyond tangible bits, toward transformable materials. *Interactions*, *19*(1), 38–51.

Ishii, H., & Ullmer, B. (1997, March). Tangible bits: Towards seamless interfaces between people, bits and atoms. In S. Pemberton (Ed.), *CHI '97: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, 22–27 March 1997, Atlanta, GA* (pp. 234–241). New York: ACM/Addison-Wesley.

Iwai, T. (1999, August). Composition on the Table. In A. Rockwood (Ed.), *SIGGRAPH '99: Proceedings of the 26th International Conference on Computer Graphics and Interactive Techniques, 8–13 August 1999, Los Angeles* (p. 10). New York: ACM Press/Addison-Wesley Publishing.

Jacob, R. J. K., Girouard, A., Horn, M. S., & Zigelbaum, J. (2008, April). Reality-based interaction: A framework for post-WIMP interfaces. In M. Czerwinski, A. M. Lund, & D. S. Tan (Eds.), *CHI '08: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 5–10 April 2008, Florence, Italy* (pp. 201–210). New York: ACM.

Jensenius, A. R., Wanderley, M. M., Godøy, R. I., & Leman, M. (2010). Musical gestures: Concepts and methods in research. In R. I. Godøy & M. Leman (Eds.), *Musical gestures: Sound, movement, and meaning* (pp. 12–35). New York: Routledge.

Jewitt, C. (2013). Multimodal methods for researching digital technologies. In B. Price S. Brown & C. Jewitt (Eds.), *The SAGE handbook of digital technology research* (pp. 250–265). London: SAGE Publications.

Jordà, S. (2004, November). Digital instruments and players: Part II–diversity, freedom and control. In G. Tzanetakis, G. Essl, & C. Leider (Eds.), *ICMC 2004: Proceedings of the 30th Annual International Computer Music Conference, 1–6 November 2004, University of Miami, Miami, FL* (pp. 706–710). San Francisco: International Computer Music Association.

Jordà, S. (2005a). *Digital Lutherie Crafting musical computers for new musics' performance and improvisation*. PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain.

Jordà, S. (2005b, May). Multi-user instruments: Models, examples and promises. In S. Fels & T. Blaine (Eds.), *NIME '05: Proceedings of the 2005 Conference on New Interfaces for Musical Expression, 26–28 May 2005, Vancouver, BC* (pp. 23–26). Vancouver, BC: University of British Columbia, Media and Graphics Interdisciplinary Center (MAGIC).

Jordà, S. (2007). Interactivity and live computer music. In N. Collins & J. d'Escriván (Eds.), *The Cambridge companion to electronic music* (4th ed., pp. 89–106). Cambridge, UK: Cambridge University Press.

Jordà, S. (2008). On stage: the Reactable and other musical tangibles go real. *International Journal of Arts and Technology*, *1*(3/4), 268–287.

Jordà, S., Kaltenbrunner, M., Geiger, G., & Bencina, R. (2005, September). The reacTable*. In A. Lewin-Richter & X. Serra (Eds.), *ICMC 2005: Proceedings of the 2005 International Computer Music Conference, 5–9 September 2005, Escola Superior de Musica de Catalunya, Barcelona, Spain* (pp. 579–582). San Francisco: International Computer Music Association.

Jordà, S., & Mealla, S. (2014, June–July). A methodological framework for teaching, evaluating and informing NIME design with a focus on expressiveness and mapping. In B. Caramiaux, K. Tahiroglu, R. Fiebrink, & A. Tanaka (Eds.), *NIME '14: Proceedings of the 14th International Conference on New Interfaces for Musical Expression, 30 June – 4 July 2014, London* (pp. 233–238). London: Goldsmiths, University of London.

Jordan, B. (1996). Ethnographic workplace studies and CSCW. *Human Factors in Information Technology*, *12*, 17–42.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39–103.

Julià, C. F., & Jordà, S. (2009, October). Songexplorer: A tabletop application for exploring large collections of songs. In K. Hirata, G. Tzanetakis, & K. Yoshii (Eds.), *ISMIR 2009: Proceedings of the 10th International Society for Music Information Retrieval Conference, 26–30 October 2009, Kobe International Conference Center, Kobe, Japan* (pp. 675–680). International Society for Music Information Retrieval.

Kaltenbrunner, M., & Bencina, R. (2007, February). reacTIVision: A computer-vision framework for table-based tangible interaction. In B. Ullmer & A. Schmidt (Eds.), *TEI '07: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, 15–17 February 2007, Baton Rouge, LA* (pp. 69–74). New York: ACM.

Kaltenbrunner, M., Bovermann, T., Bencina, R., & Costanza, E. (2005, May). TUIO: A protocol for table-top tangible user interfaces. In *GW 2005: Proceedings of the 6th international Gesture Workshop, 18–20 May 2005, Berder Island, France*.

Kaltenbrunner, M., Geiger, G., & Jordà, S. (2004, June). Dynamic patches for live musical performance. In Y. Nagashima & M. J. Lyons (Eds.), *NIME '04: Proceedings of the 2004 Conference on New Interfaces for Musical Expression, 3–5 June 2004, Hamamatsu, Japan* (pp. 19–22). Hamamatsu, Japan: Shizuoka University of Art and Culture.

Kaltenbrunner, M., O'Modhrain, S., & Costanza, E. (2004, May). Object design considerations for tangible musical interfaces. In K. Ng (Ed.), *GIMS2006: Proceedings of COST287-ConGAS 2nd International Symposium on Gesture Interfaces for Multimedia Systems, 9–10 May 2006, Leeds, UK*.

Kiefer, C. (2012). *Multiparametric Interfaces For Fine-Grained Control of Digital Music.* PhD thesis, University Of Sussex, Brighton, UK.

Kirsh, D. (2010). Thinking with external representations. *AI and Society*, 25(4), 441–454.

Klemmer, S. R., Hartmann, B., & Takayama, L. (2006, June). How bodies matter: Five themes for interaction design. In J. M. Carroll, S. Bødker, & J. Coughlin (Eds.), *DIS '06: Proceedings of the 6th Conference on Designing Interactive Systems, 26–28 June 2006, University Park, PA* (pp. 140–149). New York: ACM.

Klügel, N., Hagerer, G., & Groh, G. (2014, June–July). TreeQuencer: Collaborative rhythm sequencing: A comparative study. In B. Caramiaux, K. Tahiroglu, R. Fiebrink, & A. Tanaka (Eds.), *NIME '14: Proceedings of the 14th International Conference on New Interfaces for Musical Expression, 30 June – 4 July 2014, London* (pp. 50–53). London: Goldsmiths, University of London.

Klügel, N., Lindström, A., & Groh, G. (2014, September). A genetic algorithm approach to collaborative music creation on a multi-touch table. In A. Georgaki & G. Kouroupetroglou (Eds.), *ICMC/SMC 2014: Proceedings of the 40th International Computer Music Conference joint with the 11th Sound and Music Computing Conference, 14–20 September 2014, National and Kapodistrian University of Athens, Athens, Greece* (pp. 286–292). San Francisco: International Computer Music Association.

Korkiakangas, T., Weldon, S.-M., & Bezemer, J. (2014). Factors affecting nurse-surgeon object transfer: Video analysis of inter-professional communication and situation awareness in the operating theatre. *International Journal of Nursing Studies*.

Kreitmayer, S., Rogers, Y., Laney, R., & Peake, S. (2012, May). From participatory to contributory simulations: Changing the game in the classroom. In J. A. Konstan, E. H. Chi, & K. Höök (Eds.), *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 5–10 May 2012, Austin, TX* (pp. 49–58). New York: ACM.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.

Lee, S. W., & Freeman, J. (2013, May). echobo: A mobile music instrument designed for audience to play. In W. S. Yeo & K. Lee (Eds.), *NIME '13: Proceedings of the 14th International Conference on New Interfaces for Musical Expression, 27–30 May 2013, Daejeon + Seoul, Korea* (pp. 450–455).

Levin, G. (2006, November). The table is the score: An augmented reality interface for real time, tangible, spectrographic performance. In G. Essl & I. Fujinaga (Eds.), *ICMC 2006: Proceedings of the 2006 International Computer Music Conference, 6–11 November 2006, New Orleans, LA.* San Francisco: International Computer Music Association & Tulane University.

Linson, A. (2014). *Investigating the cognitive foundations of collaborative musical free improvisation: Experimental case studies using a novel application of the subsumption architecture.* PhD thesis, The Open University, Milton Keynes, UK.

Lofland, J., Snow, D. A., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings.* Belmont, CA: Thomson Wadsworth.

Lopes, P., Jota, R., & Jorge, J. A. (2011, November). Augmenting touch interaction through acoustic sensing. In J. Rekimoto, H. Koike, K. Fukuchi, Y. Kitamura, & D. Wigdor (Eds.), *ITS '11: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 13–16 November 2011, Kobe, Japan* (pp. 53–56). New York: ACM.

MacDonald, L. (2014, June). For every action: Interactive installations in liminal spaces. In R. Wakkary, S. Harrison, C. Neustaedter, S. Bardzell, & E. Paulos (Eds.), *DIS Companion '14: Proceedings of the 2014 Companion Publication on Designing Interactive Systems, 21–25 June 2014, Vancouver, BC* (pp. 181–184). New York: ACM.

Magnusson, T. (2010). Designing constraints: Composing and performing with digital musical systems. *Computer Music Journal*, *34*(4), 62–73.

Marquardt, N., Nacenta, M. A., Young, J. E., Carpendale, S., Greenberg, S., & Sharlin, E. (2009, November). The haptic tabletop puck: Tactile feedback

for interactive tabletops. In G. Morrison, S. Subramanian, M. S. T. Carpendale, M. Haller, & S. D. Scott (Eds.), *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, 23–25 November 2009, Banff/Calgary, AL* (pp. 85–92). New York: ACM.

Marshall, P., & Hornecker, E. (2013). Theories of embodiment in HCI. In C. Jewitt, S. Price, & B. Brown (Eds.), *The SAGE handbook of digital technology research* (pp. 144–158). London: SAGE Publications.

Marshall, P., Morris, R., Rogers, Y., Kreitmayer, S., & Davies, M. (2011, May). Rethinking 'multi-user': An in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In D. S. Tan, S. Amershi, B. Begole, W. A. Kellogg, & M. Tungare (Eds.), *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 7–12 May 2011, Vancouver, BC* (pp. 3033–3042). New York: ACM.

Mavers, D. (2012). *Transcribing video.* NCRM Working Paper. Retrieved 30 September 2014, from `http://eprints.ncrm.ac.uk/2877/4/NCRM _working_paper0512.pdf`.

McCartney, J. (2002). Rethinking the computer music language: SuperCollider. *Computer Music Journal*, *26*(4), 61–68.

Medler, B., St, C., & Magerko, B. (2010, April). The implications of improvisational acting and role-playing on design methodologies. In E. D. Mynatt, D. Schoner, G. Fitzpatrick, S. E. Hudson, W. K. Edwards, & T. Rodden (Eds.), *CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 10–15 April 2010, Atlanta, GA* (pp. 483–492). New York: ACM.

Merrill, D., Kalanithi, J., & Maes, P. (2007, February). Siftables: Towards sensor network user interfaces. In B. Ullmer & A. Schmidt (Eds.), *TEI '07: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, 15–17 February 2007, Baton Rouge, LA* (pp. 75–78). New York: ACM.

Milne, A. J., Xambó, A., Laney, R., Sharp, D. B., Prechtl, A., & Holland, S. (2011, May–June). Hex Player: A virtual musical controller. In A. R. Jensenius, A. Tveit, R. I. Godøy, & D. Overholt (Eds.), *NIME '11: Proceedings of*

the 11th International Conference on New Interfaces for Musical Expression, 30 May – 1 June 2011, Oslo, Norway (pp. 244–247). Oslo, Norway: University of Oslo, Norwegian Academy of Music.

Minuto, A., Vyas, D., Poelman, W., & Nijholt, A. (2012). Smart material interfaces: A vision. In A. Nijholt, H. Hondorp, & D. Reidsma (Eds.), *Intelligent technologies for interactive entertainment* (pp. 57–62). Berlin/Heidelberg: Springer-Verlag.

Miranda, E. R., & Wanderley, M. M. (2006). *New digital musical instruments: Control and interaction beyond the keyboard*. Middleton, WI: A-R Editions.

Monson, I. (1996). *Saying something: Jazz improvisation and interaction*. Chicago: University of Chicago Press.

Montessori, M. (1912). *The Montessori method*. New York: Frederick A. Stokes.

Morris, M. R. (2006). *Supporting effective interaction with tabletop groupware*. PhD thesis, Stanford University, Stanford, CA.

Morris, M. R., Morris, D., & Winograd, T. (2004, November). Individual audio channels with single display groupware: Effects on communication and task strategy. In J. D. Herbsleb & G. M. Olson (Eds.), *CSCW '04: Proceedings of the 2004 ACM Conference on Computer-Supported Cooperative Work, 6–10 November 2004, Chicago* (pp. 242–251). New York: ACM.

Mulder, J. (2010, June). The loudspeaker as musical instrument. In K. Beilharz, A. Johnston, S. Ferguson, & A. Y.-C. Chen (Eds.), *NIME '10: Proceedings of the 10th International Conference on New Interfaces for Musical Expression, 15–18 June 2010, Sydney, Australia* (pp. 13–18). Sydney, Australia: University of Technology Sydney.

Müller-Tomfelde, C. (Ed.). (2010). *Tabletops: Horizontal interactive displays*. London: Springer-Verlag.

Nacenta, M. A., Pinelle, D., Gutwin, C., & Mandryk, R. (2010). Individual and group support in tabletop interaction techniques. In C. Müller-Tomfelde (Ed.), *Tabletops: Horizontal interactive displays* (pp. 303–333). London: Springer-Verlag.

Nardi, B. A., Kuchinsky, A., Whittaker, S., Leichner, R., & Schwarz, H. (1995). Video-as-data: Technical and social aspects of a collaborative multimedia application. *Computer Supported Cooperative Work (CSCW)*, *4*(1), 73–100.

Neale, D. C., Carroll, J. M., & Rosson, M. B. (2004, November). Evaluating computer-supported cooperative work: Models and frameworks. In J. D. Herbsleb & G. M. Olson (Eds.), *CSCW '04: Proceedings of the 2004 ACM Conference on Computer-Supported Cooperative Work, 6–10 November 2004, Chicago* (pp. 112–121). New York: ACM.

Newton-Dunn, H., Nakano, H., & Gibson, J. (2003, May). Block Jam: A tangible interface for interactive music. In F. Thibault (Ed.), *NIME '03: Proceedings of the 2003 Conference on New Interfaces for Musical Expression, 22–24 May 2003, Montréal, QC* (pp. 170–177). Montreal, QC: Faculty of Music, McGill University.

NUI Group Authors. (2009). *Multi-touch technologies* [e-book]. NUI Group. Retrieved 30 September 2014, from `http://nuicode.com/attachments/download/115/Multi-Touch_Technologies_v1.01.pdf`.

Oh, J., Herrera, J., Bryan, N. J., Dahl, L., & Wang, G. (2010, June). Evolving the mobile phone orchestra. In K. Beilharz, A. Johnston, S. Ferguson, & A. Y.-C. Chen (Eds.), *NIME '10: Proceedings of the 10th International Conference on New Interfaces for Musical Expression, 15–18 June 2010, Sydney, Australia* (pp. 82–87). Sydney, Australia: University of Technology Sydney.

O'Modhrain, S. (2011). A framework for the evaluation of digital musical instruments. *Computer Music Journal*, *35*(1), 28–42.

Oppenheimer, F. (1968). A rationale for a science museum. *Curator*, *11*, 206–209.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Patten, J., Recht, B., & Ishii, H. (2002, May). Audiopad: A tag-based interface for musical performance. In E. Brazil (Ed.), *NIME '02: Proceedings of the 2002*

Conference on New Interfaces for Musical Expression, 24–26 May 2002 Dublin, Ireland (pp. 1–6). Dublin, Ireland: Media Lab Europe.

Piper, A. M., & Hollan, J. D. (2009, April). Tabletop displays for small group study: Affordances of paper and digital materials. In D. R. Olsen Jr., R. B. Arthur, K. Hinckley, M. R. Morris, S. E. Hudson, & S. Greenberg (Eds.), *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4–9 April 2009, Boston* (pp. 1227–1236). New York: ACM.

Plonka, L., Segal, J., Sharp, H., & van der Linden, J. (2011, May). Collaboration in pair programming: Driving and switching. In A. Sillitti, O. Hazzan, E. Bache, & X. Albaladejo (Eds.), *XP 2011: Proceedings of the 12th International Conference on Agile Software Development: Agile Processes in Software Engineering and Extreme Programming, 10–13 May 2011, Madrid, Spain* (pp. 43–59). Berlin/Heidelberg: Springer-Verlag.

Pontual Falcão, T., & Price, S. (2010). Interfering and resolving: How tabletop interaction facilitates co-construction of argumentative knowledge. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 539–559.

Pressing, J. (1984). Cognitive processes in improvisation. In W. R. Crozier & A. J. Chapman (Eds.), *Cognitive processes in the perception of art* (pp. 345–363). Amsterdam: Elsevier.

Pressing, J. (1988). Improvisation: Methods and models. In J. A. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation and composition* (pp. 130–177). London: Oxford University Press.

Puckette, M. (2002). Max at seventeen. *Computer Music Journal*, 26(4), 31–43.

Pugliese, R., Goddard, C., & Nesfield, J. (2012, May). A qualitative evaluation of augmented human-human interaction in mobile group improvisation. In G. Essl, B. Gillespie, M. Gurevich, & S. O'Modhrain (Eds.), *NIME '12: Proceedings of the 12th International Conference on New Interfaces for Musical Expression, 21–23 May 2012, Ann Arbor, MI*. Ann Arbor, MI: University of Michigan.

Raffle, H. S., Parkes, A. J., & Ishii, H. (2004, April). Topobo: A constructive assembly system with kinetic memory. In E. Dykstra-Erickson & M. Tscheligi (Eds.), *CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 24–29 April 2004, Vienna, Austria* (pp. 647–654). New York: ACM.

Rick, J., Marshall, P., & Yuill, N. (2011, June). Beyond one-size-fits-all: How interactive tabletops support collaborative learning. In T. Moher, C. Quintana, & S. Price (Eds.), *IDC '11: Proceedings of the 10th International Conference on Interaction Design and Children, 20–23 June 2011, Ann Arbor, MI* (pp. 109–117). New York: ACM.

Roads, C. (1996). *The computer music tutorial*. Cambridge, MA: The MIT Press.

Rodgers, T. (2010). *Pink noises: Women on electronic music and sound*. Durham/London: Duke University Press.

Rogers, Y. (2012). Interaction design gone wild: Striving for wild theory. *Interactions*, *18*(4), 58–62.

Rogers, Y., & Price, S. (2004). Extending and augmenting scientific enquiry through pervasive learning environments. *Children Youth and Environments*, *14*, 67–83.

Rohrhuber, J. (2007). Network music. In N. Collins & J. d'Escriván (Eds.), *The Cambridge companion to electronic music* (4th ed., pp. 140–170). Cambridge, UK: Cambridge University Press.

Rohrhuber, J., de Campo, A., Wieser, R., van Kampen, J.-K., Ho, E., & Hölzl, H. (2007, September). Purloined letters and distributed persons. In *Music in the Global Village Conference 2007, 6–8 September 2007, Budapest, Hungary.*

Roma, G., Herrera, P., & Serra, X. (2009, April). Freesound Radio: Supporting music creation by exploration of a sound database. In *CHI 2009 Workshop on Computational Creativity Support: Using Algorithms and Machine Learning to Help People Be More Creative.* Boston.

Roma, G., & Xambó, A. (2008, June). A tabletop waveform editor for live performance. In A. Camurri, S. Serafin, & G. Volpe (Eds.), *NIME '08: Proceedings of the 8th International Conference on New Interfaces for Musical Expression, 5–7 June 2008, Genova, Italy* (pp. 249–252). Genoa, Italy: Casa Paganini-InfoMus Lab.

Rooksby, J. (2013). Wild in the laboratory: A discussion of plans and situated actions. *ACM Transactions on Computer-Human Interaction*, *20*(3), 1–17.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, *2*(3), 235–276.

Russolo, L. (1913/2004). *The art of noise* (M. Tencer, Ed. & R. Filliou, Trans.) [e-book]. Ubuclassics. Retrieved 30 September 2014, from `http://www.ubu.com/historical/gb/russolo_noise.pdf`. (Original work published 1913, and translated to English 1967)

Sawyer, R. K. (2003). *Group creativity: Music, theater, collaboration*. Hove, UK: Psychology Press.

Sawyer, R. K. (2007). *Group genius: The creative power of collaboration*. New York: Basic Books.

Schmidt, D., Ramakers, R., Pedersen, E. W., Jasper, J., Köhler, S., Pohl, A., . . . Baudisch, P. (2014, April–May). Kickables: Tangibles for feet. In M. Jones, P. A. Palanque, A. Schmidt, & T. Grossman (Eds.), *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 26 April – 1 May 2014, Toronto, ON* (pp. 3143–3152). New York: ACM.

Schofield, J. W. (1993). Increasing the generalizability of qualitative research. In M. Hammersley (Ed.), *Social research: Philosophy, politics and practice* (pp. 171–203). London: SAGE Publications.

Scott, S. D., & Carpendale, S. (2010). Theory of tabletop territoriality. In C. Müller-Tomfelde (Ed.), *Tabletops: Horizontal interactive displays* (pp. 357–385). London: Springer-Verlag.

Scott, S. D., Grant, K. D., & Mandryk, R. L. (2003, September). System guidelines for co-located, collaborative work on a tabletop display. In K. Kuutti,

E. H. Karsten, G. Fitzpatrick, P. Dourish, & K. Schmidt (Eds.), *ECSCW'03: Proceedings of the 8th European Conference on Computer-Supported Cooperative Work, 14–18 September 2003, Helsinki, Finland* (pp. 159–178). Dordrecht/-Boston: Kluwer Academic Publishers.

Scott, S. D., Sheelagh, M., Carpendale, T., & Inkpen, K. M. (2004, November). Territoriality in collaborative tabletop workspaces. In J. D. Herbsleb & G. M. Olson (Eds.), *CSCW '04: Proceedings of the 2004 ACM Conference on Computer-Supported Cooperative Work, 6–10 November 2004, Chicago* (pp. 294–303). New York: ACM.

Serafin, S., de Götzen, A., Böttcher, N., & Gelineck, S. (2006, June). Synthesis and control of everyday sounds reconstructing Russolo's Intonarumori. In N. Schnell, F. Bevilacqua, M. J. Lyons, & A. Tanaka (Eds.), *NIME '06: Proceedings of the 2006 Conference on New Interfaces for Musical Expression, 4–8 June 2006, Paris* (pp. 240–245). Paris: IRCAM - Centre Pompidou in collaboration with Sorbonne University.

Shaer, O., & Hornecker, E. (2009). Tangible user interfaces: Past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*, *3*(1-2), 1–137.

Sloetjes, H., & Wittenburg, P. (2008, May). Annotation by category: ELAN and ISO DCR. In N. Calzolari et al. (Eds.), *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)* (pp. 816–820). Marrakech, Morocco: European Language Resources Association (ELRA).

Spiegel, L. (1998). Should music-making be reserved for an elite? *Computer Music Journal*, *22*(1), 6–7.

Stowell, D., & McLean, A. (2013). Live music-making: A rich open task requires a rich open interface. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 139–152). London: Springer-Verlag.

Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge, UK: Cambridge University Press.

Swift, B. (2012, May). Becoming-sound: Affect and assemblage in improvisational digital music making. In J. A. Konstan, E. H. Chi, & K. Höök (Eds.), *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 5–10 May 2012, Austin, TX* (pp. 1815–1824). New York: ACM.

Swift, B. (2013). Chasing a feeling: Experience in computer supported jamming. In S. Holland, K. Wilkie, P. Mulholland, & A. Seago (Eds.), *Music and Human-Computer Interaction* (pp. 85–99). London: Springer-Verlag.

Tanaka, A. (2004a, September). Malleable mobile music. In N. Davies, E. Mynatt, & I. Siio (Eds.), *UBICOMP '04: Adjunct Proceedings of the 6th International Conference on Ubiquitous Computing, 7–10 September 2004, Nottingham, UK.*

Tanaka, A. (2004b, June). Mobile music making. In Y. Nagashima & M. J. Lyons (Eds.), *NIME '04: Proceedings of the 2004 Conference on New Interfaces for Musical Expression, 3–5 June 2004, Hamamatsu, Japan* (pp. 154–156). Hamamatsu, Japan: Shizuoka University of Art and Culture.

Tang, A., Tory, M., Po, B., Neumann, P., & Carpendale, S. (2006, April). Collaborative coupling over tabletop displays. In R. E. Grinter, T. Rodden, P. M. Aoki, E. Cutrell, R. Jeffries, & G. M. Olson (Eds.), *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 22–27 April 2006, Montréal, QC* (pp. 1181–1190). New York: ACM.

Tuddenham, P., & Robinson, P. (2009, April). Territorial coordination and workspace awareness in remote tabletop collaboration. In D. R. Olsen Jr., R. B. Arthur, K. Hinckley, M. R. Morris, S. E. Hudson, & S. Greenberg (Eds.), *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 4–9 April 2009, Boston* (pp. 2139–2148). New York: ACM.

Turner, V. W. (1964). Betwixt and between: The liminal period in rites de passage. In J. Helm (Ed.), *The Proceedings of the American Ethnological Society* (pp. 4–20). Seattle, WA: American Ethnological Society.

Ullmer, B., & Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal, 39*(3-4), 915–931.

Underkoffler, J., & Ishii, H. (1999, May). Urp: A luminous-tangible work-bench for urban planning and design. In M. G. Williams & M. W. Altom (Eds.), *CHI '99: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, 15–20 May 1999, Pittsburgh, PA* (pp. 386–393). New York: ACM.

Vigotsky, L. (1978). *Mind in society*. London: Harvard University Press.

Vogel, D., & Casiez, G. (2012, May). Hand occlusion on a multi-touch table-top. In J. A. Konstan, E. H. Chi, & K. Höök (Eds.), *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 5–10 May 2012, Austin, TX* (pp. 2307–2316). New York: ACM.

vom Lehn, D. (2006). Embodying experience: A video-based examination of visitors' conduct and interaction in museums. *European Journal of Marketing*, *40*(11), 1340–1359.

vom Lehn, D., Heath, C., & Hindmarsh, J. (2001). Exhibiting interaction: Conduct and collaboration in museums and galleries. *Symbolic Interaction*, *24*(2), 189–216.

Walker, K. (2010). *Designing for meaning making in museums: Visitor-constructed trails using mobile digital technologies*. PhD thesis, Institute of Education, University of London, London.

Wanderley, M. M., & Vines, B. (2006). Origins and functions of clarinettists ancillary gestures. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 165–191). Hampshire, UK: Ashgate Publishing.

Wang, G. (2008). *The ChucK Audio Programming Language: A Strongly-timed and On-the-fly Environ/mentality*. PhD thesis, Princeton University, New Jersey.

Wang, G., Bryan, N., Oh, J., & Hamilton, R. (2009, August). Stanford Laptop Orchestra (SLOrk). In G. Scavone, V. Verfaille, & A. da Silva (Eds.), *ICMC 2009: Proceedings of the 2009 International Computer Music Conference, 16–21 August 2009, McGill University, Montreal, QC* (pp. 505–508). San Francisco: International Computer Music Association.

Wang, G., Essl, G., & Penttinen, H. (2008, August). Do mobile phones dream of electric orchestras? In *ICMC 2008: Proceedings of the 2008 International Computer Music Conference, 24–29 August 2008, Belfast, Northern Ireland.* Belfast, Northern Ireland: International Computer Music Association & Sonic Arts Research Centre, Queen's University.

Wang, Q., Battocchi, A., Graziola, I., Pianesi, F., Tomasini, D., Zancanaro, M., & Nass, C. (2006, November). The role of psychological ownership and ownership markers in collaborative working environment. In F. K. H. Quek, J. Yang, D. W. Massaro, A. A. Alwan, & T. J. Hazen (Eds.), *ICMI' 06: Proceedings of the 8th International Conference on Multimodal Interfaces, 2–4 November 2006, Banff, AL* (pp. 225–232). New York: ACM.

Weinberg, G. (2005). Interconnected musical networks: Toward a theoretical framework. *Computer Music Journal*, *29*(2), 23–39.

Weinberg, G., Aimi, R., & Jennings, K. (2002, May). The Beatbug network: A rhythmic system for interdependent group collaboration. In E. Brazil (Ed.), *NIME '02: Proceedings of the 2002 Conference on New Interfaces for Musical Expression, 24–26 May 2002 Dublin, Ireland* (pp. 1–6). Dublin, Ireland: Media Lab Europe.

Weinberg, G., & Driscoll, S. (2006). Toward robotic musicianship. *Computer Music Journal*, *30*(4), 28–45.

Weinberg, G., & Gan, S.-L. (2001). The Squeezables: Toward an expressive and interdependent multi-player musical instrument. *Computer Music Journal*, *25*(2), 37–45.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94–104.

Westerkamp, H. (2002). Linking soundscape composition and acoustic ecology. *Organised Sound*, *7*(1), 1355–7718.

Wigdor, D., Penn, G., Ryall, K., Esenther, A., & Shen, C. (2007, October). Living with a tabletop: Analysis and observations of long term office use of a multi-touch table. In *TABLETOP 2007: 2nd IEEE International Workshop*

on Horizontal Interactive Human-Computer Systems, 10–12 October 2007, Newport, RI (pp. 60–67). Los Alamitos, CA: IEEE Computer Society.

Wigdor, D., Williams, S., Cronin, M., Levy, R., White, K., Mazeev, M., & Benko, H. (2009, October). Ripples: Utilizing per-contact visualizations to improve user interaction with touch displays. In A. D. Wilson & F. Guimbretière (Eds.), *UIST '09: Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, 4–7 October 2009, Victoria, BC* (pp. 3–12). New York: ACM.

Wilkie, K., Holland, S., & Mulholland, P. (2010). What can the language of musicians tell us about music interaction design? *Computer Music Journal*, *34*(4), 34–48.

Wright, M., & Freed, A. (1997, September). Open Sound Control: A new protocol for communicating with sound synthesizers. In P. R. Cook & A. Bakalakos (Eds.), *ICMC 1997: Proceedings of the 1997 International Computer Music Conference, 25–30 September 1997, Aristotle University, Thessaloniki, Greece* (pp. 101–104). San Francisco: International Computer Music Association.

Yuill, N., & Rogers, Y. (2012). Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM Transactions on Computer-Human Interaction*, *19*(1), 1–25.

Zuckerman, O., Arida, S., & Resnick, M. (2005, April). Extending tangible interfaces for education: Digital Montessori-inspired manipulatives. In G. C. van der Veer & C. Gale (Eds.), *CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2–7 April 2005, Portland, OR* (pp. 859–868). New York: ACM.

Appendix A

Methodology Materials

A.1   Tabletop tangible technologies

A.1.1   What are the tabletop tangible technologies available?

There exist a number of tabletop tangible and multitouch technologies, which support either object tracking or multitouch recognition, or both. Multitouch technologies include *Frustrated Total Internal Reflection* (FTIR) or *Diffused Illumination* (DI), among others.

FTIR,[1] is a multitouch technology, developed by Han [2005], and demonstrated on a TED Talk in 2007.[2] In FTIR, there is an IR camera below the tabletop surface, and an acrylic surface works with IR light located in its sides shining into it, in which the light is trapped (total internal reflection). By pressing with a fingertip on the surface, the total internal reflection is frustrated so the light is scattered down and the IR camera can sense it. It is recommended to use a compliant surface on top of the acrylic (for example, a silicone rubber layer) to increase touch sensitivity and make the dragging smoother. A projection surface will also be needed, which can be the same compliant surface. FTIR does not track objects. Figure A.1 illustrates how FTIR works.

---

[1] `http://wiki.nuigroup.com/FTIR` (accessed 30 September 2014).
[2] `www.youtube.com/watch?v=QKh1Rv0PlOQ` (accessed 30 September 2014).
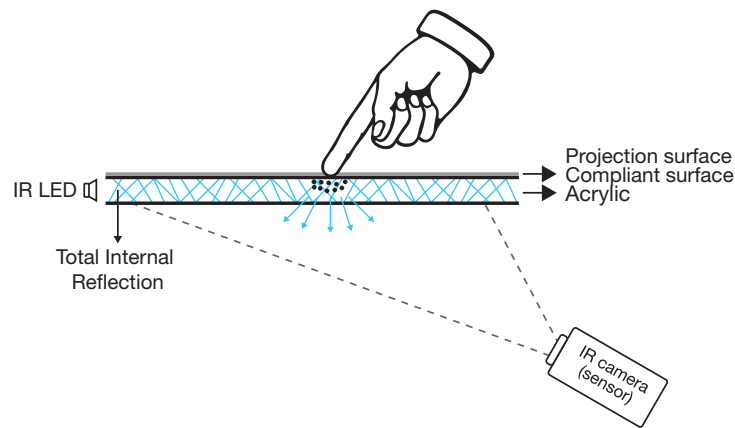
FIGURE A.1: FTIR – Frustrated Total Internal Reflection
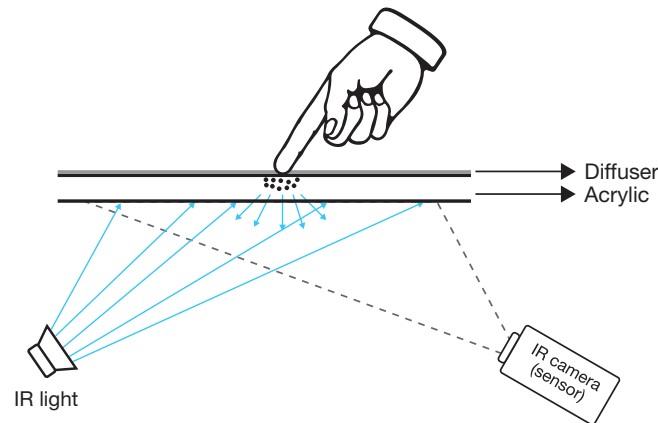
FIGURE A.2: Rear DI – Diffused Illumination with rear illumination

The DI technique,[3] refers to a multitouch technique that diffuses the light on the surface using a diffuser on top or at the bottom of an acrylic. The IR light is shined to the surface either from top (front DI) or below (rear DI). When a fingertip presses the surface, the fingertip reflects more light than the diffuser, an additional light that is sensed by the IR camera. Of the two, the rear DI can track objects as well. Figure A.2 shows how rear DI works.

---

[3] http://wiki.nuigroup.com/Diffused_Illumination (accessed 30 September 2014).

Similarly, Laser Light Plane Illumination (LLP)[4] is a multitouch technology that uses IR laser light positioned above the surface, which only tracks fingertips. An alternative to rear DI for both object tracking and multitouch is Diffused Surface Illumination (DSI),[5] which extends FTIR for object recognition by using a special acrylic that distributes the IR light evenly on the surface. With DSI, no compliant surface is needed. By contrast, a top webcam is an easy option to implement that manages object tracking, yet it lacks multitouch recognition or visual feedback. This option is suitable for early prototyping. Both the Reactable and the Woodentable use Rear DI. They can track objects. The Reactable has also multitouch input. Rear DI is also used by other commercial products such as Samsung SUR40 by Microsoft,[6] although it lacks object recognition. The DiamondTouch [Dietz & Leigh, 2001] is another multitouch tabletop based on capacitive touch technology instead, which in contrast with the previous tabletops, can recognise who is touching and where. Figure A.3 summarises in a matrix the different available multitouch and object tracking techniques for tabletops.

For more information on tabletop tangible technologies, a popular and recommended resource is the NUI Group website,[7] a worldwide leading community on multitouch technologies and techniques with many online resources, including tutorials on how to build your own tabletop systems, or an e-book on multitouch technologies [NUI Group Authors, 2009].

A.1.2 How rear diffused illumination works?

Figure 3.1 shows the basic design principles of a tabletop system implementing the rear DI technique. The DI technique refers to positioning an infrared

---

[4]`http://wiki.nuigroup.com/Laser_Light_Plane_Illumination_(LLP)` (accessed 30 September 2014).

[5]`http://wiki.nuigroup.com/Diffused_Surface_Illumination` (accessed 30 September 2014).

[6]`www.microsoft.com/en-us/pixelsense/default.aspx` (accessed 30 September 2014).

[7]`www.nuigroup.com` (accessed 30 September 2014).

OBJECT TRACKING

FIGURE A.3: Matrix of object tracking and multitouch tabletop technologies

(IR) camera and a projector below the tabletop for object tracking and projecting visual feedback respectively. A suitable material for the surface is required, transparent enough to allow the identification of objects, and opaque enough to permit a visible projection. As explained above, an acrylic combined with a diffuser is the most common solution. In addition, IR lighting is required because otherwise the IR camera will not be able to track the objects.

Depending on the quality of the material, IR lighting and camera used, it will affect considerably to the responsiveness of the system in terms of object tracking and visual feedback. Responsiveness is closely related to the cost of the items, particularly the IR camera, the IR lighting, and the surface material. For example, the low-cost camera used with the Woodentable (see Appendix A.2 for further details) has a responsiveness of 30 fps speed, which is lower than a camera for a commercial tabletop system that can reach 60 fps. However, 30 fps is an acceptable standard level for prototype testing.

A.2   Woodentable inventory

Table A.1 lists the commercial products used to build the Woodentable (2011–2012). All URL links were accessed on 30 September 2014.

TABLE A.1: List of products used to build the Woodentable

Concept	Product	Link	Cost
Short throw projector	Optoma EW605ST short throw projector (throw ratio = 0.52:1).	www.optoma.co.uk/ uploads/brochures/ EW605ST-B-en-GB.pdf	£ 576.00
High power IR LED	ACULED®VHL^TM IR ACL01-SC-IIII-005-C01-L-R000	www.perkinelmer.com/ CMSResources/Images/ 44-3465DTS_ACULEDVHL _IR.pdf	£ 12.00
Heatsink	Heatsink + thermal compound bought from the Maplin store (an electronics supply store in the UK).	-	£ 40.00
IR camera + lens	Sony PS3 Eye camera modified (with CS mount, IR filter and custom mountable housing) and a vari-focal CS lens (2.8 mm–10 mm).	http:// peauproductions.com/ store/index.php ?main_page=product _info&products_id=68	£ 160.00
Acrylic surface	Acrylic sheet 90 cm×180 cm (cut down to 87 cm × 87 cm) bought from B&Q (a home and garden supply store in the UK)	-	£ 30.00
Diffuser/projection surface	Lexan® 8B35-112 film 20 mils Velvet/Matte 60 cm × 90 cm. This material was already in the lab, it was bought from Sam Flax (an art supply store in Atlanta). The cost provided is approximate.	-	£ 30.00
Mac mini	Mac mini 2.5GHZ dual core Intel I5 - 4GB - 500GB serial ATA drive	-	£ 628.80
Total cost			**£** 1476.80

A.3   How to build a rear DI tabletop

An enclosed frame is required for optimal performance. As shown in Fig. A.4, the Woodentable's layers from top to bottom are:

- **Clear acrylic or glass surface** for supporting the pressure of tangible objects or fingertips.

- **Diffuser/projection surface** for stopping the image from the projector.
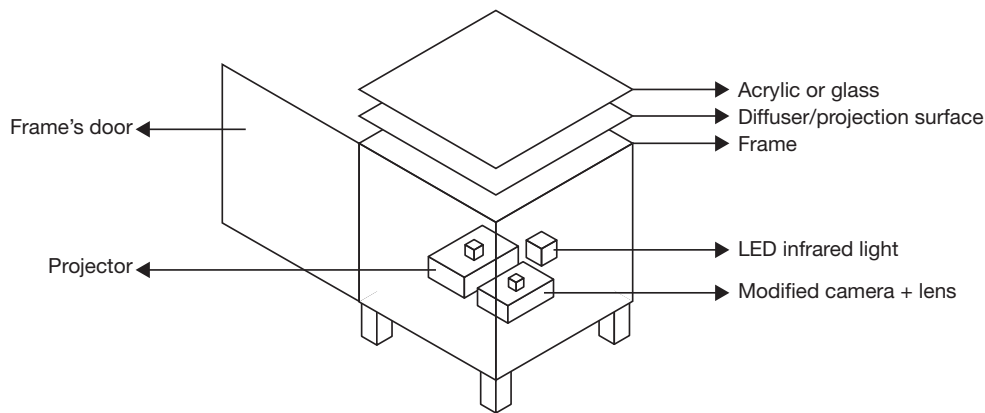
FIGURE A.4: Rear DI set-up

- **Infrared light** for illuminating the camera's view, positioned below the touch surface.

- **Modified camera + lens**, positioned below the touch surface.

- **Projector**, positioned below the touch surface.

For more information, a good explanation is provided by Peau Productions.[8]

A.3.1 How to position the camera

As documented in §A.2, we opted to use a camera with a vari-focal CS lens, so we could explore the position of the camera in relation to the area visible to the camera with less constraints when compared to a fixed-focal lens. The lens used had a focal length of $2.8\,\text{mm}$–$10\,\text{mm}$, a range of values that relates the focal distance and the *field of view* (FOV) or angle of view (area that a camera sees): the FOV and focal length are inversely proportional for a given distance. A focal length of $2.8\,\text{mm}$ provides a wide angle, whilst a larger focal length provides a FOV with a narrower angle. From the CS lens distance datasheet (see Table A.2, adapted from Peau Productions[9]), we see that we need to set the lens to $2.7\,\text{mm}$ at $60.96\,\text{cm}$ distance to the projection

---

[8]www.peauproductions.com/reardi.html (accessed 30 September 2014).
[9]http://peauproductions.com/store/index.php?main_page=product _info&products_id=321 (accessed 30 September 2014).

TABLE A.2: CS lens distances

Focal length (mm)	Field of view (°)	Distance = 60.96 cm		Distance = 91.44 cm		Distance = 182.88 cm	
		Horizontal (cm)	Vertical (cm)	Horizontal (cm)	Vertical (cm)	Horizontal (cm)	Vertical (cm)
2.7 mm	93	82.30	60.96	121.92	91.44	243.84	182.88
3.5 mm	80	64.01	45.72	94.49	70.10	188.98	140.21
8 mm	72	27.43	21.34	42.67	30.48	82.30	60.96
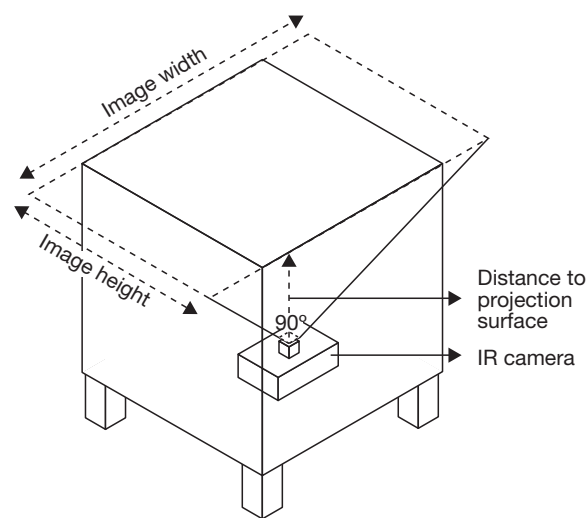
FIGURE A.5: Camera position

surface in order to obtain an image of 82 cm × 60.96 cm. Figure A.5 illustrates how it works. For more information, a useful resource is provided by Peau Productions.[10]

A.3.2 How to position the projector

A short throw projector provides a short *throw distance* (the distance between the projector lens and the projection surface). A *short throw* projector produces larger images at a close distance compared to a normal projector. This allows placing the projector opposite side of the touch surface and point directly to the surface (Fig. A.6, right), whereas a normal projector requires a mirror to

---

[10]www.peauproductions.com/cameras.html (accessed 30 September 2014).
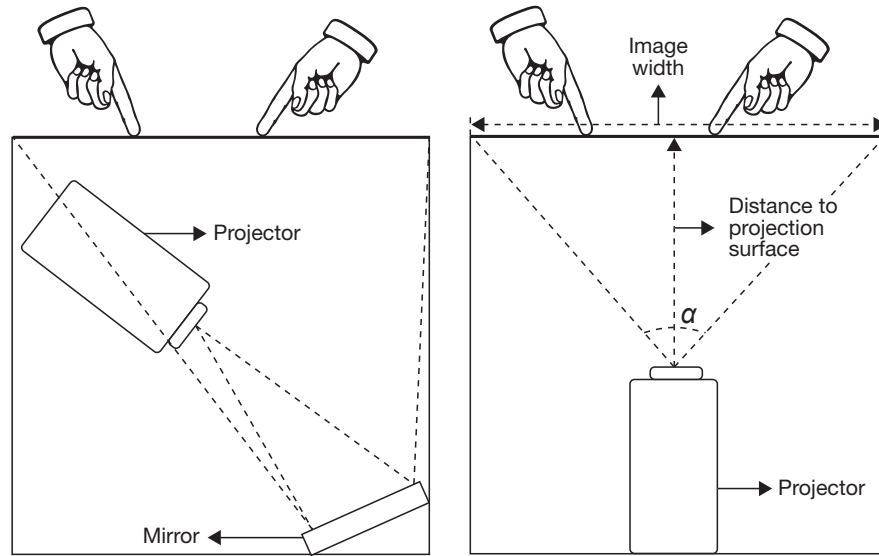
FIGURE A.6: Left: Reduction of the throw distance using a mirror. Right: Direct throw distance

reduce the throw distance (Fig. A.6, left). The latter option is more affordable, yet it will require finding the right angle of the mirror, and calibrating the projected image as well. We opted for the former to simplify the set-up.

The *throw ratio* (TR) of the projector refers to the ratio of the distance to the projection surface (D) to the projected image width (W): $TR = \frac{D}{W}$. The TR measure is helpful to know the image width thrown at a particular distance, and, conversely, the distance needed to project the image with a particular width.

For the Woodentable, we had in mind to build a square table of approximately 70 cm height (100 cm height in total including its legs) with a square surface of 60 cm × 60 cm. The short throw projector used had a throw ratio of 0.52:1. We selected an aspect ratio of 4:3.[11] Therefore, we needed an image width of 80 cm, and the projector had to be placed at a distance of 41.6 cm. There exist useful tools, such as projection calculators, for calculating the image size in relation to the throw distance depending on the type of projector.[12]

---
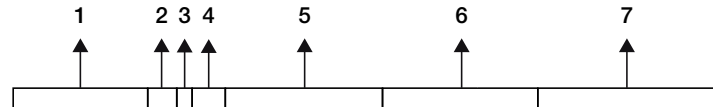
[11] `www.projectorcentral.com/Optoma-EW605ST.htm` (accessed 30 September 2014).
[12] `www.projectorcentral.com/Optoma-EW605ST-projection-calculator-pro`
`.htm` (accessed 30 September 2014).

For more information, a recommended explanation is provided by Peau Pro-

ductions.[13]

---

[13]`www.peauproductions.com/proj_lcds.html` (accessed 30 September 2014).

A.4   Example of interaction log file

```
      1         2 3 4      5           6           7
      ↑         ↑ ↑ ↑      ↑           ↑           ↑

3185.003901808,set,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.039428751,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.072332145,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.106854861,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.139551097,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.171525316,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.20576496,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.237945813,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.273953476,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.305801818,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.337662371,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.371799006,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.405050204,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.437630904,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.473914331,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.504991804,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.540936604,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.572793481,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.603925246,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.639126511,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.670476906,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.705829569,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.737112982,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.772973884,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.804460593,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.837899876,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.869354962,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.902552396,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.937898191,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3185.970971117,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.006949967,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.040641981,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.073230737,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.102583776,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.137310965,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.171759236,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.206618688,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.239035628,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.270449653,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.304696125,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.338779794,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.3714947,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.404684938,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
3186.438371396,update,1,366,0.49283760786057,0.10404283553362,3.9054114818573
```

1 Timestamp (time tag in seconds, float32)

2 Messages (range set/update/del)

3 Class ID (fiducial ID number, int32)

4 Session ID (temporary object ID, int32)

5 Position-x (x-axis position, range 0...1, float32)

6 Position-y (y-axis position, range 0...1, float32)

7 Rotation-x (euler angle alpha, range 0..2PI, float32)

A.5 Example of consent form

A.5.1 Informed consent form for use in a lab

Consent Form

[Title of the study]

If you are willing to take part in this research project please tick the box, complete the details below and return the signed form.

At any time during the research you are free to withdraw and to request the destruction of any data that have been gathered from you, up to the point at which data are aggregated for analysis.

The results of any research project involving [University Name] staff constitute personal data under the Data Protection Act. They will be kept secure, confidential and not released to any third party.

☐ I am willing to take part in this research, I consent to being video recorded, and I give my permission for the data collected to be used in an anonymous form in any written reports, presentations and published papers relating to this study. My written consent will be sought separately before any identifiable data are used in such dissemination.

Signing this form indicates that you understand the purpose of the research, as explained in the covering letter, and accept the conditions for handling the data you provide.

There are no risks associated with this experiment. This project has the approval of the [University's Human Research Ethics Committee] and it is in compliance with the Data Protection Act.

If you have any questions you may contact:

• [Name and Surname] (PhD student), [email address], [phone number]
• [Name and Surname] (Principal Supervisor) [email address], [phone number]

Name:..

Signed: ....

Date:...

Please return completed form to: [email address PhD student]

A.5.2 Public sign for use in a museum

[Title of the study]
([City], [Country])

[Name and Surname][1] [Name and Surname][2] [Name and Surname][3]
1 Affiliation 2 Affiliation 3 Affiliation

ABSTRACT

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.

STUDY TIME/DATE

[Day] DD/MM/YY
Morning: HHam-HHam
Afternoon: HHpm-HHpm

[Day] DD/MM/YY
Morning: HHam-HHam
Afternoon: HHpm-HHpm

[Photo]

© Photo by [Name and Surname]

VIDEO RECORDINGS

In order to undertake this study, we are currently video- and audio-recording in a corner of the lower exhibition area. If you have any concerns about being recorded, please inform our researcher or a member of staff and the camera will be switched off immediately. If you have been recorded but decide that you would prefer that the recording be destroyed, again, please inform us and the footage will be deleted. The material will be used for research purposes only.

CONTACT

[Photo]

The PhD student [Name and Surname] is going to be available during the intervals of filming for questions and/or reservations. For further enquiries, please contact Dr. [Name and Surname] ([email address]).

IN COLLABORATION WITH   [Logo]   [Logo]

[Logo]

A.6 Student's t-Test

The paired Student's t-Test for statistical means comparison between conditions is a suitable statistical test for comparing two samples of small size ($n < 30$) [Crawley, 2005, pp. 67] of the same length, in which each value in one sample can be sensibly paired with a measurement in the other sample. For example, this test is suitable if measures are taken from the same group twice as it is the case with a within subjects design [Crawley, 2005, pp. 81–83]. The *t-value* indicates how different are two samples from the same population (the larger the value, the larger the difference), and the *p-value* indicates the probability of this difference, that is, whether we can accept (a value above the significant level) or reject (a value below the significant level) the null hypothesis that the means of two samples are equal.

The t-Test was used in Study 3 in order to compare space use between conditions. Results are reported at a significant level of 0.05 (that is from the value we can reject the null hypothesis that the two means are equal). In this study, the degrees of freedom or number of values in the dataset is 7 (number of groups minus 1). The less data you have, the larger your sampling error is likely to be. We complemented statistical analysis with detailed qualitative analysis of the videos for consistency.

Appendix B

Supplementary Videos

Table B.1 shows links to videos discussed in Chapter 4 and Chapter 6 of this thesis. Please use the following password to access to those that are protected videos: *thesis2014ax*. All URL links were accessed on 30 September 2014. A DVD copy is also available at The Open University Library.

TABLE B.1: Code, description of and links to videos discussed in Studies 1 and 3

Study	Code (Ch.-#)	Description	Link
	4-1A	Dynamic reconfiguration of individual territory (Fig. 4.9)	https://vimeo.com/72203848
	4-2A	Complex invasion with the radar trigger object (Fig. 4.10)	https://vimeo.com/72286054
	4-3A	Triggering unintentional effects with dynamic patching (Fig. 4.7)	https://vimeo.com/72286223
	4-4A	Handover (Fig. 4.12)	https://vimeo.com/72286313
	4-5A	Bobbing heads (Fig. 4.13)	https://vimeo.com/72286386
	4-6A		https://vimeo.com/72286510
	4-6B		https://vimeo.com/72287743
	4-6C		https://vimeo.com/72288309
Study 1	4-6D	Basic vs. complex dialogues (Table 4.3)	https://vimeo.com/72288524
	4-6E		https://vimeo.com/72288588
	4-6F		https://vimeo.com/72288662
	4-6G		https://vimeo.com/72288752
	4-6H		https://vimeo.com/72288952
	4-7A	Basic vs. complex intro (Fig. 4.15)	https://vimeo.com/72289103
	4-7B		https://vimeo.com/72289380
	4-8A	Basic vs. complex ending (Fig. 4.16)	https://vimeo.com/72289617
	4-8B		https://vimeo.com/72289829
	4-9A		https://vimeo.com/72289829
	4-9B	Mimicking (Fig. 4.18)	https://vimeo.com/72290004
	4-9C		https://vimeo.com/72290106
	4-10A		https://vimeo.com/72290637
	4-10B	Explicit peer learning (Fig. 4.19)	https://vimeo.com/72290742
	4-10C		https://vimeo.com/72290897
Study 3	6-1A	Video demo of SoundXY4	https://vimeo.com/70693984

Appendix C

Study 1 Materials

C.1   Themes on verbal and non-verbal communication

Figure C.1 shows the emergent themes in a two by two matrix of verbal and non-verbal behaviours vs. individual and group units. Table C.1 describes these themes, which appear in alphabetical order.

C.2   Territories and thread ownership categories

Table C.2 describes the identified territorial interaction categories.

FIGURE C.1: Overview of themes on verbal and non-verbal communication. Themes straddling the centre line relate to both individual and group behaviour

TABLE C.1: Description of verbal and non-verbal communication themes

Theme	Description
Comments/dialogues	Verbal conversation between at least two individuals.
Coordination	Non-verbal joint action in group (e.g. handovers).
Dynamic patching	Those interaction events related to the automatic connections present in the Reactable interface.
Error/repair situations	A person starts an unintended effect (e.g. a mistake) and then potentially tries to fix it.
Explicit peer learning	A group exchanges verbal information to solve a problem in collaboration (e.g. solving an error/repair situation in team).
Explorations	Dragging, rearranging, strobing objects to ones' own and/or others' threads, with both musical and interactional impact.
Gestures	Hands and body actions (e.g. sound producing gestures, ancillary gestures, communicative gestures, sound accompanying gestures).
Intros/endings	A group starts or finishes the session.
Isolationism	An individual interacts alone with the Reactable.
Mimicking (implicit peer learning)	At least one person imitates another person's interactions with both musical and interactional impact.
Musical dialogues	Conversation between at least two leading melodic and/or rhythmic voices operated from at an individual thread each.
Non-participation	Member of a group stops actively participating for a while (i.e. standing back).
Solos	Leading melodic and/or rhythmic voice operated from an individual thread.
Special objects	Use of global and special objects such as a radar trigger (a local tempo controller with local/global effects on all objects in its range) or programmer (used to re-program the instruments with samples).
Configurations	Basic configurations (e.g. one sound generator only) vs. complex configurations (e.g. one filter at the end of a thread) with objects.
Techniques	Basic interaction styles (e.g. dragging) vs. complex interaction styles (e.g., dragging and swapping) with the objects.
Territories & threads ownership	Events related to the Reactable's distribution of territories and ownership of audio threads.
Thinking alouds	A person exposes his/her thoughts aloud (e.g. commenting to him/herself about the effect of an action with an object).
Transition/changes	The individual or collective process of moving from one motif to another motif (e.g. the process of moving from one individual leading melody to another or the group process of moving from one consistent set of leading melodies and/or rhythms to another consistent set).
User manual reading	A person reads the Reactable manual with no explicit exchange of verbal communication.

TABLE C.2: Description of territorial interaction categories and sub-types

Category	Description	Sub-type
Thread owner-ship	A thread can 'belong' to an individual, or be shared. A shared thread is built in collaboration with at least one other person. A thread can start out as shared, or can become shared over time. Similarly, it can stop being shared, either becoming the property of one of the collaborators, or disappearing. The essential characteristic of a shared thread is a dialogue of actions (moves or turns with interface objects) between two or more people: that is, if A and B are two people, actions on a shared thread would be carried out in the order A-B-A or B-A-B. We defined a threshold of 5 actions with no response from other participants in order for a thread to become personal again.	*Individual*: Thread built by a single person. *Shared*: Thread built in collaboration.
Invasions	'Interfering' in somebody else's thread via an action (e.g., adding an object, 'strobing on and off' i.e. adding/removing an object repeatedly (rhythmically or non-rhythmically), dragging an object around, swapping it for another object). If one object is just removed as a single action from an active thread of someone else, with no previous or post sequence of actions, then it is counted as a take. Invasions can be carried out with any of the tangible objects that have local effects, such as sound generators, sound effects or control generators. Global objects are thus not counted in this case (except for the special object radar trigger, which is a local tempo controller with local/global effects to all objects in its range).	*Rejected*: The thread owner undoes the contribution immediately (within 5 seconds). *Accepted*: The thread owner does not undo the contribution immediately (within 5 seconds).
Takes	Taking an object that belongs to somebody else for individual use. 'Belonging' could mean either that the object is currently used in a personal thread (active take from the surface), or it is positioned in the rim area close to another person, who is not using it, and requires an extensive reach movement to get the object, intruding into the other person's personal territory (passive take). The object is taken without asking for permission and without the intention of being returned (in contrast with asking somebody to pass a personal storage item, cf. Scott et al. [2004]).	*Active*: Taken from surface table area. *Passive*: Taken from rim area.
Gives	Handing an object to somebody else for individual use. The object is given without asking which is either left on the surface as an active give (e.g., handover) or in the rim area as a passive give. If the object is added to an active thread of somebody else, then it is counted as invasion.	*Active*: Given from surface table area. *Passive*: Given from Rim area.

Appendix D

Study 3 Materials

D.1 Sound credits of SoundXY4

This section includes a selection of sounds from Freesound.org used in Study 3. The selection was made by Gerard Roma. Sounds are grouped in six categories inspired by Russolo's taxonomy of sounds based on the landscape of the urban city: explosions (D.1.1), percussion (D.1.2), screeches (D.1.3), voices (D.1.4), whispers (D.1.5), and whistles (D.1.6). All URL links were accessed on 20 September 2013.

D.1.1 Explosions

- Splash.wav by Kayyy:
  http://freesound.org/people/Kayyy/sounds/61015
- SUBSONIC RUMBLE.wav by sandyrb:
  http://freesound.org/people/sandyrb/sounds/84347
- Large Crash.wav by CGEffex:
  http://freesound.org/people/CGEffex/sounds/99960
- explosion4.wav by sarge4267:
  http://freesound.org/people/sarge4267/sounds/102734
- Thunder_close_Boem.wav by hantorio:
  http://freesound.org/people/hantorio/sounds/121946
- Gun Shot sound_02 by GregsMedia:
  http://freesound.org/people/GregsMedia/sounds/150139

D.1.2 Percussion

- prac - tom.wav by TicTacShutUp:

  `http://freesound.org/people/TicTacShutUp/sounds/449`

- rocks.aif by splashzooka:

  `http://freesound.org/people/splashzooka/sounds/21789`

- BatonRes.wav by garogourou:

  `http://freesound.org/people/garogourou/sounds/144522`

- Teller 01.wav by Peter Lustig:

  `http://www.freesound.org/people/Peter%20Lustig/sounds/150098`

- falling metal 1 - 20.3.11.wav by toiletrolltube:

  `http://freesound.org/people/toiletrolltube/sounds/179854`

- Cymbal Impact.wav by LloydEvans09:

  `http://freesound.org/people/LloydEvans09/sounds/185818`

D.1.3 Screeches

- spoon_bowl2.wav by Corsica_S:

  `http://freesound.org/people/Corsica_S/sounds/64383`

- 00736 rustling plastic 1.wav by Robinhood76:

  `http://freesound.org/people/Robinhood76/sounds/67267`

- Neon Light.wav by Julien Matthey:

  `http://freesound.org/people/Julien%20Matthey/sounds/118340`

- Monster Screech.wav by thegoose09:

  `http://freesound.org/people/thegoose09/sounds/125388`

- train screech.wav by ognito perceptu:

  `http://freesound.org/people/cognito%20perceptu/sounds/181868`

- fire.crackling.mp3 by dobroide:

  `http://freesound.org/people/dobroide/sounds/4211`

D.1.4 Voices

- tibetan chant 1.wav by djgriffin:

  `http://freesound.org/people/djgriffin/sounds/15362`

- insane laughter man reverb.wav by Leady:

  `http://freesound.org/people/Leady/sounds/26729`

- psycho scream 1.wav by FreqMan:

  `http://freesound.org/people/FreqMan/sounds/42847`

- 01948 kid laughter.wav by Robinhood76:

  `http://freesound.org/people/Robinhood76/sounds/103592`

- catHowling2.wav by Zabuhailo:

  `http://freesound.org/people/Zabuhailo/sounds/146965`

- Groan_male_normalised.wav by Adam_N:

  `http://freesound.org/people/Adam_N/sounds/171758`

D.1.5   Whispers

- wind_gurgle.ogg by Halleck:

  `http://freesound.org/people/Halleck/sounds/2878`

- CatMathilda01_-18dBFS.wav by DanGasior:

  `http://freesound.org/people/DanGasior/sounds/90742`

- Stomach_Gurgling - SeveralSamples by Vosvoy:

  `http://freesound.org/people/Vosvoy/sounds/149120`

- whisper.wav by Erdie:

  `http://freesound.org/people/Erdie/sounds/165617`

- audience becomes still 01.wav by klankbeeld:

  `http://freesound.org/people/klankbeeld/sounds/171550`

- Whisper of Wind by Hephaestus:

  `http://freesound.org/people/Hephaestus/sounds/181183`

D.1.6   Whistles

- 00773 leaking gas 1.wav by Robinhood76:

  `http://freesound.org/people/Robinhood76/sounds/66248`

- whistle of boat.aif by nextmaking:

  `http://freesound.org/people/nextmaking/sounds/86045`

- short whistle it tunnel.aiff by SoundCollectah:

  `http://freesound.org/people/SoundCollectah/sounds/109354`

- pipe hiss 003.WAV by DJ Chronos:

  `http://freesound.org/people/DJ%20Chronos/sounds/130289`

- Air_escaping.wav by Adam_N:

  `http://freesound.org/people/Adam_N/sounds/164623`

- Snort 2 long.flac by bigfriendlyjiant:

  `http://freesound.org/people/bigfriendlyjiant/sounds/183915`

D.2 Questionnaire in SoundXY4

Questionnaire SoundXY4

Supporting real-time auditory feedback in musical tabletops

Please, fill in the form by answering the questions. This might help us understanding your interaction with the tangible user interface. Many thanks for participating in the experiment!

Name:

Surname:

Gender: F / M

Age:

Country of origin:

Musical background

How many years do you have of musical training?
None
1-2
2-4
4-6
More than 6

Additional comments about your musical training (e.g., do you play any instrument?)

Are you familiar with electronic music? (e.g. listening, playing…)
None
Occasionally
Regular basis

Have you used multitouch technology before?

Smartphones
Never
Once or twice
Regular basis

Tablets
Never
Once or twice
Regular basis

Interactive tabletops
Never
Once or twice
Regular basis

Have you used tangible user interfaces (TUI) before?
Never
Once or twice
Regular basis

Are you...?
Left-handed
Right-handed
Ambidextrous

Collaboration

Did you know the other members of the group?

How did you find improvising together?
Easy
Regular
Difficult

Why?

Any additional comments about the experience?

D.3   Description of the effects in SoundXY4

Table D.1 details the effects used in SoundXY4, including the modifiable parameter and its range of values.

TABLE D.1: Description of the effects used in SoundXY4. For more information see SuperCollider Help

UnitGenerator	Description	Parameter	Range
In*SinOsc	Amplitude modulation	Frequency	1–2
BFP	Band pass filter	Centre frequency (in Hertz)	100–10100
CombC	CombC delay (with cubic interpolation)	Decay time	(-1)–(-0.5)*
HPF	High pass pilter	Cutoff frequency	440–7440
RLPF	Resonant low pass filter	Cutoff frequency	20–460
PitchShift	A time domain granular pitch shifter	Ratio of the pitch shift	0–4

*negative values emphasise odd harmonics at an octave lower

D.4   How the implementation of ambisonics works

Ambisonics can be implemented in the horizontal plane (2D, minimum four speakers), and also in the vertical plane (3D, minimum six speakers). The ambisonics technique is founded on theories of psychoacoustic perception, centred on how the listener perceives sound when decoding an audio signal. As explained by Gerzon [1974a], decoding an audio signal is based on perception principles such as differences between localisation of low, mid-high, and above 5kHz frequencies; and the use of reverberation as a localisation aid.

D.5   Used colours by category

Table D.2 shows the colour used for each category.

TABLE D.2: Used colours by category in SoundXY4

Category	Colour
Explosions	Red
Percussion	Yellow
Screeches	Orange
Voices	Magenta
Whispers	Grey
Whistles	Green

D.6 Physical interface design

We here briefly describe the physical interface design of SoundXY4. We mapped both sound players and filters to cubes of 5 cm size each side, which is the default size of the fiducial markers that are available with the reacTIVision software, and is a standard size for hand-reach objects. The material used for the tangible cubes was translucent plastic, which was flexible enough to draw the layout of a cube, cut it, fold it and mount it with glue (see Fig. D.1). Then we stuck a fiducial marker, printed on non-transparent adhesive, to each side: black over white for sound players, and white over black for filters.

FIGURE D.1: Layout and assemblage of the SoundXY4 tangible cubes

D.7 Themes on workspace awareness

Table D.3 describes the emerging themes on workspace awareness, non-mutually exclusive, related to analysing verbal communication from video data.

TABLE D.3: Description of themes on workspace awareness

Theme	Description
Identification of categories [1]	When a particular group of sounds of a cube is explicitly identified using general concepts (e.g. [We hear a sound of a ceramic impact, a tom drum, a falling metal, metallic impact] "This is all percussion sounds") or characteristic attributes (e.g. "Mine ['screeches' white cube/sound] is really nasty").

Impressions of sounds [2]	The first reactions of identifying a sound. We include here clear expressions (e.g. laughing) or exclamations (e.g. onomatopoeias) as a first reaction to a sound, by at least one team member. These reactions indicate implicit workspace awareness. Feelings of like or dislike are also included here as it shows implicit workspace awareness as well (e.g. "I don't like that"). We also include here attributes that characterise the sound output (e.g. "That's a horror film now") but not necessarily refer to the sound source. Also included here are generic concepts of the sound source, which lack the detail of the specific sound source (e.g. "[We hear a sound of a water splash, a scream, and a tunnel short whistle] It's noise all that").
Identification of source [3]	This theme includes identifying the sound source using general concepts or characteristic attributes. We include here when a particular sound source is explicitly identified using terms related to the sound source (e.g. "[We hear a sound of a cat howling] A kitty! a kitty!") or the sound source is described ("[We hear a sound of a tunnel short whistle] (...) a squeaking metal") or characterised (e.g. "[We hear a sound of a man laugh] You've got the evil laugh"). If a quote refers to a sound source and a feeling, then it will appear both in *Impressions sounds* and *Identification source* (e.g. "Silence, oh that's fantastic"). If there is a clear mismatch between the identified sound source and the actual sound source the quote will be excluded because even though trying to define the sound indicates a workspace awareness of the sound, we consider that a certain precision in understanding the sound is important when trying to define it. For example, if the sound is a purring cat and someone says "Oh that's the... crickets!", it is classified as non-identification of sound. We may also find examples of both identifying a sound and another theme, for example "The fire sounds really nice, it sounds as it is actually fire" exemplifies both *Identification of sounds* and *Stories, realism and ambisonics*.
Non-identification of person or object [4]	An explicit question or comment about who or what produces a sound, typically starting with "Who?", "What?", or "Which" (e.g. "Oh who's got that one?").
Non-identification of source [5]	An explicit question or comment about the sound source (e.g. "Yeah I can't hear this wind at all [while lifting up a white cube]"), or when a sound source is clearly misunderstood (e.g. "[We hear a sound of a purring cat] Oh that's the... crickets!").
Identification of filters [6]	A particular effect or the collection of effects available in a black cube is noticed (e.g. "There is a difference!"), identified (e.g. "There's an echo"), or described (e.g. "Uh that's very spooky"). If there is a related conversation and some of the team members perceive the effects and others do not perceive the effects, the quote would appear in both *Identification of filters* and *Non-identification of filters*. For example, here P2 and P4 are understanding the effect of a black cube, whilst P3 is not: P2 "Yours black one is not going to make any noises (...) you have to connect it to a white one to make noise" P3 "what sort of noise?" P2 "I think is the same noise as the white cube, which changes the noise saturance" P4 "It's like a transition".
Non-identification of filters [7]	An explicit question or comment about a particular effect or the collection of effects available in a black cube because of a lack of understanding of how it works (e.g. "What these modifiers are supposed to do? what they actually do?").

Stories, realism and ambisonics [8]	A connection made 1) between sounds or associations; 2) to storytelling; 3) to musical composition; 4) to realism; or 5) to spatialisation in terms of associations with the position of sounds and the position of speakers (ambisonics), or associations with no-spatialisation. These subtypes are not mutually exclusive.

1. Sounds associations [8.1]: Comments that connect or associate the sounds to real or imaginary characters (e.g. "ghosts", "Shrek"), or sound sources (e.g. "evil laugh", "cat panther") or situations (e.g. "sounds like a teapot", "sounds like a supermarket") or abstract concepts (e.g. "Whispering, underground noises"), beyond the sound source.

2. Storytelling [8.2]: Comments that narrate or build stories connected to the sounds beyond a mere description and qualification of the used sounds (e.g. "the snoring man with a purring cat by the fire"), including connections to memories or lived situations (e.g. "do you remember this song called popcorn?").

3. Musical composition [8.3]: Comments about music-related stories, in particular references to musical roles (e.g. "I am gonna give you guys this rhythm [while lifting up and down rhythmically a sound of a wood impact] (...) See if you can have some sounds on top of that rhythm"), musical style (e.g. "we've arrived, we are doing Stockhausen now, we've arrived, that's definitely Stockhausen, take it back down, stepping in, boom"), or musical structures (e.g. "we can keep the ghosts but we can decrease the volume").

4. Realism [8.4]: Comments that characterise the sounds, effects or the overall sound output as realistic (e.g. "A bit more realistic, it is a bit more realistic the snoring...") or as a scene (e.g. "We created a scene that's all we wanted to do, wasn't it?")

5. Ambisonics [8.5]: Comments that associate the speakers with the location of the objects on the surface (e.g. "I wonder whether it is interesting where we put these [tangible objects on the table] cause we wanted the fire in the corner if it had some significance, we created a geographical representation as well as the sound one"). Or when there are comments about the lack of spatialisation (e.g. "So the sound comes mainly from this speaker, yeah? [pointing to his back right speaker]").

Immersive musical experience [9]	Explicit comments about enjoying the activity by characterising particular actions (e.g. "This is nice, it's a different way" [after pressing on and off a cube side with the sound of a tunnel short whistle at different parts of the table rhythmically]) or the overall sound output (e.g. "I like what we've got, it's good!").

D.8 Transcripts by themes

Transcripts are shown here as quotes classified by themes. Usually the quotes only appear once, but in some cases they belong to different themes. For example, when there is awareness of a sound or a set of sounds, and it is characterised as realistic, it can be found in *Identification of sounds* and *Stories, realism and ambisonics* as in "The fire sounds really nice, it sounds as it is actually

fire". Or if there is awareness of a sound, and it is characterised as a concept that matches a white cube, it can be found in *Identification of sounds* and *Identification of categories* as in "You've got the people noises".

The quotes which are directly topic-related and subsequent in time are grouped as one quote. For example, '[We hear the sound of a kid laugh] P2 "That's a happy baby" P3 "happy baby" P2 "Shall we find a nice sound to go with the happy baby?"'. If the quote has different topics it is split into these themes, if possible. For example, 'P2 "Is that the gong? P1 "That's the gong"' is split into "Is that the gong?" grouped in *Non-identification of source*, and "That's the gong" grouped in *Impressions of sounds*. If the quote cannot be split without losing its meaning, then it will be repeated. For example, 'P2 "Yours black one is not going to make any noises (...) you have to connect it to a white one to make noise" P3 "What sort of noise?" P2 "I think is the same noise as the white cube, which changes the noise saturance" P4 "It's like a transition"' will be classified as *Identification of filters* and *Non-identification of filters*.

D.8.1 Identification of categories

TABLE D.4: Transcripts of *Identification of categories*

Group	Condition	Quotes
G1	NSP: 0	–
	SP: 4	12:16 P4 "Sounds like humans [inaudible] animals (...) This is like loud noises, and explosions and [inaudible]" P2 "I wonder this one" P4 "That's in animal objects" P3 "The pink one is animals and people?" ● 12:30 P4 "So red is... hum..." P2 "gat sounds" P4 "gat sounds, and gunshots, and stuff" P3 "That's red, danger, warning" ● *17:09 P3 "Explosions" [in response to P1's question "What is that one [white cube/sound]?" after hearing a sound of a war boom explosion] [1,3]* ● 17:26 P3 "the explosions" P2 "that one" ●
G2	NSP: 1	*21:14 P2 laughs and says "You've got the people noises" [1,3]* ●
	SP: 0	–
G3	NSP: 0	–
	SP: 0	–

G4	NSP: 11	*06:29 [We hear a sound of a metallic whine, and a sound of a metallic-ceramic friction] P2 "Mine ['screeches' white cube/sound] is really nasty" [1,2]* ● 07:07 P2 "Does this baby have other things of voices?" [looking at P4] [1,3] ● 08:04 P2 [We hear no sound from the 'screeches' white cube] "Mine ['screeches' white cube/-sound] is not doing anything anymore I've got a really... [now we hear a sound of a lamp buzz]" [1,2] ● 08:28 P2 "Mine ['screeches' white cube/sound] seems to just disconnect" [1,2] ● 09:05 [We hear a sound of a tom drum] P1 "No, I, mine ['percussion' white cube/sound] is the strong [and demonstrates by lifting up and down the white cube] that one" P3 "Ok" P2 "Uh" [in response to P2's question "Who's got this thing? This voice? This lament?" [referring to a sound of a groan] [1,2] ● 09:28 [We hear a sound of plastic rustling] P1 "So your block makes all kind of everyday noises [pointing to P2's 'screeches' white cube], rustling paper and all kinds of these" [1,5] ● 09:35 P2 "Yes, mine [confirming P1's comment about P2's white cube on everyday noises e.g. rustling paper, referring to the 'screeches' white cube] but I don't think the voice [referring to a sound of a groan] comes from my block" [1,3] ● 10:24 P3 "This is the exploding block" ● 18:38 P3 "This one has percussive sounds" P1 "So it's just mine too are percussive actually" P3 "I am talking about explosions" ● 18:54 P2 "Mine is very inconsistent" ● 19:04 P3 "It's got shots" ●
	SP: 0	–
G5	NSP: 1	27:26 P4 "Yeah, and they are all here" [talking about a white cube that has a sound of a tibetan chant, in response to P4's comment that "That drone is the winner] P4 turns a white cube and it changes the sound from a tibetan chant to a cat howling P3 "Not only [tibetan chants]!" [1,2] ●
	SP: 0	–
G6	NSP: 0	–
	SP: 0	–
G7	NSP: 15	26:45 P2 "So we discover all the sounds of this one ['whistles' white cube] (...) The interesting one is snoring" [1,2,3] ● 28:45 P4 [We hear a sound of a ceramic impact] "Not very exciting that cube isn't" ['percussion' white cube] [1] ● 29:19 P2 [We hear a sound of a ceramic impact] "We know everything here ['percussion' white cube]" P4 "It is very exciting" P4 "This is still scenario, or like still things" ['percussion' white cube] P2 "Yes" [1,2] ● 29:32 P2 [We hear a sound of a long snort] "Let's keep this one ['whistles' white cube], that one is not interesting ['percussion' white cube]" [1] ● 29:39 [We hear a sound of a war boom explosion] P2 "Dangerous" P4 "It's the boring one" P4 "Yes this is the boring one [pointing to 'percussion' white cube] and this is the bad one [pointing to 'explosions' white cube]" P4 laughs [1,2] ● 30:17 [We hear a sound of water splash, a bomb boom explosion, a boom rumble] P2 "Okay I think this is the creepy one ['explosions' white cube]" [laughs] P2 "Maybe put it [next] to the boring one [in the rim area]" ● 30:23 P4 "Snoring, boring, creeping [recap of the identified white cubes]" P2 "Yes" P4 "I like some blurbs" ● 31:41 [We hear a sound of a man laugh] P2 "Maybe you can keep this one as the best one as the child side" ['voices' white cube] 31:44 P4 "The cat, scream" P2 "Exactly" ● 31:48 [We hear a sound of a wind mouth gurgle] P2 "This is another one ['whispers' white cube/sound]" [1,2] ● 33:05 [We hear a sound of a cat purring with an effect] P2 "Ok we've discovered this one, another boring category" [laughs] ['whispers' white cube] ● 33:51 [We hear a sound of a metallic whine, a lamp buzz] P3 "This is the sci-fi one" [the 'screeches' white cube] P4 "Yeah" P2 "Yes" P4 "Synthetic, yeah maybe" ● 33:59 P2 "What sounds have we got?" P4 "We've got snoring boring creepy cats another like boring" P2 "The cat one is good" P3 "Cats and babies" 34:13 P3 "And the... " P4 "Oh yeah, and the exciting one" [1,2,8] ● 34:45 P4 "Okay, baby, cat, laughing guy" [while listening to the 'voices' white cube] [1,3] ● 36:22 P4 "This one is the creepy one, [we hear a sound of a gun shot]" 36:27 P2 "Kill the baby" 36:30 P4 "And it's so wrong" [1,2,8] ● 39:08 P2 Yeah this one is red [pointing to the explosions white cube while P4 is trying whether it is working or not], probably, maybe we can..." [then we hear a sound of a water splash from the explosions white cube, and P4 leaves it] [1,2] ●
	SP: 2	11:05 [We hear a sound of a car crash] P4 "Is that the red?" P2 "I think it's the red [inaudible] the gat sounds" [1,2] ● 14:46 [We hear a sound of a gun shot from the red white cube] P2 "This one [white cube/sound] is negative [pointing to the red white cube]" P4 "The red" P2 "Yes" [1,2] ●
G8	NSP: 3	08:22 [We hear a sound of a kid laugh] P2 Okay so quite a lot of animal noises on that one [sound of a cat howling, a groan, and kid laugh] P1 "Animal noises the baby crying? [laughs]" ● 08:24 [We hear a sound of a ceramic impact, a tom drum, a falling metal, a metallic impact] P3 "This is all percussion sounds" ● 09:43 [We hear a sound of a man laugh] P2 "Ok so that's the cats, and animals, and things on" P1 "Ok" ●

| | SP: 1 | 28:04 [We hear a sound of a scream] P2 "I like that one" P1 "This is the funny box [laughs]" P3 laughs [1,2] ● |

D.8.2 Impressions of sounds

TABLE D.5: Transcripts of *Impressions of sounds*

Group	Condition	Quotes
G1	NSP: 8	24:25 [We hear a sound of a scream] All laugh ● 24:53 [We hear a sound of a groan] P1 "That's quite dodgy" ● 25:49 [We hear a sound of a groan] P1 "Some crawl" ● 28:34 P1 [We hear a sound of a water splash, a scream, and a tunnel short whistle] "It's noise all that" ● 29:25 [We hear a sound of a water splash played lifting it up and down rhythmically] P2 "That's paranoing" P1 "Uh that's annoying" ● 30:26 [We hear a sound of a group whispering] P1 "I like that one" P3 "It's frightening" P3 "I like that [sound]" ● 31:10 [We hear a sound of a group whispering] P4 "The market" [2,8] ● [We hear a sound of a scream] 31:58 P3 "Including this freak" ●
	SP: 9	06:11 [We hear a sound of a car crash] P4 "That wasn't mine" P1 laughs ● 06:24 [We hear a sound of a gun shot] P1 laughs ● 06:55 [We hear a sound of a purring cat] All laugh ● 07:04 [We hear a sound of a tom drum] P3 "Nice this one" ● 08:52 [We hear a sound of a long snort] P2 "Is that one?" P3 "Yeah" ● [We hear a sound of a cat howling] 09:51 P3 "This is me [replying to P4's question about "Who's got the cat one?"] (...) it's one of them cause I was using it before" ● [We hear a sound of a car crash] 11:22 P1 "Such a window" [2,Ambisonics and realism] ● 12:05 P3 "I've got it" [we hear a sound of a cat howling] [in response to P2 "Who is the cat?"] ● 17:20 [We hear a sound of a kid laugh] P1 "Oh I don't like that" P2 "It's terrifying" ●
G2	NSP: 19	09:54 P1 "Let's try a different sound, if you flip the cube" P2 "Ok" [P4 changes the side of the cube, now we hear a metallic-ceramic friction] P2 "[laughing] Silly sounds [turning her head towards her left back speaker]" ● [We hear a sound of a cat purring] 12:16 P2 "Nice" ● 12:25 [We hear a sound of a wind mouth gurgle] P2 "Actually, to be fair, the volume is very low" ● 12:58 [We hear a sound of a crowd murmur] P2 "That's good actually" ● 13:48 P2 "Anyone remembers which was the nicer [sounds from the 'whispers' cube]" P2 "that's thrilling, there is only one interesting sound I think in that one [white cube] I'll put that way up then" [P2 positions a white cube she is manipulating in the rim area next to her] ● 14:37 P2 "Cause if I take that one off does it make any difference? [P2 lifts up a white cube] not really" ● 15:13 P3 puts an object [we hear a sound of a long snort] All laugh P3 "That's quite nice" ● 15:25 P2 "Why don't you put this one on the winding" [we hear a sound of a water pipe hiss] ● 16:26 [We hear a sound of a tunnel short whistle] P2 "Uhh is it that one?" P4 "Yeah" P1 "That's just a short sound" P2 "This sound it's supposed to be a fitty noise" ● 16:40 [We hear a sound of a lamp buzz] P4 "this is a (...) buzzing" [in response to P1's question "which sound is which now?"] P1 "that's a buzz (...) they take a while, they take a second before we ... [hear the sound] (...) That's horrible, can we change it? That is like the computer humming in the background" P4 "It's like a generator" P2 "It's slightly irritating this" P4 "All right, ok" [2,3,8] ● 17:04 P2 "It's also an irritating noise, really" P4 "which one?" [while changing again to another face of a white cube] ● 18:33 [We hear a sound of a stomach gurgling] All laugh. P2 "Oh nice, I like this one" ● 18:50 P3 "The volume down [while pointing to a white cube]" ● 19:37 [We hear the sound of a scream] P1 "That's a horror film now" [2,8] ● 19:51 P2 "Someone is being murdered" [2,8] ● 20:18 [We hear a sound of a metallic impact] P1 "This is the ding" P2 "That's the ding" ● 20:25 P2 "Is that keep making a ding?" [pointing to a white cube operated by P1 that was producing a a sound of a metallic impact a few seconds before] ● 21:33 P4 "That's me" [replying to P2's question about "Is that your snoring?"] P2 "Oh it's you snoring" [2,3] ● 23:43 [We hear a sound of a tunnel short whistle] P2 "Uh it's yours right? [pointing at a white cube next to P4] you've got the fluty noise" [looking at P4] P4 "Yeah" [2,3,8] ●

	SP: 18	24:45 [We hear a sound of a fire crackle] P4 "We can just put down the campfire" P2 "yeah we just love the campfire in the corner" [2,3,8] ● 25:42 [We hear a sound of a group whispering] P4 "and then... some ghosts [laughing]" P1 "we can keep the ghosts but we can decrease the volume" P2 "or we can try a different one" [2,8] ● 26:06 [We hear a sound of a stomach gurgling] P2 "Someone's rumbling tummy, I don't like that one, what about you?" [replying to P4's question about "What is this supposed to do?"] [2,3] ● 26:21 P4 "That's a cat" [and moves towards his right back speaker] P3 "Oh yeah I can hear it now" P4 "It's like... 'catish'" [everyone laughes] [2,3] ● 26:32 [We hear a sound of a cat purring and a fire crackle] P4 "You know, next to the fire" [2,3] ● 26:36 P4 "It's just very deep, almost like 'prrrrr'" [replying to P2's comment "I can't hear the cat"] P2 "Purring is it? Alright, okay" [2,3] ● 27:32 P2 "Where's the volume in that one? it's a little bit loud isn't?" [referring to a sound of a long snort] ● 27:51 P4 "It's quite loud" P2 "it's still quite loud, it changes quite randomly, isn't? (...) now we are getting there" ● 28:08 P2 "Yeah that sounds right, that sounds better" ● 28:52 P2 "Wasn't this a very useless one" [talking about the 'percussion' white cube] P4 "You can sometimes make it gong" [2,3] ● 30:34 [We hear a sound of a man laugh] P2 laughs ● 29:04 P1 "That's the gong [turning the face of a white cube, we hear a sound of a metallic impact]" [in response to P2's question "is that the gong?"] P3 "Yeah" ● 30:12 P1 "The fire sounds really nice, it sounds as it is actually fire" [2,3,8] ● 30:38 [We hear a sound of a cat howling] P2 "Uhh" ● 33:39 [We hear a sound of snoring] P2 "Make it a little bit louder" [laughs] ● 34:42 [We hear a sound of a man snoring, and of a kid laugh] P2 "It's a little bit irritating isn't? P4 "Yeah the baby's sound" [2,3] ● 35:34 P2 "Let's go for the snoring and the baby" P4 "The noisy baby (...) it is too loud" P2 "It is too loud, isn't?" [2,3,8] ● 38:37 P4 "there is no change" ●
G3	NSP: 0	–
	SP: 1	07:22 P2 "Cool!" [P1 laughs] [2,9] ●

G4 NSP: 29 06:07 [We hear the sound of a ceramic impact operated by P1] P2 "That's a nice... to go with the happy baby" P1 nods his head ● 06:09 [We hear a sound of a thunderstorm] P3 "That's a thunder line that's not good" P2 "That goes with the scream of the baby, with the baby crying" [2,3,8] ● 06:21 [We hear a sound of a kid laugh] P2 "This laugh is a little evil" [2,3,8] ● 06:29 [We hear a sound of a metallic whine, and a sound of a metallic-ceramic friction] P2 "Mine ['screeches' white cube/sound] is really nasty" [1,2] ● 06:56 [We hear a sound of a purring cat] P2 "Must be yours" [looking at P3] P2 "Ohhh that's nice" ● 07:12 P4 changes the face of a white cube [we hear a sound of a groan] All laugh ● 07:20 [We hear a sound of a long snort] P4 "That's boring" ● 08:04 P2 [We hear no sound from the 'screeches' white cube] "Mine ['screeches' white cube/sound] is not doing anything anymore I've got a really... [now we hear a sound of a lamp buzz]" [1,2] ● 08:13 [We hear a sound of a groan] P1 "That's just disturbing" [laughs] ● 08:28 P2 "Mine ['screeches' white cube/sound] seems to just disconnect" [1,3] ● 08:32 We hear a sound of a stomach gurgling, P3 and P4 laugh ● [We hear a sound of a groan] 08:58 P2 "Who's got this thing? This voice? This lament?" [2,4,8] ● 09:02 [We hear a sound of a scream from P4's white cube] P3 "[inaudible] creatures, does it? [pointing to P4's white cube]" ● 09:05 [We hear a sound of a tom drum] P1 "No, I, mine ['percussion' white cube/sound] is the strong [and demonstrates by lifting up and down the white cube] that one" P3 "Ok" P2 "Uh" [in response to P2's question "Who's got this thing? This voice? This lament?" [referring to a sound of a groan] [1,2] ● 09:24 [We hear plastic rustling] P2 "again!" P1 "What's that sound...[looking at P2] That sound is coming from your block [pointing to P2's white cube]" [2,5] ● 09:40 [We hear a sound of plastic rustling] "That's me, kijkijkijkij" ● 09:44 [P2-P4 wonder who has the white cube with a sound of a groan] P3 "Ok turn yours [P2] off" [P2 lifts up her white cube, there is now P1 and P4 manipulating cubes, P1 is lifting up and down a cube with a sound of a ceramic impact, and we still hear a sound of a groan] P2 "Uh no that's not me" P4 "Might be this one [pointing to a white cube he is manipulating, he lifts it up and we stop hearing a sound of a groan]" P2 "Aha! [pointing to P4's 'voices' white cube]" P4 "Uh [it was] that one [while changing the face of the 'voices' white cube]" ● 10:22 P1 "I am gonna give you guys this rhythm [while lifting up and down rhythmically a sound of a wood impact] (...) See if you can have some sounds on top of that rhythm" [2,8] ● 10:41 P2 "So I am giving you for 3 bits of that I am giving you a length of noise [we hear a sound of plastic rustling]" [2,8] ● 10:52 [We hear a sound of a kid laugh] P4 "I should only laugh at the end (...) oh actually it is still laughing even though I lifted it up [we can see the shade of his white cube]" P3 "yeah I can see the reflection" P4 "Yes it is still reflecting, oh I see" [2,3,8] ● 11:52 [We hear a sound of a war boom explosion] P3 "The finale" P4 "We are still alive" [2,8] ● 14:02 P4 "This one" [in response to P2's question "Which one is your 'amplifier'?"] P2 "That one?" P4 "That is the tikitikitiki [pointing to P2's sound] P4 "That's the amplifier" P2 "Uhhh nice!" P3 "it's also the loop into it" P2 "Oh nice! This goes nice together" [2,6,7] ● 15:15 [We hear a sound of a wind mouth whisper] P2 "Wow" P4 "How did you call it? spooky sound" ● 15:41 [We hear a sound of a water splash with an effect] P3 "[inaudible] is coming" P2 laughs ● 17:19 [We hear a sound of a scream] P4 "That's this one" [while lifting a white cube up] P2 "Ah you were the scream [pointing to P4]" [2,3] ● 17:37 [We hear a sound of a tunnel short whistle] P2 "No, that's me" [in response to P3's question "Uh is that mine?"] ● 18:12 [We hear the sound of a scream] P1 and P2 laugh ● 19:05 [We hear a sound of a tibetan chant] P2 "Who's got this? ... you? [pointing to P4]" P4 "This one yeah" P2 "Oh cool" P1 "Didgeridoo" P3 "[inaudible]" P2 "I love that one" [2,5] ● 20:32 [We hear a cat howling] P2 "Uhhh" ●

SP: 25 [We hear a sound of a falling metal] 21:22 P3 "Uh that was a nice noise" 21:25 P2 "Doong!" • 21:38 P1 "This is this horn [while lifting up and down a white cube with a sound of a tibetan chant]" P2 "Nice!" P3 "A great sound" [2,5] • 22:13 P4 "It is this one" [showing his white cube with a sound of a thunderstorm] • 22:27 P3 "It's a really amazing sound" [we hear a sound of a tibetan chant] • 22:32 [We hear a sound of a leaking gas hiss from the 'whistles' white cube] P2 "I ended up... did I ended up with the same thing? no!" • 22:36 [We hear a sound of an explosion] P3 and P4 laugh • 23:45 [We hear a sound of a gun shot] P2 "You put some nice percussion..." • 25:55 We hear a sound of a pipe hiss that appears and disappears] P2 "So if you rotate constantly it goes to a maximum and then it goes to zero so you have like a..." • 26:52 [We hear a sound of a pipe hiss that appears and disappears] P2 "My role is to turn this one [laughs] I think it's a bit [inaudible]" [2,8] • 27:46 [Each of the four team members is producing sounds] P2 laughs • 28:13 P4 "They turned all to one single colour" P3 "Do they? the cubes [in6,audible] colour" P4 "so these are all white now" [pointing to a white cube and two black cubes next to the white cube] • 32:28 [We hear a sound of a gun shot] P3 "This is a nice noise" • 32:47 P2 "Right, I think I've done that long enough" [making a sound of a pipe hiss appearing and disappearing] P3 "You got bored!" P2 "Maybe I can try another..." [2,8] • 32:56 P3 "That sound doesn't work! [looking at P4 while P4 is stacking 4 cubes]" [P4's stack of cubes fall apart] P2 laughs • 33:49 P4 is dragging 2 groups of objects until they touch to each other P3 "You are not crashing things into one, are you?" [looking at P4] P2 "What are you doing?" P4 "Composition various different composition" [2,8] • [We hear a sound of a tibetan chant] 34:12 P3 "I think we agree that the [tibetan chant] singing is the best sound" P2 "It's definitely very atmospheric" • 34:24 [We hear a sound of a long snort] P2 "Who is got this noise?" [2,4] • 34:26 [We hear a sound of a long snort] [P2 lifts up and down and realises that she has this noise after asking herself "Who is got this noise?"] P2 "I have this noise? that doesn't go with the [inaudible]" • 35:51 P1 "Now it has become a cacophony I think (...) this is whatever John Cage's Silence but the opposite" P3 "Yeah [laughs]" [2,8] • 36:40 [We hear a sound of a long snort, and of a cat howling, and of a kid laugh] P2 "These are neighborhood sounds" 36:44 [We hear of a tibetan chant] P2 "Not that one" [2,8] • 37:28 P4 "0 and 9 are very close" [while turning a white cube] • [We hear a sound of a long snort] 37:32 P3 "You really like that noise, don't you? [looking at P2]" P2 laughs • 37:56 P2 "You can't do it fast it doesn't... [looking at P1 while P1 is lifting up and down rhythmically a white cube with a sound of a tom drum and P2 is lifting up and down a white cube with a sound of a tunnel short whistle] • 38:02 [We hear a sound of a war boom explosion] P4 "I've got 4, I've got 4" • 39:10 [In response to P2's question "Who's got this?"] P3 "[These are all these things] in the middle" P2 "all together?" P3 "Yeah" P2 "That's very cool!" [2,9] •

G5	NSP: 15	24:11 [We hear a sound of a gun shot] P3 "Perfect" ● 24:33 [We hear a sound of a tibetan chant] P1 "It's a good one" P2 "Boring" P1 "The boring chant or the orthodoxy drowning" P2 "Heavy metal" P1 "No unfortunately one of the [inaudible] just can't sleep" P2 and P4 laugh [2,3,8] ● 24:55 [We hear the sound of a ceramic impact] P2 "Uh that's the same than before" ● 25:26 P1 "Take the percussion now [inaudible], industrial based" [2,8] ● [We hear a sound of a tibetan chant] 27:13 P1 "That drone is the winner" P2 "Yeah" P1 "Are you making that drone?" P4 "What?" P2 laughs P1 "Are you making the drone yourself?" P4 "This one [pointing to a white cube]" P1 "I think you need to mark somehow that box" P2 "Yeah" P1 "That's the winner" [2,4,8] ● 27:26 P4 "Yeah, and they are all here" [talking about a white cube that has a sound of a tibetan chant, in response to P4's comment that "That drone is the winner] P4 turns a white cube and it changes the sound from a tibetan chant to a cat howling P3 "Not only [tibetan chants]!" [1,2] ● 27:34 [We hear a sound of a tibetan chant] P1 "Keep with the drone is good" [2,8] ● 28:16 [We hear a sound of a leaking gas hiss] All laugh ● 29:14 P3 turns a white cube and we hear the sound of a cat howling, P1 nods his head, P2 says "hum" and laughs, P3 laughs and imitates the sound of a cat howling "ummm" ● 31:53 [We hear the sound of a cat howling] P4 and P1 are laughing ● 31:59 [We hear the sound of a kid laugh] P1 says "No" with his head while laughing ● 34:14 [P4 is lifting up and down rhythmically a white cube with a sound of a tunnel short whistle with an effect while P1 is nodding his head] P4 "It seems to pick it up you know? [pointing to a set of black cubes]" [2,6] ● 35:19 [We hear a sound of a cat howling with an effect] P2 laughes P3 "That's good" P2 & P3 are laughing P4 "Not sure if it is a baby or a cat" P3 "It's different" [2,6] ● 36:33 [We hear a sound of a car crash] P4 raises his two hands 36:36 P3 "Strike!" [2,8] ● 38:48 [We hear a sound of a cat howling, and then a sound of a kid laugh] P1 "There's something quite chilly about the cat, it sounds like a baby, ours it's quite a bit like that, especially when it wags" P2 "When I was a kid and used to hear a cat crying outside I used to take my dad it sounded like a baby on the doorstep making my dad look" P1 "True, it's absolutely chilly" [2,3,8] ●
	SP: 13	08:16 [We hear a sound of a boom rumble] P1 "Uhh I like that one" ● 10:36 P1 "I'm just doing a pulse [while lifting up and down a white cube with a sound of a gun shot]" P2 "Well done" [2,8] ● 10:48 [P4 lifts up a white cube and a sound of a leaking gas hiss stops] P4 "I think it's this one" [in response to P3's question "Who's doing this... gas? [laughs]" ● 11:00 P1 "That works [nodding his head following a rhythm lifting up and down a white cube with a sound of water splash]" [2,8,9] ● 11:56 [We hear a sound of a man snoring] P2 and P3 laugh P3 smiles ● 12:00 [We hear a sound of a man snoring with an effect] [P4 lifts up and down a white cube with a sound of a man snoring] P2 "It's that one [pointing to P4's white cube]" [in response to P1's question "We've got a snore, who's got the snore?"] ● 12:10 P1 [We hear a sound of a man snoring with an effect] "There's something happening there" [after asking who's got the snore and P2 pointing to the white cube with a sound of a man snoring that is next to a black cube] P2 "It's like it picks up the rhythm by itself" [2,6] ● 13:10 [We hear a sound of a metallic impact] P2 "Wow" ● 13:18 P1 "We got a few..." P2 "Drumming [while lifting up and down a white cube with a sound of a metallic impact]" [2,8] ● 13:36 [P2 is switching between two sounds of a cube rhythmically, all of the sudden there is silence] P2 "Uhh" ● 13:41 P2 "Nice [inaudible] [while lifting up and down a white cube alternating between a sound of a metallic impact and a tom drum]" P3 "[inaudible] slow movement" [2,8] ● 15:05 [We hear a sound of a car crash] P2 "That's good" P2 "Well done" ● 20:49 [We hear a sound of a water splash with an effect] P1 "I like that [while turning a white cube left and right next to a black cube and nodding his head rhythmically]" P1 & P2 are nodding their heads [2,6,9] ●
G6	NSP: 11	09:16 [We hear a sound of a ceramic impact with an effect] P2 & P1 laugh ● 09:53 [We hear a sound of a groan] P2 laughs looking at P3 ● 09:26 [We hear a sound of a kid laugh with an effect] P3 "That's creepy" P1 & P2 laugh [2,6] ● 10:42 [We hear a sound of a groan] P2 & P3 laugh ● 10:44 [We hear a sound of a groan] P3 "No, it's not me (...) It's stunning" [in response to P2's question "Is that you?"] ● 10:50 [We hear a sound of a metallic whine] P3 "I've got the highest" [2,8] ● 10:51 P1 "It's this one [lifting up and down a white cube with a sound of a groan]" [in response to P2's question "Is that you?"] P2 & P3 laugh P2 "Ah" P3 "yeah" ● 15:10 P1 & P3 smile with the sound of a kid laughing with an effect [2,6] ● 16:10 [We hear a sound of a car crash] P2 "That's me! [while lifting it up and down]" P1 "Yeah" ● 17:19 [We hear a sound of rocks smashing] P4 "I like that one" ● 22:22 [We hear a sound of snoring] P1 laughs ●

	SP: 5	26:29 P2 "No, another [inaudible] sound" [and swaps one white cube with a black cube] ● 30:42 [We hear a sound of a groan] P1 & P2 smile ● 32:02 P2 lifts up and down rhythmically a white cube with a sound of a gun shot, P1 laughs ● 38:16 P2 "Let's put all white ones in full mode" [2,8] ● 39:01 [We hear a sound of a snoring] P1 & P2 laugh ●
G7	NSP: 22	26:40 [We hear a sound of a air escaping hiss] P2 "Nice" ● 26:45 P2 "So we discover all the sounds of this one ['whistles' white cube] (...) The interesting one is snoring" [1,2,3] ● 27:15 P2 "Ok maybe we can substitute it [white cube with a sound of a long snort] with another white one [as maybe] it sounds better" [we hear a sound of a tom drum] ● 29:19 P2 [We hear a sound of a ceramic impact] "We know everything here ['percussion' white cube]" P4 "It is very exciting" P4 "This is still scenario, or like still things" ['percussion' white cube] P2 "Yes" [1,2] ● 29:39 [We hear a sound of a war boom explosion] P2 "Dangerous" P4 "It's the boring one" P4 "Yes this is the boring one [pointing to 'percussion' white cube] and this is the bad one [pointing to 'explosions' white cube]" P4 laughs [1,2] ● 30:27 [We hear a sound of a man laugh] P2 "This is not bad" ● 30:37 [We hear a sound of a kid laugh] P2 "Ok this is the good one" P3 "I like the baby" [2,3] ● 30:40 [We hear a sound of a groan] P4 "So after that" [Just before there was the sound of a kid laugh and P2 said "Ok this is the good one" and P3 said "I like the baby"] P2 "Exactly" ● 30:46 [We hear a sound of a cat howling] P4 "I like every cat" [2,3] ● 31:48 [We hear a sound of a wind mouth gurgle] P2 "This is another one [white cube/sound]" [1,2] ● 32:13 [We hear a sound of a group whispering] P3 "This is my favorite [sound]" P2 & P3 laugh ● 32:21 [We hear a sound of a stomach gurgling with an effect] P2 "What is this? something under water I think" [2,6] ● 33:59 P2 "What sounds have we got?" P4 "We've got snoring boring creepy cats another like boring" P2 "The cat one is good" P3 "Cats and babies" 34:13 P3 "And the... " P4 "Oh yeah, and the exciting one" [1,2,8] ● 34:57 [We hear a sound of a kid laugh] P2 "Oh yes yes baby is also nice" [2,3] ● 35:46 [We hear a sound of a kid laugh with an effect] P3 "That's freaky" P2 & P4 laughs [2,6] ● 35:52 [We hear a sound of a kid laugh with an effect] P2 "Ok so this is the sound of the parent making jokes out about himself" P3 "It looks an old man laughing" All laugh P3 "Or maybe a bit obsessed" P4 "Sounds like an exorcist stuff" P2 "So maybe we can try if it is snoring to have some fun" [2,3,6,8] ● 36:22 P4 "This one is the creepy one, [we hear a sound of a gun shot]" 36:27 P2 "Kill the baby" 36:30 P4 "And it's so wrong" [1,2,8] ● 36:37 [We hear a sound of a car crash] P3 "Just shot the baby" P2 laughs [2,8] ● 36:48 [We hear a sound of a gun shot] P4 laughs ● 37:19 [We hear the sound of a man laugh] P2 laughs P2 "Let's make a horror movie" [2,8] ● 37:25 P2 "So we have the cat, we have the shooting and..." P4 "We have like the sony baby" P3 "We have a creepy background" P2 "Exactly we have the [inaudible] baby" P3 "It's walking down a park" P2 "And maybe someone is sleeping and gets shot in the sleep [we hear a sound of a gun shot], you can also add the snoring one" P3 laughs P4 "Yes, so there's a shot as long as I know" P2 "This is the snoring [we hear a sound of a long snort]" P4 "Somewhere yeah" [we hear a sound of a groan] P2 "This is the shot" [we hear a sound of a gun shot] P3 "It's like [inaudible]" P2 "And then burn it" P2 "Ok I think we left them not very good" [2,3,8] ● 39:08 P2 Yeah this one is red [pointing to the explosions white cube while P4 is trying whether it is working or not], probably, maybe we can..." [then we hear a sound of a water splash from the explosions white cube, and P4 leaves it] [1,2] ●

	SP: 22	11:05 [We hear a sound of a car crash] P4 "Is that the red?" P2 "I think it's the red [inaudible] the gat sounds" [1,2] ● 11:19 [We hear a sound of a car crash, a boom rumble, a pipe hiss] P2 "So let's make this positive" P1 & P3 laugh P2 "Happy now" [2,8] ● 11:47 P2 "This is the problem [while turning down volume of the red white cube]" P1 "Thank you" ● 12:06 [We hear a sound of a cat howling] P4 laughs ● 13:42 [We hear a sound of a scream] P3 laughs P2 "Again and again" P1 "Sorry" ● 14:25 [We hear sound of a kid laugh] P4 laughes ● 12:32 [We hear a sound of a cat howling] P2 "I like this cat" [2,3] ● 14:31 P2 "Let's put the volume low" 14:36 [We hear a sound of a ceramic impact from a white cube next to P3] P1 "Maybe it's too loud [while turning down the volume of P3's white cube with a sound of a ceramic impact]" 14:37 P4 "And quiet" P1 "And quiet" ● 14:46 [We hear a sound of a gun shot from the red white cube] P2 "This one [sound/white cube] is negative [pointing to the red white cube]" P4 "The red" P2 "Yes" [1,2] ● 16:55 [We hear a sound of a cat purring with an effect, and a cat howling] P3 laughs P4 smiles ● 17:39 [We hear a sound of a tibetan chant, a ceramic impact, a fire crackle, a boom rumble] P3 "It's like noises from films" [2,8] ● 17:52 [We hear a sound of a crowd murmur] P4 "Oh dear [laughs]" ● 18:11 [We hear a sound of a crowd murmur with four black cubes affecting it] P2 "This is more scaring" P3 "Yeah [laughs]" [2,6] ● 21:09 [We hear a sound of a kid laugh from the 'voices' white cube with four black cubes next to the white cube, manipulated by P3] P2 "Let's keep it like this [laughs looking at P3]" [2,6] ● 21:40 [We hear a sound of a kid laugh and a group whispering] P3 "This is my favourite sound I think" P2 "The baby?" P3 "This thing, the whispering [pointing to the 'whispers' white cube]" P2 "This is the scary one" [2,3] ● 22:15 [We hear a sound of kid laugh with an effect] P4 "Is this the echo?" P3 "Creepy" P2 "This is the echo [pointing to a black cube]" [2,6] ● 22:25 [We hear a sound of a gas hiss] P2 "This is the pheeeew [pointing to the 'whistles' white cube]" ● 22:30 [We hear a sound of a kid laugh with an effect, and a gas hiss with an effect] P4 "[laughs] This is so creepy" [2,6] ● 23:21 [We hear a sound of a man laugh] P3 "That" [while lifting up and down her 'voices' white cube in response to P2's question "Which one is laughing?"] ● 23:23 [We hear a sound of a man laugh] P3 "It is so [adventurous] [laughs]" [2,8] ● 23:26 P4 "Like this one [lifting up a white cube with a sound of a tom drum and then lifting it down]" ● 24:02 [We hear a sound of a howl] P2 laughs ●
G8	NSP: 17	08:06 [We hear a sound of a cat howling] P3 "This is the cat" All laugh 08:07 P1 "Yeah" [2,3] ● 08:13 [We hear a sound of a groan] P1 and P3 laugh ● 08:37 [We hear a sound of a ceramic impact] P3 "Okay" [and leaves the white cube on the tabletop surface] ● 08:44 [We hear a sound of a stomach gurgling] P2 laughs ● 09:29 [We hear a sound of a metallic whine] "Noises" ● 09:31 [We hear a sound of a howl] P1 "That's scary" ● 09:06 P3 [We hear a sound of a crowd murmur] "If we rotate it? [to P2]" P2 "Crowds" P4 "It just changes the volume" [2,3] ● 09:49 [We hear a sound of a kid laugh] P3 laughs ● 09:50 [We hear a sound of a cat howling] P2 laughs P3 "What's the other one then? [pointing to the other white cube that P2 is holding]" [Impression sounds, 5] ● 15:09 [We hear a sound of a plastic rustling with an effect] P3 "[inaudible] very loud" P2 "[inaudible] it's too very loud" [Impression sounds, Identification filter] ● 15:29 P3 "Now we should turn it loud again [while rotating the white cube with a sound of plastic rustling and then a black cube that increases the volume of the sound of a plastic rustling]" [2,6] ● 15:39 [We hear a sound of stomach gurgling] P1 laughs ● 16:27 [We hear a sound of a man laugh with an effect, before it was a kid laugh from the same cube] P1 "This is the other one [laughs]" [2,6] ● 17:10 [We hear a sound of a tibetan chant, and a howl] P1 "It's scary" ● 18:43 We hear a very loud sound [of a car crash], P3 and P4 look at each other ●[All participants are manipulating white cubes and black cubes] 20:22 P2 "Funny noises [inaudible]" ● 21:45 [We hear a sound of a kid laugh] P3 laughs ●

	SP: 14	23:13 [P3 is holding a black cube and P4 is changing the sides of the 'whispers' white cube] P3 "Nice" ● 23:30 P2 [We hear a sound of a crowd murmur] "Ok that's the crowd noise, it's got a good one actually" [2,3] ● 24:44 [P2 adds a new white cube, the 'explosions' cube, with a sound of a boom rumble next to a black cube and the 'whispers' white cube] P3 "Let's do this one [the new white cube], let's take this one out [moving to the rim area the 'whispers' white cube with a sound of a crowd murmur]" [Each of the participants puts a black cube next to the white cube with a sound of a boom rumble] P3 "Quiet quiet" P3 "It's that everything down there?" 25:19 P2 "It's possible" P2 "Ooops" P3 "Maybe it does... maybe those things up" [2,7] ● 25:51 P2 "Uh I quite like that [creating a rhythm by lifting up and down the 'whistles' white object with a sound of an air escaping hiss]" [2,9] ● 26:47 P2 "I give you the bit [while lifting up and down rhythmically the 'whistles' white object with a sound of an air escaping hiss]" [2,8] ● 27:36 [We hear a sound of a cat howling] P3 laughs "once again" ● 28:04 [We hear a sound of a scream] P2 "I like that one" P1 "This is the funny box [laughs]" P3 laughs [1,2] ● 28:43 [We hear a sound of a man laugh with an effect] P3 "That makes a difference" P1 there is a second voice?" P1 "There is a second voice" [2,6] ● 30:53 [We hear a sound of a scream] P1 "Screaming" P2 & P4 laugh [2,3] ● 31:18 P2 "It's that one [white cube]" [in response to P3's question "Is that the one?"] ● 31:29 [We hear a groan with an effect] P1 & P3 laugh 31:42 P2 "We've got a man groaning" P2 "It's a little bit..." P3 "...Spooky" [2,3,6] ● 35:25 [We hear a sound of a long snort] P2 "I don't quite like that [while changing from the sound of a long snort to a sound of a pipe hiss, to a long snort, to a tunnel short whistle]" ● 35:34 [We hear a sound of a tunnel short whistle] P3 "Haven't we met that one before?" ● 37:38 [Silence] P3 "Silence, oh that's fantastic" [2,3] ●

D.8.3   Identification of source

TABLE D.6: Transcripts of *Identification of source*

Group	Condition	Quotes
G1	NSP: 4	29:46 [We stop hearing the sound of a water splash] P2 "Stop splashing now" ● 30:01 [We hear a sound of a group whispering] P1 "It's a crowd" ● 30:16 [We hear a sound of a cat purring] P1 "Oh that's purring" ● 34:36 [We hear a sound of a long snort] P2 "Snoring" ●
	SP: 18	06:24 [We hear a sound of a gun shot] P4 "The gunshot" ● 06:29 [We hear a sound of a ceramic impact] P1 "You got the bell" [looking at P2] [3,8] ● 09:43 [We hear a sound of a gun shot] P3 "This is the gunshot" P4 "That's the gunshot" ● [We hear a sound of a man laugh] 11:13 P2 "You've got the evil laugh" P2 "Yeah, ok, I've got the evil laugh" [3,8] ● 09:48 [We hear a sound of a cat howling] P4 "Who's got the cat one?" [3,4] ● 11:17 P1 "I've got the baby" P3 "The killer baby" [3,8] ● 11:28 [We hear a sound of a gun shot] P2 "Who got the gunshot?" [3,4] ● 11:30 P4 "That's the gunshot" [we hear a sound of a gun shot] ● 11:32 P3 "That's the baby, I've got the evil baby" [3,8] ● 11:34 [We hear a sound of a water splash] P4 "[inaudible] The splash" ● 12:07 P1 [We hear a sound of a cat howling] "You've got the cat" P3 "Cat? it's a huge cat" [3,8] ● 12:08 [We hear a sound of gun shot] P2 "The shoot cat" [3,8] ● 12:13 P3 "I've got the baby" ●13:14 [We hear a sound of man laugh] P3 "Take the baby evil up" [3,8] ● 15:24 [We hear a sound of fire crackle] P3 "Sound effects" P4 "Just crackles" [3,8] ● 17:09 P3 "Explosions" [in response to P1's question "What is that one [white cube/sound]?" after hearing a sound of a war boom explosion] [1,3] ● 17:26 [We hear a sound of a war boom explosion] P3 "Explosions" ● [We hear a sound of a man laugh] 19:46 P2 "Evily laugh" [3,8] ●

G2 NSP: 14 [We hear a sound of a thunderstorm] 14:04 P2 "thunder wind" P1 "so we do a storm" [3,8] ● 14:10 [We hear a sound of a fire crackle] P2 "Wow is that fire? (...) what you got there, a fire? [pointing to a white cube]" P4 "Yeah I think it is the fire [while lifting up and down again a white cube] P3 "oh yeah" ● [We hear a sound of a water pipe hiss] 15:37 P2 "A humming or something" ● 15:51 P1 "Should try enhancing the wind" P2 "Yeah I can't hear this wind at all [while lifting up a white cube with a sound of a wind mouth gurgle]" [3,4] ● 17:56 P4 "I think it's the fire" [replying to P2's question about "What's that one then?"] P3 "Yeah" ● 16:40 [We hear a sound of a lamp buzz] P4 "this is a (...) buzzing" [in response to P1's question "which sound is which now?"] P1 "that's a buzz (...) they take a while, they take a second before we ... [hear the sound] (...) That's horrible, can we change it? That is like the computer humming in the background" P4 "It's like a generator" P2 "It's slightly irritating this" P4 "All right, ok" [2,3,8] ● 18:22 P2 [We hear a sound of a crowd murmur] "That is not wind is it? This is more like people" ● 19:43 P4 "Let's put down the fire just to keep" P2 "[laughing] Now it's gonna burn in the background" [3,8] ● 21:13 P3 "That's me" [replying to P2's question about "Who's got that one?"] ● 21:14 P2 laughs and says "You've got the people noises" [1,3] ● 21:30 P2 "Is that your snoring?" [asking to P1] [3,4] ● 21:33 P4 "That's me" [replying to P2's question about "Is that your snoring?"] P2 "Oh it's you snoring" [2,3] ● 23:43 [We hear a sound of a tunnel short whistle] P2 "Uh it's yours right? [pointing at a white cube next to P4] you've got the fluty noise" [looking at P4] P4 "Yeah" [2,3,8] ●

SP: 36 24:45 [We hear a sound of a fire crackle] P4 "We can just put down the campfire" P2 "yeah we just love the campfire in the corner" [2,3,8] ● 24:54 [We hear a sound of a fire crackle] P1 "So the fire was nice for the background so we leave it" [3,8] ● 25:17 [We hear a sound of a thunderstorm] P1 "That's the thunder right?, so we'll leave it with the fire" ● 25:22 [We hear a sound of a wind mouth gurgle] P2 "Oh it's the wind" ● 26:06 [We hear a sound of a stomach gurgling] P2 "Someone's rumbling tummy, I don't like that one, what about you?" [replying to P4's question about "What is this supposed to do?"] [2,3] ● 26:21 P4 "That's a cat" [and moves towards his right back speaker] P3 "Oh yeah I can hear it now" P4 "It's like... 'catish'" [everyone laughs] [2,3] ● 26:32 [We hear a sound of a cat purring and a fire crackle] P4 "You know, next to the fire" [2,3] ● 26:36 P4 "It's just very deep, almost like 'prrrr'" [replying to P2's comment "I can't hear the cat"] P2 "Purring is it? Alright, okay" [2,3] ● 26:44 [We hear a sound of a tunnel short whistle] P1 "This could be like a sort of a squeaking metal" P2 "Oh yeah it could be" ● 26:53 P2 "Oh that's your purring. Oh we've got a proper purring cat, that's better." ● 27:07 [We hear a leaking gas hiss] P1 "This could be something on the stove" [3,8] ● 27:16 P1 "A guy sleeping" P2 "Snoring" ● 27:18 P2 "A bit more realistic, it is a bit more realistic the snoring..." [Identification sounds, 8] ● 30:12 P1 "The fire sounds really nice, it sounds as it is actually fire" [2,3,8] ● 30:45 [We hear a sound of a cat howling] P1 "It sounds more like a cat" ● 30:59 P1 "That's a crying baby" P2 "A screaming one, scared of the thunder" [3,8] ● 30:49 [We hear a sound of a kid laugh] P1 "That's a baby" P3 "Baby" P2 "I don't mind having a baby, but should be laughing or crying?" ● 31:35 [We hear a sound of a cat howling] P2 "Uh that, whatever that is, that's another cat" ● 31:42 [We heard a sound of a kid laugh] P1 "That's a crying baby, we should decide whether a crying baby or a laughing baby" P4 "I aim for the laughing baby" [3,8] ● 31:55 P4 "I was thinking maybe we could put out the snoring" ● 32:56 [We hear a sound of air escaping hiss] P1 "Sounds like a teapot [laughs]" [3,8] ● 33:08 [We hear a sound of a pipe hiss] P1 "Like a plane taking off [3,8] ● 33:25 [We hear a cat purring] P2 "Is that the cat purring?" [pointing to a 'wrong' white cube] [3,4] ● 33:27 [We hear a sound of a purring cat] P4 "That is the cat purring" [replying to P2's question "Is that the cat purring?"] ● 33:30 [We can't hear a sound of a long snort] P2 "I've lost the snoring now (...) [while turning the face of a white cube]" [We hear a sound of a long snort] "There's the snoring!" ● 34:42 [We hear a sound of a man snoring, and of a kid laugh] P2 "It's a little bit irritating isn't?" P4 "Yeah the baby's sound" [2,3] ● 34:58 [We hear a kid laugh] P3 "That's the baby" ● 34:57 P3 points to a white cube and says "That's the baby" [in response to P4's action of lifting up a white cube thinking that it was the sound of the baby] ● 35:34 P2 "Let's go for the snoring and the baby" P4 "The noisy baby (...) it is too loud" P2 "It is too loud, isn't?" [2,3,8] ● 36:00 [We hear a sound of a thunderstorm, a tom drum, a fire crackle, a kid laugh, and a purring cat] P2 "This sound is awesome" P4 "We can make some noises" P2 "We created a scene that's all we wanted to do, wasn't it?" P1 "That's the purpose, we have to combine them" P2 "We have our urban noises we've got a purring cat by a fire, and a baby, and some snoring, and a thunder, and a banging door; it seems realistic at least in terms of urban noises" [3,8,9] ● 36:28 [We hear a sound of a groan] P2 "What we've done now?" Uh you have changed the noise, you've got a groaning man instead" ● 37:32 P2 "This one.. this one was the snoring, the snoring man is not actually by the fire though [pointing to a white cube with a sound of a long snort]" P1 "It's closer to the door because the door is slipping [pointing to a white cube with a sound of a long snort and a white cube with a sound of a tom drum]" P2 laughs [Identification sounds, 8] ● 37:42 [We hear a sound of a thunderstorm] P2 "The thunder wants to be aware I guess somewhere" [3,8] ● 38:04 P4 "I've just changed the fire" [in response to P2's question "What are you doing?"] ● [We hear a sound of a groan] 38:56 P4 "Can we change the groaning?" [laughs] ● 39:05 [We hear a sound of a cat howling] P2 "The cat is alright, the cat will do" ●

G3 NSP: 0 –

SP: 0 –

G4	NSP: 20	06:00 [We hear the sound of a kid laugh] P2 "That's a happy baby" P3 "happy baby" P2 "Shall we find a nice sound to go with the happy baby?" [3,8] ● 06:09 [We hear a sound of a thunderstorm] P3 "That's a thunder line that's not good" P2 "That goes with the scream of the baby, with the baby crying" [2,3,8] ● 06:21 [We hear a sound of a kid laugh] P2 "This laugh is a little evil" [2,3,8] ● 06:48 [We hear a sound of a purring cat] P3 "A purring cat" ● 06:56 [We hear a sound of a purring cat] P3 "A purring cat" P2 & P1 "It's a purring cat!" [in response to P4's comment "Oh that's the... crickets!"] P2 "Ohhhh [smile]" ● 07:07 [We hear a sound of a kid laugh] P2 "Does this baby have other things of voices?" [looking at P4] [1,3] ● 07:29 [We hear a sound of a cat howling] P2 "a kitty! a kitty!" ● [We hear a sound of a long snort] 07:32 P1 "...a snoring yeah" P2 "a snoring" ● 07:36 [We hear a sound of a cat purring] P2 "Who's got the cat?" Is that you? [looking at P3] [3,4] ● 07:46 P2 "There was a cat purring..." P3 "a cat purring" P2 "and then there is the cat 'miauing' it was pissed off" P3 "I don't know whether I have that one" P2 "somebody has that pissed off cat" 3,4,8] ● 09:08 P2 "Who's got the other voice? [referring to a sound of a groan]" [2,4]; 09:35 P2 "Yes, mine [confirming P1's comment about P2's white cube on everyday noises e.g. rustling paper, referring to the 'screeches' white cube] but I don't think the voice [referring to a sound of a groan] comes from my block" [1,3] ● 10:10 [We hear a sound of a car crash] P3 "That was mine" P2 "The glasses" [in response to P1's question "Are you? [looking at P3]" ● 10:52 [We hear a sound of a kid laugh] P4 "I should only laugh at the end (...) oh actually it is still laughing even though I lifted it up [we can see the shade of his white cube]" P3 "yeah I can see the reflection" P4 "Yes it is still reflecting, oh I see" [2,3,8] ● 15:59 [We hear a sound of a cat purring with an effect] P2 "This is the purring cat... in a dimension... this is like a purring panther" [3,6,8] ● 16:59 [We hear a sound of a scream] P2 "Who's been screaming?" P3 "Who's screaming?" P3 "Yes I don't know who, it's not me" [3,4] ● 17:22 [We hear a sound of a wind mouth gurgle] P2 "Who's got the nice wind?" [3,4] ● 18:32 [We hear a sound of a water splash] P2 "Wow somebody is splashing" [3,4] ● 20:18 [We hear a sound of a cat purring] P3 "That's the cat" P2 "What do you call it? like Shrek, what are they called?" P3 "Ogre" P2 "Ogre" [3,8] ● 17:19 [We hear a sound of a scream] P4 "That's this one" [while lifting a white cube up] P2 "Ah you were the scream [pointing to P4]" [2,3] ●
	SP: 5	22:09 [We hear a sound of a thunderstorm] P3 "A thunder again (...)" ● 22:18 We hear a sound of a cat purring] P3 and P2 "It's the cat!" P2 "The purring cat" ● 24:56 [We hear a rhythmic sound of a ding from a ceramic impact] P2 "That sounds like a supermarket (...) a cashier of a supermarket" [Identification of sounds, 8] ● 34:44 [We hear a sound of a long snort] P2 "No this is the cat snoring [showing a white cube she is holding], I've got the snoring" [in response to P3's comment that another white cube changes the sound of a cat purring to snoring] P3 "Have you?" P2 "I am making it really loud, look" [Identification sounds, 8] ● 34:55 [We hear a sound of a long snort] P2 "So there could be somebody laughing at me because I am snoring" [Identification sounds, 8] ●
G5	NSP: 3	24:33 [We hear a sound of a tibetan chant] P1 "It's a good one" P2 "Boring" P1 "The boring chant or the orthodoxy drowning" P2 "Heavy metal" P1 "No unfortunately one of the [inaudible] just can't sleep" P2 and P4 laugh [2,3,8] ● 38:46 [We hear a sound of a cat howling] P4 "That was a cat" ● 38:48 [We hear a sound of a cat howling, and then a sound of a kid laugh] P1 "There's something quite chilly about the cat, it sounds like a baby, ours it's quite a bit like that, especially when it wags" P2 "When I was a kid and used to hear a cat crying outside I used to take my dad it sounded like a baby on the doorstep making my dad look" P1 "True, it's absolutely chilly" [2,3,8] ●
	SP: 6	10:44 [We hear a sound of a leaking gas hiss] P3 "Who's doing this... gas? [laughs]" [3,4] ● 11:58 [We hear a sound of a man snoring with an effect] P1 "We've got a snore, who's got the snore?" [3,4] ● 12:05 [We hear a sound of a kid laugh] P2 "A baby" ● 13:26 P1 "I seem to have the thunder claps [while lifting up and down a white cube with a sound of a thunderstorm]" [3,8] ● 16:32 [We hear a sound of a water splash] P2 "That one [inaudible] splashing" ● [We hear a sound of a scream] 16:57 P2 "Screaming" P1 "Yeah bringing the ending planet" [Identification sounds, 8] ●
G6	NSP: 1	13:01 [We hear a sound of a cat howling] P4 "A cat" P1 "That's the cat again" ●
	SP: 0	–

G7	NSP: 15	26:13 [We hear a sound of a howl] P2 "There is a lion sample" ● 26:22 [We hear a sound of a long snort] P1 "Okay so this is the snoring" ● 27:25 [We hear a sound of a tom drum] P2 "A drum" ● 26:45 P2 "So we discover all the sounds of this one ['whistles' white cube] (…) The interesting one is snoring" [1,2,3] ● 27:40 [We hear a sound of rocks smashing] P2 "Working with boots or something" [3,8] ● 28:52 P2 [We hear a sound of a tom drum] "Maybe this is drum" ● 30:33 [We hear a sound of a cat howling] P4 "Oh kitty!" ● 30:37 [We hear a sound of a kid laugh] P2 "Ok this is the good one" P3 "I like the baby" [2,3] ● 30:46 [We hear a sound of a cat howling] P4 "I like every cat" [2,3] ● 32:48 [We hear a sound of a wind mouth gurgle] P2 "Ok this is the wind" ● 33:11 [We hear a sound of fire crackle] P3 "Fire" 33:15 P2 "Fire" ● 34:45 P4 "Okay baby, cat, laughing guy" [while listening to the 'voices' white cube] [1,3] ● 34:57 [We hear a sound of a kid laugh] P2 "Oh yes yes baby is also nice" [2,3] ● 37:25 P2 "So we have the cat, we have the shooting and…" P4 "We have like the sony baby" P3 "We have a creepy background" P2 "Exactly we have the [inaudible] baby" P3 "It's walking down a park" P2 "And maybe someone is sleeping and gets shot in the sleep [we hear a sound of a gun shot], you can also add the snoring one" P3 laughs P4 "Yes, so there's a shot as long as I know" P2 "This is the snoring [we hear a sound of a long snort]" P4 "Somewhere yeah" [we hear a sound of a groan] P2 "This is the shot" [we hear a sound of a gun shot] P3 "It's like [inaudible]" P2 "And then burn it" P2 "Ok I think we left them not very good" [2,3,8] ● 35:52 [We hear a sound of a kid laugh with an effect] P2 "Ok so this is the sound of the parent making jokes out about himself" P3 "It looks an old man laughing" All laugh P3 "Or maybe a bit obsessed" P4 "Sounds like an exorcist stuff" P2 "So maybe we can try if it is snoring to have some fun" [2,3,6,8] ●
	SP: 10	12:10 [We hear a sound of a groan] P2 "This is the snatch? [pointing to a white cube manipulated by P1]" P1 "Yes" ● 12:12 [We hear a sound of a cat howling and then a sound of a groan] P2 "This is the sleepy cat" [3,8] ● 12:32 [We hear a sound of a cat howling] P2 "I like this cat" [2,3] ● 15:34 [We hear a sound of a cat howling] P2 "This is the cat one [pointing to the 'voices' white cube], exactly" ● 17:32 [We hear a sound of a tibetan chant, before there was a sound of a cat howling] P3 "The cat is praying [laughs]" [3,8] ● 18:48 [We hear a sound of a crowd murmur with four black cubes affecting it] P4 "I reckon it's a crowd (…) together" [3,6] ● 18:52 [We hear a sound of a crowd murmur with four black cubes affecting it] P2 "Which one is a crowd?" 18:55 P4 "It seems it's this one there" [P4 and P3 point to the 'whistles' white cube, although the sound of a crowd murmur comes from the 'whispers' white cube] [3,4] ● 21:40 [We hear a sound of a kid laugh and a group whispering] P3 "This is my favourite sound I think" P2 "The baby?" P3 "This thing, the whispering [pointing to the 'whispers' white cube]" P2 "This is the scary one" [2,3] ● 22:47 P3 "This is the cat [we now hear a sound of a cat howling from the 'voices' white cube of magenta colour, manipulated by P3]" P2 "The pink one [pointing to P3's white cube]" P3 "Yeah" ● 23:17 [We hear a sound of a man laugh] P2 "Which one is laughing [pointing to P3's cubes and looking at her]" [3,4] ●
G8	NSP: 18	07:30 [We hear a sound of a cat howling] P3 "This is the cat" P2 "[inaudible] is the cat, pink is the cat I think, although are the same colour the same sound?" [3,4] ● 08:06 [We hear a sound of a cat howling] P3 "This is the cat" All laugh 08:07 P1 "Yeah" [2,3] ● 08:56 [We hear a sound of a group whispering] P1 "whispering, underground noises [laughs]"[3,8] ● 09:06 P3 [We hear a sound of a crowd murmur] "If we rotate it? [to P2]" P2 "Crowds" P4 "It just changes the volume" [2,3] ● 09:32 [We hear a sound of a fire crackle] P3 "Fire" ● 12:35 [We hear a kid laugh] P2 "Another baby laughing" ● 14:10 [P3 is lifting up and down a white cube with a sound of a tom drum] P2 "We've got drums" ● 15:26 P2 "Let's turn the baby down [while turning down the volume of the sound of the kid laugh]" P3 "It's enough baby [laughs]" ● 11:12 [We hear a sound of a boom rumble, and a car crash] P3 "Is that a thunderstorm?" [we hear a sound of a gun shot] P1 "Yeah or the shot & the thunderstorm" P3 "It's like things happening" [3,5,8] ● 16:24 [We hear a sound of a kid laugh with an effect] P1 "Can't believe it's a baby" [3,6] ● 16:35 [We hear a sound of a thunderstorm with an effect] P2 "Ok this is the thunder one" ● 16:43 P3 "Is that the cat one? [moving a black cube next to a white cube with a sound of a cat howling]" [3,4] ● 16:54 P2 "Stormy sound, I think it's this one [pointing to the 'explosions' white cube with a sound of a thunder storm]"; 17:35 P1 "Fire" [in response to P3's question "What's that sound?"] ● 19:41 [We hear a sound of a water splash with an effect] P2 "There's splashing always" [lifting it up once and then down again] P2 "Maybe that's better that splashing" ● 20:27 [We hear a sound of a long snort] P3 "So which is the snoring" P2 "That?" [3,4] ● 21:45 [We hear a sound of a kid laugh] P3 laughs P3 "A baby.." ●

SP: 7		23:30 P2 [We hear a sound of a crowd murmur] "Ok that's the crowd noise, it's got a good one actually" [2,3] ● 28:37 [We hear a sound of a man laugh] P3 "He's just laughing at us" [Identification sounds, 8] ● 29:13 [We hear a sound of a cat howling and an effect] P3 "Yes something else in there" [in response to P3's question "There is a second voice?"] P1 "It's like, so they are two cats now" [P3 adds another black cube next to a white cube with a sound of a cat howling] P1 "3 cats?" P1 "Wow, 3 cats!" P4 smiles P2 "More" P1 "4 cats" P3 laughs [3,6,8] ● 30:53 [We hear a sound of a scream] P1 "Screaming" P2 & P4 laugh [2,3] ● 31:29 [We hear a groan with an effect] P1 & P3 laugh 31:42 P2 "We've got a man groaning" P2 "It's a little bit..." P3 "...Spooky" [2,3,6] ● 35:17 [We hear a sound of a long snort] P2 "That's the snore" ● 37:38 [Silence] P3 "Silence, oh that's fantastic" [2,3] ●

D.8.4 Non-identification of person or object

TABLE D.7: Transcripts of *Non-identification of person or object*

Group	Condition	Quotes
G1	NSP: 2	24:46 [We hear a sound of a groan] P3 "What's that?" ● 31:34 P3 [We hear a sound of a long snort] "Is this this noise?" ●
	SP: 4	11:28 [We hear a sound of a gun shot] P2 "Who got the gunshot?" [3,4] ● 11:52 P2 "Who is the [inaudible]?" ● 17:07 [We hear a sound of a war boom explosion] P1 "What is that one?" ● 09:48 [We hear a sound of a cat howling] P4 "Who's got the cat one?" [3,4] ●
G2	NSP: 11	15:51 P1 "Should try enhancing the wind" P2 "Yeah I can't hear this wind at all [while lifting up a white cube with a sound of a wind mouth gurgle] [3,4] ● 16:13 P2 "I don't know what this one is doing now, this one might go to sleep [while moving a white cube to her next rim area]" ● 16:40 [We hear a sound of a lamp buzz] P1 "Which sound is which now?" ● 17:21 P2 "What is that one?" ● 17:54 P2 "What's that one then?" ● 18:58 P4 "Which one?" ● 20:25 P2 "Is that keep making a ding?" [pointing to a white cube operated by P1 that was producing a a sound of a metallic impact a few seconds before] ● 21:11 P2 "Oh who's got that one?" ● 21:30 P2 "Is that your snoring?" [asking to P1] [3,4] ● 23:19 [We hear a sound of a tunnel short whistle] P2 "Who is that one? (...) ● 23:26 P2 Is that your fluty noise?" [pointing to a white cube next to P1] P1 lifts up another white cube and says "This one" [although it is a different sound than the fluty noise] ●
	SP: 8	25:58 P4 "What is this supposed to do?" ● 26:23 P2 "Is it? I can't actually hear what it is" [replying to P4's comment "That's a cat"] ● 32:42 P2 "I can't hear what that is" ● 33:25 [We hear a cat purring] P2 "Is that the cat purring?" [pointing to a 'wrong' white cube] [3,4] ● 34:52 P4 lifts up a white cube with a sound of a tunnel short whistle and says "No I thought it was the baby" lifting down the object ● 35:04 P2 "I can't really tell what these sounds are (...) what that is?" ● 38:04 P2 "What are you doing?" ● 38:08 P2 "What did you get then?" P4 "(...) I don't know" ●
G3	NSP: 0	–
	SP: 0	–

G4	NSP: 21	06:50 [We hear a sound of a cat purring] P2 "Is that yours?" [looking at P3 after P3 saying "purring cat"] P3 "Not sure" ● 07:12 [We hear a sound of a groan] [laughs] P3 "I don't know what that was" ● 07:25 P3 "so yours...?" ● 07:30 P3 "Have you got..? [looking at P2] ● 07:33 P3 "I've got..." ● 07:55 P2 "Uh because you've got..." ● 08:13 [We hear a sound of a groan] P2 "Who's got that?" [laughs] ● 07:36 [We hear a sound of a cat purring] P2 "Who's got the cat?" Is that you? [looking at P3] [3,4] ● 07:46 P2 "There was a cat purring..." P3 "a cat purring" P2 "and then there is the cat 'miauing' it was pissed off" P3 "I don't know whether I have that one" P2 "somebody has that pissed off cat" 3,4,8] ● [We hear a sound of a groan] 08:58 P2 "Who's got this thing? This voice? This lament?" [2,4,8] ● 09:08 P2 "Who's got the other voice? [referring to a sound of a groan]" [2,4] ● 09:11 P3 "It's you!" [in response to P2's question "Who's got the other voice?", although referring to a sound actually operated by another team member] ● 09:12 P2 "Me? Oh my gosh, I do apologize, it is me! [laughs] I had a different sound before" [actually another person is operating the groan sound P2 is attributing to herself] ● 10:09 P1 [We hear a sound of a car crash] "Are you? [looking at P3]" ● 16:59 [We hear a sound of a scream] P2 "Who's been screaming?" P3 "Who's screaming?" P3 "Yes I don't know who, it's not me" [3,4] ● 17:22 [We hear a sound of a wind mouth gurgle] P2 "Who's got the nice wind?" [3,4] ● 17:36 [we hear a sound of a tunnel short whistle] P3 "Uh is that mine?" P1 "Yes I can see the combination of both" P2 "No, that's me" ● 18:22 P4 "Who's like more than one person?" ● 18:25 P2 "What was the...?" ● 20:18 P2 "Where is the...¿' ● 18:32 [We hear a sound of a water splash] P2 "Wow somebody is splashing" [3,4] ●
	SP: 8	21:23 P2 "Who's got the one?" P3 "I don't know" ● 22:16 P3 "I wonder what this one does then" ● 34:24 [We hear a sound of a long snort] P2 "Who is got this noise?" [2,4] ● 34:36 P3 "This makes that [pointing to a white cube on the table], whatever that is, this is not the cat purring, I think this turns the cat purring into snoring" [actually P2 points out that herself is holding the sound of snoring that P3 is attributing to another white cube] [4,8] ● 34:51 P3 "What else do we have?" ● 37:31 P3 "Who's delighted that one?" ● 38:56 P2 "This is me?" ● 39:09 P2 "Who's got this?" ●
G5	NSP: 3	[We hear a sound of a tibetan chant] 27:13 P1 "That drone is the winner" P4 "Yeah" P1 "Are you making that drone?" P4 "What?" P2 laughs P1 "Are you making the drone yourself?" P4 "This one [pointing to a white cube]" P1 "I think you need to mark somehow that box" P2 "Yeah" P1 "That's the winner" [2,4,8] ● 24:27 P2 "Different sounds, isn't? [in this session compared to the previous session] " ● 32:28 P2 "Who's this? (...) Uhhh" ●
	SP: 2	10:44 [We hear a sound of a leaking gas hiss] P3 "Who's doing this... gas? [laughs]" [3,4] ● 11:58 [We hear a sound of a man snoring with an effect] P1 "We've got a snore, who's got the snore?" [3,4] ●
G6	NSP: 1	10:43 [We hear a sound of a groan] P2 "Is that you? [looking at P3]" ●
	SP: 1	39:03 [We hear a sound of a snoring] P2 "Who's this?" ●
G7	NSP: 3	27:36 [We hear a sound of a falling metal, a sound of rocks smashing] P1 "What is this?" P4 says no with her head P3 "A tsunami [inaudible]" [4,5,8] 32:07 [We hear a sound of a stomach gurgling] P2 "What is this?" ['whispers' white cube] ●35:40 [We hear a sound of a kid laugh with an effect] P2 "What is the second sound?" ● 40:20 P2 "What is this? Yellow one?" ●
	SP: 5	19:21 [P1 is manipulating the 'explosions' white cube and we hear a sound of a car crash] P3 "Who's like the vampire car?" P2 "It's probably..." 19:25 P4 "It might be someone with [inaudible] anything that fill over [inaudible] cause it does steeping" P2 "I prefer the curtain" [4,5,8] ● SP: 13:55 [We hear a sound of rocks smashing from the yellow white cube] P3 "Orange?" 13:59 P2 "The pink one or something" ● 18:52 [We hear a sound of a crowd murmur with four black cubes affecting it] P2 "Which one is a crowd?" 18:55 P4 "It seems it's this one there" [P4 and P3 point to the 'whispers' white cube, although the sound of a crowd murmur comes from the 'whispers' white cube] [3,4] ● 22:41 [We hear a sound of a cat howling from the 'voices' white cube of magenta colour] P2 "This is the cat I think the green one [pointing to the 'whistles' cube of green colour]" ● 23:17 [We hear a sound of a man laugh] P2 "Which one is laughing [pointing to P3's cubes and looking at her]" [3,4] ●
G8	NSP: 4	07:30 [We hear a sound of a cat howling] P3 "This is the cat" P2 "[inaudible] is the cat, pink is the cat I think, although are the same colour the same sound?" [3,4] ● 16:43 P3 "Is that the cat one? [moving a black cube next to a white cube with a sound of a cat howling]" [3,4] ● 20:27 [We hear a sound of a long snort] P3 "So which is the snoring" P2 "That?" [3,4] ● 22:10 [We hear a sound of a tibetan chant] P2 "Who's got that didgeridoo sort of noise?" [4,5] ●

| | SP: 2 | 26:13 P3 "Is that this one then? [lifting up and down once the 'percussion' white cube with a sound of a falling metal]" ● 31:13 P3 "Is that the one [white cube]? [pointing to a white cube next to P2 and adding a black cube next to it]... Or is this the one [pointing to another white cube]?" ● |

D.8.5 Non-identification of source

TABLE D.8: Transcripts of *Non-identification of source*

Group	Condition	Quotes
G1	NSP: 2	24:23 [We hear a sound of a man scream] P2 "Terrifying pig" [5,8] ● 28:30 [We hear a sound of a water splash, a scream, and a tunnel short whistle] P2 "That's crackles" P1 "Yes" ●
	SP: 1	[We hear a sound of a cat howling] 09:45 P3 "Little cow or something" [5,8] ●
G2	NSP: 1	19:16 [We hear a sound of a metallic whine] P2 "That's not the wind, is it? that's more like the airplane" [5,8] ●
	SP: 10	26:06 [We hear a sound of a stomach gurgling] P4 "Uhh" P3 "A cat or a lion?" ● 26:35 [P4 comments there is a sound of a cat] P2 "I can't hear the cat" [turning her head towards her left back speaker] ● 29:03 [We hear a sound of a ceramic impact] P2 "Is that the gong?" ● [We hear a sound of a tom drum] 29:13 P2 "Like a door banging or something, over the wind" P1 "it happens every once in a while" [Identification sound, 8] ● 30:43 [We hear a sound of a cat howling] P2 "Is that a baby?" P4 "Not sure" ● 33:01 [We can't hear a sound of snoring] P2 "Or is this snoring the best we are gonna get?" ● 32:19 P4 [We hear a sound of air escaping hiss] "some static" ● 32:26 [We hear a sound of air escaping hiss] P1 "It's like a washing-machine or a dishwasher" [5,8] ● 34:59 [We hear a sound of a tom drum] P2 "That's the batting door" [5,8] ● 38:27 [We hear a sound of rustling plastic] P4 "Uh this is the paper" P1 "Paper?" ●
G3	NSP: 0	–
	SP: 0	–
G4	NSP: 5	[We hear a sound of a cat purring] 06:56 P4 "Oh that's the... crickets!" ● [We hear a sound of a tibetan chant] 07:20 P1 "That's didgeridoo" ● 09:28 [We hear a sound of plastic rustling] P1 "so your block makes all kind of everyday noises [pointing to P2's 'screeches' white cube], rustling paper and all kinds of these" [1,5] ● 09:24 [We hear plastic rustling] P2 "again!" P1 "What's that sound...[looking at P2] That sound is coming from your block [pointing to P2's white cube]" [2,5] ● 19:05 [We hear a sound of a tibetan chant] P2 "Who's got this? ... you? [pointing to P4]" P4 "This one yeah" P2 "Oh cool" P1 "Didgeridoo" P3 "[inaudible]" P2 "I love that one" [2,5] ●
	SP: 2	21:38 P1 "This is this horn [while lifting up and down a white cube with a sound of a tibetan chant]" P2 "Nice!" P3 "A great sound" [2,5] ● 36:32 P4 "That's the sound of fire right?" P1 & P3 "No that's the modifier" ●
G5	NSP: 0	–
	SP: 0	–
G6	NSP: 0	–
	SP: 0	–
G7	NSP: 4	27:36 [We hear a sound of a falling metal, a sound of rocks smashing] P1 "What is this?" P4 says no with her head P3 "A tsunami [inaudible]" [4,5,8] ● 27:44 [We hear a sound of rocks smashing] P3 "A fire" ● 28:54 P2 [We hear a sound of rocks smashing] "This is wood work" ● 24:08 P2 "This is explosive sounds" [moving instead the 'whistles' white cube next to him] ●
	SP: 3	19:21 [P1 is manipulating the 'explosions' white cube and we hear a sound of a car crash] P3 "Who's like the vampire car?" P2 "It's probably..." 19:25 P4 "It might be someone with [inaudible] anything that fill over [inaudible] cause it does steeping" P2 "I prefer the curtain" [4,5,8] ● 20:36 P2 "What is the sound of this one? [pointing to the 'explosions' white cube manipulated by P1 with a sound of a boom rumble]" P1 lifts up the 'explosions' white cube, P2 "Maybe you can put it [inaudible]" ● 20:57 [We hear a sound of a long snort] P2 "Ok this is the I think sleeping dog [pointing to the 'whispers' white cube with a sound of a long snort]"[5,8] ●

G8	NSP: 6	08:46 [We hear a sound of a cat purring] P1 "Tiger or something?" P4 "Or some-one's stomach" P1 & P3 laugh P1 "Yeah I agree" • 09:50 [We hear a sound of a cat howling] P2 laughs P3 "What's the other one then? [pointing to the other white cube that P2 is holding]" [Impression sounds, 5] • 10:52 [We hear a sound of a boat whistle] P2 "[Pinky] noise (...) or the cat purring" • 11:12 [We hear a sound of a boom rumble, and a car crash] P3 "Is that a thunderstorm?" [we hear a sound of a gun shot] P1 "Yeah or the shot & the thunderstorm" P3 "It's like things happening" [3,5,8] • 17:26 [We hear a sound of a group whispering, a tibetan chant, and a fire] P3 "What's that sound?" • 22:10 [We hear a sound of a tibetan chant] P2 "Who's got that didgeridoo sort of noise?" [4,5] •
	SP: 1	35:09 P2 "I can't work out what sound that is, I don't know [while manipulating the 'whispers' white cube by lifting it up and down]" •

D.8.6 Identification of filters

TABLE D.9: Transcripts of *Identification of filters*

Group	Condition	Quotes
G1	NSP: 4	23:56 P4 "you combine 2" P2 "2 modi" P4 "2 modifiers" • 29:11 P3 "loud, too noisy, this one" P2 "that was the black one" • 30:03 P4 "leave it with the modi-fier" • 30:51 P2 "Yours black one is not going to make any noises (...) you have to connect it to a white one to make noise" P3 "What sort of noise?" P2 "I think is the same noise as the white cube, which changes the noise saturance" P4 "It's like a transition" [6,7] •
	SP: 3	07:33 P1 "[inaudible] black one is pink, and now yellow [inaudible]" [in re-sponse to P2's question "I wonder what the black ones do?"] • 14:12 P4 "Take everything off and see what it does" P1 "Don't do anything on its own, I think" P4 "Put a white one and then with that" [All the objects are removed, P4 tries one white cube and one black cube, changing to different faces of the cube] P4 "Here is slight different, different place level there is I think" P3 "Yes" P2 "It's hard to hear" P3 "Yes it does" P4 "See?" [while changing to another face of the black cube] P4 "Different parts sound different (...) here the splash is louder" • 19:06 P3 seems to be discovering the role of the modifier by dragging it to different sound players positioned in the middle •
G2	NSP: 5	P4 08:58 "That's the amplifier" • 10:33 [P4 is manipulating a black cube next to a white cube] P4 "There was a change somewhere there" • 11:08 [P4 is manipu-lating a black cube next to a white cube] P4 "But here it becomes red ... uh ok" • 11:49 [P4 is manipulating a black cube next to a white cube lifting it up and down] P1 "it is increasing the volume" P4 "Yeah" • 14:58 [P2 is manipulating a black cube next to a white cube with a sound of a fire crackle] P1 "It's making it louder a bit" •
	SP: 2	33:56 [P1 positions a black cube next to a white cube with a sound of a man snoring] P2 "That's really loud" [P1 reduces the volume of the white cube using a black cube] P1 "different size of the filter" P4 "good good, an echo" • 34:03 P1 "There's an echo" P3 "(...) interesting" P2 "That's too much (...) That's a bit much that snoring (...) Wasn't that loud before" •
G3	NSP: 0	–
	SP: 0	–

G4	NSP: 15	08:15 P2 and P3 "It changes" P1 "It changes the sound in some way" [in response to P3's question "What does this black does?"]; 08:40 P4 "It didn't change in mine" [black cube on white cube] ● 08:43 P1 "No, it is based on the proximity to the other blocks it attaches to the closest, so this filter is also applying to that sound" ● 11:14 P2 "I have no interest in the modifiers" ● 11:20 P3 "That's the volume [while approaching a black cube next to a white cube with a sound of a pipe hiss]" ● 11:25 P2 "There is a delay" ● 11:33 [P4 is dragging a black cube to a white cube] P4 "Oh the colour changed to red" P3 "Oh yes" P4 "And this is red I believe (...) Oh it's amplified" P3 "Uhhhh" P4 "Wow (...) It's echoing" ● 12:09 [P4 is dragging a black cube next to a white cube with a sound of a baby crying] P3 "Uh that's very spooky" P4 "It's a bit spooky" [laughs] P3 "[inaudible] an echo" ● 12:23 P2 "This doesn't seem to do very much" P1 "She said two of the modifiers are band pass filters so they are not going to change much apart from the high pass filter or low pass filter but he has this kind of reverb" P2 "That's nice" ● 13:46 P4 "[inaudible] What the orange filter?" P3 "I think every white block has its own colour (...) and the shadow underneath says which block is attached to" P2 "Yeah now it's attached to mine, can you do amplify mine?" ● 14:02 P4 "This one" [in response to P2's question "Which one is your 'amplifier'?"] P2 "That one?" P4 "That is the tikitikitiki" [pointing to P2's sound] P4 "That's the amplifier" P2 "Uhhh nice!" P3 "it's also the loop into it" P2 "Oh nice! This goes nice together" [2,6,7] ● 14:44 [We hear the effect of a black cube next to a white cube] P2 "Gosh" P3 "Dr. Who" [6,8] ● 15:10 P1 "Yeah they are both [black cubes] attached to this one [white cube]" ● 15:59 [We hear a sound of a cat purring with an effect] P2 "This is the purring cat... in a dimension... this is like a purring panther" [3,6,8] ● 17:30 P4 "This modifier by itself doesn't make any sound"●
	SP: 6	22:54 P3 [inaudible] reverb [putting a black cube next to a white cube] ● 23:08 P2 "Look I try to steal the amplifiers, the modifiers [laughs], they don't stick to me (...) Ok let's engage with yours and it wouldn't" ● 23:29 P1 "This combination is quite cool" [a black cube and a white cube] ● 28:15 P2 "The modifiers have become white because..." P4 "So the other white cube doesn't change" P2 "No, this modifier is become yellow like mine it's nicely amplifying the thunder" ● P1 "It's that block I think" [in response to P3's question "What sort of noise are you making, have you got an amplifier?"] ● 36:34 P1 and P3 "No that's the modifier" [in response to P4's question "That's the sound of fire right?"] ●
G5	NSP: 4	31:06 P2 "Maybe we can try different things, filters" P1 nods his head and adds a black cube next to a white cube in the centre 31:18 P3 "It's good" ● 33:15 [P3 adds a black cube next to a white cube that increases considerably the volume of a sound of a tibetan chant] P3 "Sorry!" P1 smiles ● 34:14 [P4 is lifting up and down rhythmically a white cube with a sound of a tunnel short whistle with an effect while P1 is nodding his head] P4 "It seems to pick it up you know? [pointing to a set of black cubes]" [2,6] ● 35:19 [We hear a sound of a cat howling with an effect] P2 laughes P3 "That's good" P2 & P3 are laughing P4 "Not sure if it is a baby or a cat" P3 "It's different" [2,6] ●
	SP: 6	12:10 P1 [We hear a sound of a man snoring with an effect] "There's something happening there" [after asking who's got the snore and P2 pointing to the white cube with a sound of a man snoring that is next to a black cube] P2 "It's like it picks up the rhythm by itself" [2,6] ● 12:38 P1 "That's very nice wood blocking for the sound" ● 16:08 P1 "There are too many pictures, are there? they are mostly, fairly sound effected" P1 "I like that [nodding his head]" [6,9] ● 18:33 P4 "There is sort of a delay here" [in response to P2's question "What shall these black ones do?"] P2 "Uh ok" P4 "That's the only one I think" P1 "Yeah I think there is an extra reverb, sort of a..." ● 18:54 P2 "This one [effect of a black cube] seems to be round" P2 "slight delayed" P2 "Ahh an echo as well" ● 20:49 [We hear a sound of a water splash with an effect] P1 "I like that [while turning a white cube left and right next to a black cube and nodding his head rhythmically]" P1 & P2 are nodding their heads [2,6,9] ●
G6	NSP: 5	09:26 [We hear a sound of a kid laugh with an effect] P3 "That's creepy" P1 & P2 laugh [2,6] ● 13:35 P2 "Wow, good [while turning left and right a black cube next to a white cube]" [6,9] ● 15:10 P1 & P3 smile with the sound of a kid laughing with an effect [2,6] ● 21:57 P2 "Now let's delay" ● 22:28 P2 "Here is delay, look" ●
	SP: 3	26:06 P3 "So this is the delay" P2 "The 3 is the delay I think" ● 27:30 P2 "That's the... pitch shifter!" P3 "Aha" P2 "What number is it?" P3 "The 4" ● 32:33 P3 You can try to automatic cutoff by having an extreme position P2 "Aha (...) okay I see" [we hear the result of the effect] P2 "Nice idea! (...)" P1 laughs P3 "Yeah" P2 "Flattering between the..." ●

G7	NSP: 13	26:28 [P3 adds a black cube next to a white cube with a sound of a gas hiss] P2 "This just increases the sound, maybe another one [and adds another black cube next to the white cube]" • 27:56 P2 "It's probably this one increases the frequency of all of the samples [pointing to a black cube]" • 28:06 P2 "Yes another black maybe" P4 "The black" • 29:48 P2 "Maybe another filter?" • 30:52 [We hear a sound of a cat howling with an effect] P2 "Maybe we can increase the echo" • 31:10 [We hear a sound of a cat howling with an effect] P2 "Maybe we can even increase the echo, many cats [laughs]" • 31:26 [We hear a sound of a scream with effect, and the sound decreases all of the sudden] P3 "Are you trying to pulling down? [while rotating the white cube with a sound of a scream that increases the volume]" • 32:21 [We hear a sound of a stomach gurgling with an effect] P2 "What is this? something under water I think" [2,6] • 32:40 P3 "This one is black magic" [laughs] • 35:07 P2 "Cat has the echo filter but baby didn't" • 35:19 P4 "I think so" [in response to P4's question of whether the filter turned on"]; 35:46 [We hear a sound of a kid laugh with an effect] P3 "That's freaky" P2 & P4 laughs [2,6] • 35:52 [We hear a sound of a kid laugh with an effect] P2 "Ok so this is the sound of the parent making jokes out about himself" P3 "It looks an old man laughing" All laugh P3 "Or maybe a bit obsessed" P4 "Sounds like an exorcist stuff" P2 "So maybe we can try if it is snoring to have some fun" [2,3,6,8] •
	SP: 9	12:37 [After saying "I like this cat", P2 moves a black cube next to a white cube with a sound of a cat howling] P2 "With echo" P2 & P4 laugh • 14:54 P1 "This one [pointing to a black cube]" 14:55 P1 "[I think you will have fun] [while passing a black cube to P3, who was asking what filters do] with this [pointing to the black cube]" • 17:07 [We hear a sound of a cat purring with an effect that increases the volume] P2 "It increased the voice" • 18:11 [We hear a sound of a crowd murmur with four black cubes affecting it] P2 "This is more scaring" P3 "Yeah [laughs]" [2,6] • 18:48 [We hear a sound of a crowd murmur with four black cubes affecting it] P4 "I reckon it's a crowd (...) together" [3,6] • 20:39 P3 "I think the black ones are affecting the delay [pointing to two black cubes next to the 'voices' cube]" 20:42 P2 "Sorry?" P3 "The black ones are affecting the delay" • 21:09 [We hear a sound of a kid laugh from the 'voices' white cube with four black cubes next to the white cube, manipulated by P3] P2 "Let's keep it like this [laughs looking at P3]" [2,6] • 22:15 [We hear a sound of kid laugh with an effect] P4 "Is this the echo?" P3 "Creepy" P2 "This is the echo [pointing to a black cube]" [2,6] • 22:30 [We hear a sound of a kid laugh with an effect, and a gas hiss with an effect] P4 "[laughs] This is so creepy" [2,6] •
G8	NSP: 8	15:09 [We hear a sound of a plastic rustling with an effect] P3 "[inaudible] very loud" P2 "[inaudible] it's too very loud" [Impression sounds, Identification filter] • 15:29 P3 "Now we should turn it loud again [while rotating the white cube with a sound of plastic rustling and then a black cube that increases the volume of the sound of a plastic rustling]" [2,6] • 16:12 P2 "Yeah you can get a different sound [while rotating a black cube next to a white cube]" • 16:24 [We hear a sound of a kid laugh with an effect] P1 "Can't believe it's a baby" [3,6] • 16:27 [We hear a sound of a man laugh with an effect, before it was a kid laugh from the same cube] P1 "This is the other one [laughs]" [2,6] • 19:04 P3 "Is that [black cube] duplicated? We've got that twice isn't?" P3 "This one [pointing to a black cube that P4 is holding], try that duplicated" [P4 positions a black cube next to the 'percussion' white cube with a sound of a ceramic impact] P3 "Hear the difference, it is easy to hear the difference [while lifting up and down the 'percussion' white cube with a sound of a ceramic impact, and positioning it next to different black cubes]" • 20:06 P3 "That's affecting it [pointing to a black cube next to the 'explosions' white cube with a sound of a water splash]" • 20:54 [All drag a black object next to a white object with a sound of a long snort] P3 laughs P4 laughs P2 laughs [All are affecting the snoring sound with a black cube each] •

| | SP: 7 | 23:16 P2 "P1-P4 are modifying the sounds slightly differently [while manipulating a black cube next to a white cube]" ● 23:49 [P2 & P4 are lifting up and down a black cube each next to a white cube with a sound of a crowd murmur] P2 "That's different!" [P3 also tries to lift up and down a black cube next to a white cube with a sound of a crowd murmur] P3 "That's better [pointing to P2's black cube]" ● 24:08 [P2 lifts down his black cube] P2 "That one is [difference]" ● 24:19 P4 "Uh! There is a difference! [while lifting up and down again his black cube next to a white cube with a sound of a crowd murmur]" P2 "I can reckon sort of a [difference]" ● 28:43 [We hear a sound of a man laugh with an effect] P3 "That makes a difference" P1 there is a second voice? P1 "There is a second voice" [2,6] ● 29:13 [We hear a sound of a cat howling and an effect] P3 "Yes something else in there" [in response to P3's question "There is a second voice?"] P1 "It's like, so they are two cats now" [P3 adds another black cube next to a white cube with a sound of a cat howling] P1 "3 cats?" P1 "Wow, 3 cats!" P4 smiles P2 "More" P1 "4 cats" P3 laughs [3,6,8] ● 31:29 [We hear a groan with an effect] P1 & P3 laugh 31:42 P2 "We've got a man groaning" P2 "It's a little bit..." P3 "...Spooky" [2,3,6] ● |

D.8.7 Non-identification of filters

TABLE D.10: Transcripts of *Non-identification of filters*

Group	Condition	Quotes
G1	NSP: 1	30:51 P2 "Yours black one is not going to make any noises (...) you have to connect it to a white one to make noise" P3 "What sort of noise?" P2 "I think is the same noise as the white cube, which changes the noise saturance" P4 "It's like a transition" [6,7] ●
	SP: 5	07:31 P2 "I wonder what the black ones do?" ● 08:36 P2 "What the black ones do? I can't get [it]" ● 08:56 [P1 is adding a first object in the middle of the table] P2 "That's not a sound" ● 13:03 [P4 is manipulating a black cube next to a white cube lifting it up and down] P4 "I don't hear difference" ● 13:34 [P1 and P4 are manipulating a black cube each next to a white cube lifting them up and down] P2 "I can't hear any difference whether there is one or two" ●
G2	NSP: 6	09:44 [P4 is manipulating a black cube next to a white cube] P4 "I just want to know what does it change... however I don't see any change... I don't hear" ● 10:01 P2 [P4 is manipulating a black cube next to a white cube] "I don't think it matters which size you put it on, it's just putting it near" P4 "I think she [the facilitator] said the colour should change it" ● 14:56 [P2 is manipulating a black cube next to a white cube with a sound of a fire crackle] P4 "What's the change?" ● 16:00 P2 "I am not too sure about these enhancing things [the black cubes] whether they do... they must do something or just increasing the volume but... it's quite subtle if they are" ● 20:49 P2 "What these modifiers things are supposed to do? what they actually do?" ● 2:04 P2 "I don't really know what are these for" [black cubes] ●
	SP: 1	29:37 P1 "Let's try a filter with it, and see what it does" [P3 passes a black cube to P1] P2 "Take it off again (...) [P1 moves a black cube that was next to a white cube] I can't really tell (...) perhaps it is more precise on the bang?" ●
G3	NSP: 0	–
	SP: 0	–
G4	NSP: 4	08:19 P3 "What does this black do?" ● 14:02 P4 "This one" [in response to P2's question "Which one is your 'amplifier'?"] P2 "That one?" P4 "That is the tikitikitiki" [pointing to P2's sound] P4 "That's the amplifier" P2 "Uhhh nice!" P3 "it's also the loop into it" P2 "Oh nice! This goes nice together" [2,6,7] ● 13:48 P2 "Which one is your 'amplifier'?" ● 14:32 P4" What does this one do?" [black cube] [inaudible] ●
	SP: 2	27:19 P2 "Is there an amplifier to take which is unemployed?" [P4 passes a black cube to P2] ● 37:34 P3 "What sort of noise are you making, have you got an amplifier?" ●
G5	NSP: 0	–
	SP: 1	18:22 P2 "What shall these black ones do?" P4 "Filters" P1 "It's supposed to subtly alter sound right?" P2 "They don't seem to add much" P1 "Not very much, does it?" ●

G6	NSP: 2	11:05 P2 "I can't modify it, the modification doesn't seem to work" ● 13:08 P2 "I can't figure out the effect" P3 "Yes, it's very subtle and..." ●
	SP: 0	–
G7	NSP: 6	28:22 P4 "I can't tell what the filters are doing" ● 28:27 P2 "Perhaps each of those blacks... no I don't think so" ● 28:31 P4 "Can you [white cube] pick up all the filters or can only pick up one of the same colour?" ● 34:58 P2 "Can we can echo baby sound?" P4 "I am not sure if there is echo filter so much as this one of the cats that has an echo effect on it" ● 35:13 P4 "It turned on the filter is that?" P3 "This sort of filter has a different size" ● 38:58 P4 "I wonder whether the black one is doing anything" P2 "This one [pointing to the explosions white cube] didn't do anything" P4 "...It's a little bit quiet ●
	SP: 4	11:32 P1 "I cannot see what the filters I don't know what the black cubes..." P3 "I don't know either" ● 14:00 P4 "I don't understand the filters" ● 14:43 P3 "For me the filters they don't seem to do anything [while manipulating a black cube next to a white cube with a sound of a ceramic impact]" ● 18:28 P2 "So if you put them close together... just one... maybe if you put everything near this one... sound didn't change but the colour changed" [trying now a white cube with a sound of a crowd murmur and four black cubes next to it] ●
G8	NSP: 4	10:26 P1 "This is the filter? [while P3 is manipulating a black cube next to a white cube at P2's area]" P2 "It's quite subtle isn't?" ● 14:56 P3 "Can you [inaudible] make it louder to make it really go out?" [We hear a sound of plastic rustling very loud because it has a black cube next to it] P1 "Loud" P2 "I think [inaudible] it is anti-clockwise to turn it down" [although P2 is rotating a black cube that is not affecting the loud sound of plastic rustling] ● 15:55 P3 "I wonder what these blacks do" P2 "It's supposed to modify the sound, it's quite subtle isn't?" ● 20:35 P3 "Can we affect the snoring? [dragging a black cube next to a white cube with a sound of a long snort]" ●
	SP: 8	23:01 P3 "I think cause we can't work out what the black ones do and whether it matters where we put one or the other, does it matter?" 23:08 P2 "No it's very subtle" ● 23:20 P1 "That black one is the same than that black one?" P3 "Shall we try it?" P1 "Shall we try this one?" ● 23:35 P1 "I can't understand the difference [the effect of black cubes]" ● 23:42 [P2 manipulates a black cube next to a white cube with a sound of a crowd murmur] P3 "That's [neutral]" ● 23:56 P4 "Or maybe is this [pointing to his black cube while manipulating it]" 24:05 [P2 lifts up his black cube] P2 "It doesn't seem to make any difference to me" ● 24:44 [P2 adds a new white cube, the 'explosions' cube, with a sound of a boom rumble next to a black cube and the 'whispers' white cube] P3 "Let's do this one [the new white cube], let's take this one out [moving to the rim area the 'whispers' white cube with a sound of a crowd murmur]" [Each of the participants puts a black cube next to the white cube with a sound of a boom rumble] P3 "Quiet quiet" P3 "It's that everything down there?" P2 "It's possible" P2 "Ooops" P3 "Maybe it does... maybe those things up" [2,7] ● 26:29 P3 "If there are too many blocks it just gets too confused [while manipulating a black cube alternating between two white cubes]" ● 28:15 [We hear a sound of a man laugh] P3 "So have you put that... does it make a difference [while lifting down a black cube next to a white cube with a sound of a man laugh]?" P3 "Does it slight...? (...) Is that link different?" [while lifting up the black cube]" P1 "I don't see any difference [saying no with her head]" P2 "Maybe add an over line [while lifting down a black cube next to a white cube with a sound of a man laugh, and then saying no with his head]" ●

D.8.8 Stories, realism and ambisonics

TABLE D.11: Transcripts of *Stories, realism and ambisonics*

Group	Condition	Quotes
G1	NSP: 2	[8.1]:2 – [8.1] 24:23 [We hear a sound of a man scream] P2 "Terrifying pig" [5,8] ● [8.1] 31:10 [We hear a sound of a group whispering] P4 "The market" [2,8] ●

	SP: 13	[8.1]:11, [8.5]:2 – [8.1] 06:29 [We hear a sound of a ceramic impact] P1 "You got the bell" [looking at P2] [3,8] ● [8.5] 09:03 P2 "[inaudible] something is slightly different" [turning her head towards her back] ● [8.1] [We hear a sound of a cat howling] 09:45 P3 "Little cow or something" [5,8] ● [8.1] [We hear a sound of a man laugh] 11:13 P2 "You've got the evil laugh" P2 "Yeah, ok, I've got the evil laugh" [3,8] ● [8.1] 11:17 P1 "I've got the baby" P3 "The killer baby" [3,8] ● [8.1] [We hear a sound of a car crash] 11:22 P1 "Such a window" [2,Ambisonics and realism] ● [8.1] 11:32 P3 "That's the baby, I've got the evil baby" [3,8] ● [8.1] 12:07 P1 [We hear a sound of a cat howling] "You've got the cat" P3 "Cat? it's a huge cat" [3,8] ● [8.1] 12:08 [We hear a sound of gun shot] P2 "The shoot cat" [Identification sound, 8] ● [8.1] 13:14 [We hear a sound of man laugh] P3 "Take the baby evil up" [3,8] ● [8.1] 15:24 [We hear a sound of fire crackle] P3 "Sound effects" P4 "Just crackles" [3,8] ● [8.5] 17:59 [P1 and P2 operate with several cubes positioned at the centre of the table] P4 "There was a change with speaker design" [and P1 and P3 turn their heads towards the speakers] ● [8.1] [We hear a sound of a man laugh] 19:46 P2 "Evily laugh" [3,8] ●
G2	NSP: 8	[8.1]:6, [8.2]:3, [8.4]:1 – [8.1] [We hear a sound of a thunderstorm] 14:04 P2 "thunder wind" P1 "so we do a storm" [3,8] ● [8.1] 16:40 [We hear a sound of a lamp buzz] P4 "this is a (...) buzzing" [in response to P1's question "which sound is which now?"] P1 "that's a buzz (...) they take a while, they take a second before we ... [hear the sound] (...) That's horrible, can we change it? That is like the computer humming in the background" P4 "It's like a generator" P2 "It's slightly irritating this" P4 "All right, ok" [2,3,8] ● [8.1] 19:16 [We hear a sound of a metallic whine] P2 "That's not the wind, is it? that's more like the airplane" [5,8] ● [8.1] 19:37 [We hear the sound of a scream] P1 "That's a horror film now" [2,8] ● [8.2] 19:43 P4 "Let's put down the fire just to keep" P2 "[laughing] Now it's gonna burn in the background" [3,8] ● [8.2] 19:51 P2 "Someone is being murdered" [2,8] ● [8.1], [8.2], [8.4] 21:36 P2 "You should got the fire" P1 "the fire is a nice background [while laughing]" P2 "we've got some screaming babies and we've got some groaning man" P4 "we can create..." P2 "a scene" P4 "a urban scene, campfire with [inaudible] ghosts or something (...) with some thunder as well" ● [8.1] 23:43 [We hear a sound of a tunnel short whistle] P2 "Uh it's yours right? [pointing at a white cube next to P4] you've got the fluty noise" [looking at P4] P4 "Yeah" [2,3,8] ●

SP: 29   [8.1]:12, [8.2]:10, [8.3]:3, [8.4]:7, [8.5]:2 – [8.2] 24:45 [We hear a sound of a fire crackle] P4 "We can just put down the campfire" P2 "yeah we just love the campfire in the corner" [2,3,8] ● (storytelling, musical composition) 24:50 P1 "We should experiment with all the noises just to develop a theme and just make it by leaving the stuff on the table" ● [8.2] 24:54 [We hear a sound of a fire crackle] P1 "So the fire was nice for the background so we leave it" [3,8] ● (musical composition, realism) 24:59 P2 "Find the series of 6 noises that we don't mind having in a sort of (...) like a (...) scene" ● (storyteling) 25:30 P2 "Ok so we are sitting in our little house with our fire getting in the background, so you can hear the storm outside, that's the idea" ●[8.2] 25:44 P2 "Some ghosts in the house, okay" ● ([8.1], [8.3]) 25:42 [We hear a sound of a group whispering] P4 "and then... some ghosts [laughing]" P1 "we can keep the ghosts but we can decrease the volume" P2 "or we can try a different one" [2,8] ● [8.1] 27:07 [We hear a leaking gas hiss] P1 "This could be something on the stove" [3,8] ●[8.4] 27:18 P2 "A bit more realistic, it is a bit more realistic the snoring..." [Identification sounds, 8] ● [8.2] 27:20 P4 "I think this house is going to be... people sleeping and..." P2 "the snoring man with a purring cat by the fire" ● [8.1] 29:09 P1 "[it] could be noises of things outside, in the storm" ● [8.4] 30:12 P1 "The fire sounds really nice, it sounds as it is actually fire" [2,3,8] ● ([8.1], [8.2]) [We hear a sound of a tom drum] 29:13 P2 "Like a door banging or something, over the wind" P1 "it happens every once in a while" [Identification sound, 8] ● [8.1] 31:42 [We heard a sound of a kid laugh] P1 "That's a crying baby, we should decide whether a crying baby or a laughing baby" P4 "I aim for the laughing baby" [3,8] ● [8.1] 30:59 P1 "That's a crying baby" P2 "A screaming one, scared of the thunder" [3,8] ● [8.4] 31:52 P2 "Although possibly the crying one [sound of a crying baby] is more realistic" ● [8.1] 32:26 [We hear a sound of air escaping hiss] P1 "It's like a washing-machine or a dishwasher" [5,8] ● [8.1] 32:56 [We hear a sound of air escaping hiss] P1 "Sounds like a teapot [laughs]" [3,8] ● [8.1] 33:08 [We hear a sound of a pipe hiss] P1 "Like a plane taking off [3,8] ● (story-telling, realism) 36:00 [We hear a sound of a thunderstorm, a tom drum, a fire crackle, a kid laugh, and a purring cat] P2 "This sound is awesome" P4 "We can make some noises" P2 "We created a scene that's all we wanted to do, wasn't it?" P1 "That's the purpose, we have to combine them" P2 "We have our urban noises we've got a purring cat by a fire, and a baby, and some snoring, and a thunder, and a banging door; it seems realistic at least in terms of urban noises" [3,8,9] ●[8.4] 35:29 P2 "We just had our scene" ● [8.1] 34:59 [We hear a sound of a tom drum] P2 "That's the batting door" [5,8] ● [8.1] 35:34 P2 "Let's go for the snoring and the baby" P4 "The noisy baby (...) it is too loud" P2 "It is too loud, isn't?" [2,3,8] ● [8.2] 36:38 P1 "He is sleeping on the ground and the door is banging on his head [laughs], he can't get up, and the baby is just slapping" [laughs] P2 "He just decapitated himself with the door (...) yeah the baby is jumping up and down on him, maybe, and he is breaking his grips" maybe he is just having fun of seeing him banging by the door" ● [8.5] 37:08 P2 "I wonder whether it is interesting where we put these [tangible objects on the table] cause we wanted the fire in the corner if it had some significance, we created a geographical representation as well as the sound one if you see what I mean P1 "it could be a suggestion [to be added to the software] geographical representations that change something, now it doesn't" P2 "I don't think it does" ● [8.2] 37:32 P2 "This one... this one was the snoring, the snoring man is not actually by the fire though [pointing to a white cube with a sound of a long snort]" P1 "It's closer to the door because the door is slipping [pointing to a white cube with a sound of a long snort and a white cube with a sound of a tom drum]" P2 laughs [Identification sounds, 8] ● ([8.1], [8.5]) 37:42 [We hear a sound of a thunderstorm] P2 "The thunder wants to be aware I guess somewhere" [3,8] ● [8.4] 37:51 P2 "We don't have the time or patience to create another scene, have we?" ● [8.2] 38:30 P2 "I suppose you can be reading a paper but it hasn't got a fire to be reading anymore" P4 "It's a dream" ●

G3	NSP: 0	–
	SP: 0	–

G4	NSP: 14	[8.1]:9, [8.2]:4, [8.3]:7 – ([8.1], [8.3]) 05:38 P2 "We can make an audio movie" P4 "Let's go for natural sounds" ● ([8.1], [8.3]) 06:00 [We hear the sound of a kid laugh] P2 "That's a happy baby" P3 "happy baby" P2 "Shall we find a nice sound to go with the happy baby?" [3,8] ● [8.1] 07:46 P2 "There was a cat purring..." P3 "a cat purring" P2 "and then there is the cat 'miauing' it was pissed off" P3 "I don't know whether I have that one" P2 "somebody has that pissed off cat" 3,4,8] ● [8.2] 10:15 P2 "Uh it could be the glasses breaking with the cat being pissed off" ● [8.1] 06:09 [We hear a sound of a thunderstorm] P3 "That's a thunder line that's not good" P2 "That goes with the scream of the baby, with the baby crying" [2,3,8] ● [8.1] 06:21 [We hear a sound of a kid laugh] P2 "This laugh is a little evil" [2,3,8] ● [8.1] [We hear a sound of a groan] 08:58 P2 "Who's got this thing? This voice? This lament?" [2,4,8] ● [8.3] 10:22 P1 "I am gonna give you guys this rhythm [while lifting up and down rhythmically a sound of a wood impact] (...) See if you can have some sounds on top of that rhythm" [2,8] ● (musical composition, storytelling) 10:41 P2 "So I am giving you for 3 bits of that I am giving you a length of noise [we hear a sound of plastic rustling]" [2,8] ● (storytelling, musical composition) 11:52 [We hear the sound of an explosion] P3 "The finale" P4 "We are still alive" [2,8] ● (storytelling, musical composition) 10:52 [We hear a sound of a kid laugh] P4 "I should only laugh at the end (...) oh actually it is still laughing even though I lifted it up [we can see the shade of his white cube]" P3 "yeah I can see the reflection" P4 "Yes it is still reflecting, oh I see" [2,3,8] ● ([8.1], [8.3]) 14:44 [We hear the effect of a black cube next to a white cube] P2 "Gosh" P3 "Dr. Who" [6,8] ● [8.1] 15:59 [We hear a sound of a cat purring with an effect] P2 "This is the purring cat... in a dimension... this is like a purring panther" [3,6,8] ● [8.1] 20:18 [We hear a sound of a cat purring] P3 "That's the cat" P2 "What do you call it? like Shrek, what are they called?" P3 "Ogre" P2 "Ogre" [Identification sounds, 8] ●

SP: 14   [8.1]:7, [8.2]:3, [8.3]:7, [8.4]:1, [8.5]:1 – [8.1] 24:56 [We hear a rhythmic sound of a ding from a ceramic impact] P2 "That sounds like a supermarket (...) a cashier of a supermarket" [Identification of sounds, 8] ● [8.3] 26:52 [We hear a sound of a pipe hiss that appears and disappears] P2 "My role is to turn this one [laughs] I think it's a bit [inaudible]" [2,8] ● [8.3] 30:13 P2 "But we are producing music" P3 "Music? Futurist music" P4 "[inaudible] noise, still music" [8.3] 32:47 P2 "Right, I think I've done that long enough" [making a sound of a pipe hiss appearing and disappearing] P3 "You got bored!" P2 "Maybe I can try another..." [2,8] ● [8.2] 33:06 P3 starts humming a song and asks "Do you remember this song called popcorn?" P3 "popcorn? yes" P2 "you do? P3 "That's so old" P2 "I know, well, the song I meant [laughs]" P3 "Oh you are so... well thanks for the compliment [laughs]" P2 "I meant the song, well I remember it so there you go" P3 "so you were going to say on prerelease weren't you?" P2 "no i was very very very very small.." P3 "Ohhhh" P2 "[laughs] but i remember it, on the 45, my parents had a 45, what you call it in english, the vinyl" P1 "the single vinyl" P2 "Yeah" ● [8.3] 33:49 P4 is dragging 2 groups of objects until they touch to each other P3 "You are not crashing things into one, are you?" [looking at P4] P2 "What are you doing?" P4 "Composition various different composition" [2,8] ● [8.2] 34:55 [We hear a sound of a long snort] P2 "So there could be somebody laughing at me because I am snoring" [Identification sounds, 8] ● [8.1] 34:36 P3 "This makes that [pointing to a white cube on the table], whatever that is, this is not the cat purring, I think this turns the cat purring into snoring" [actually P2 points out that she holds the sound of snoring that P3 is attributing to another white cube] [5, 8] ● [8.1] 34:59 [We hear a sound of a long snort] P2 "It's like my grandfather" ●[8.3] 34:44 [We hear a sound of a long snort] P2 "No this is the cat snoring [showing a white cube she is holding], I've got the snoring" [in response to P3's comment that another white cube changes the sound of a cat purring to snoring] P3 "Have you?" P2 "I am making it really loud, look" [Identification sounds, 8] ● ([8.1], [8.3]) 35:51 P1 "Now it has become a cacophony I think (...) this is whatever John Cage's Silence but the opposite" P3 "Yeah [laughs]" [2,8] ● [8.1] 36:40 [We hear a sound of a long snort, and of a cat howling, and of a kid laugh] P2 "These are neighborhood sounds" 36:44 [We hear of a tibetan chant] P2 "Not that one" [2,8] ([8.1], [8.2], [8.4]) ● 36:46 P2 "See when we manipulate this all I can think of is scenes of places, and things happening, to me it's like the soundtrack of you know of some movie or something" P3 "Check out number 3 please" P2 "what's that number 3?" P3 "check out number three please, ding, check out number three please, ding" P2 "That's it! P1 "[inaudible]" P2 "That's what I was saying a cashier or a supermarket like that" ● [8.5] 38:04 P1 "I am curious if position affects which speaker sound comes out but it doesn't" P3 "No, there is no surround sound [looking at the four speakers]" P4 "It could be stereo sound [looking at the four speakers]" P1 "It could be 4 channel, you could actually orient the proximity of the cube to the speakers, which channel to use, but... [looking at his right back and front speakers]" ● ([8.1], [8.3]) 39:14 P3 "It's definitely Dr. Who" ●

G5 NSP: 8   [8.1]:6, [8.2]:2, [8.3]:3, [8.5]:1 – [8.5] 24:08 [We hear a sound of a tibetan chant] P3 "There is a little [change]..." ● ([8.1], [8.2], [8.3]) 24:33 [We hear a sound of a tibetan chant] P1 "It's a good one" P2 "Boring" P1 "The boring chant or the orthodoxy drowning" P2 "Heavy metal" P1 "No unfortunately one of the [inaudible] just can't sleep" P2 and P4 laugh [2,3,8] ● ([8.1], [8.3]) 25:26 P1 "Take the percussion now [inaudible], industrial based" [2,8] ● [8.1] [We hear a sound of a tibetan chant] 27:13 P1 "That drone is the winner" P4 "Yeah" P1 "Are you making that drone?" P4 "What?" P2 laughs P1 "Are you making the drone yourself?" P4 "This one [pointing to a white cube]" P1 "I think you need to mark somehow that box" P2 "Yeah" P1 "That's the winner" [2,4,8] ● [8.3] 27:34 [We hear a sound of a tibetan chant] P1 "Keep with the drone is good" [2,8] ● [8.1] 27:36 P1 "We've got this peacefulness" [referring to a sound of a tibetan chant, named 'the drone'] ● [8.1] 36:33 [We hear a sound of a car crash] P4 raises his two hands 36:36 P3 "Strike!" [2,8] ● ([8.1], [8.2]) 38:48 [We hear a sound of a cat howling, and then a sound of a kid laugh] P1 "There's something quite chilly about the cat, it sounds like a baby, ours it's quite a bit like that, especially when it wags" P2 "When I was a kid and used to hear a cat crying outside I used to take my dad it sounded like a baby on the doorstep making my dad look" P1 "True, it's absolutely chilly" [2,3,8] ●

	SP: 11	[8.1]:8, [8.2]:1, [8.3]:7, [8.4]:1) – [8.1] 08:11 P1 "[To the facilitator] Yeah give me a folder with the sounds [used in the session]" P3 "For the door bell bzzzz bzzzz" • [8.1] 08:32 P1 I've been watching "The Sopranos" this week • ([8.1], [8.3]) 10:28 P2 "It's interesting it's a kind of gothic type of sound, right?" P4 "Yeah [and smiles]" • [8.3] 10:36 P1 "I'm just doing a pulse [while lifting up and down a white cube with a sound of a gun shot]" P2 "Well done" [2,8] • [8.3] 11:00 P1 "That works [nodding his head following a rhythm lifting up and down a white cube with a sound of water splash]" [2,8,9] • ([8.1], [8.3]) 13:18 P1 "We got a few..." P2 "Drumming [while lifting up and down a white cube with a sound of a metallic impact]" [Identification sounds, 8] • [8.1] 13:26 P1 "I seem to have the thunder claps [while lifting up and down a white cube with a sound of a thunderstorm]" [3,8] • [8.3] 13:41 P2 "Nice [inaudible] [while lifting up and down a white cube alternating between a sound of a metallic impact and a tom drum]" P3 "[inaudible] slow movement" [2,8] • ([8.1], [8.3], [8.4]) 16:48 P1 "It was the piece for a real scene really, isn't? It was a mix of the London Olympic game ceremony" [laughs] [8,9] • ([8.1], [8.2]) [We hear a sound of a scream] 16:57 P2 "Screaming" P1 "Yeah bringing the ending planet" [Identification sounds, 8] • ([8.1], [8.3]) 20:23 P1 "We've arrived, we are doing Stockhausen now, we've arrived, that's definitely Stockhausen, take it back down, stepping in, boom" [laughs] [8,9] •
G6	NSP: 2	[8.1]:2, [8.3]:1 – ([8.1], [8.3]) 10:50 [We hear a sound of a metallic whine] P3 "I've got the highest" [2,8] • [8.1] 20:06 P2 "It's like the horse in the rain" 20:07 P3 "Sorry?" 20:08 P2 "The horse in the rain" •
	SP: 3	[8.1]:1, [8.3]:3, [8.4]:1 – ([8.1], [8.3]) 25:57 P3 "We can make a nice drumpad with this" [P2 and P3 alternate rhythmically between lifting up and down a white cube each, one with the sound of a gun shot, and the other with a sound of a tom drum] • (musical composition, realism) 37:56 P3 "We should record a song like this [with] 2.6 or..." [8,9] • [8.3] 38:16 P2 "Let's put all white ones in full mode" [2,8] •
G7	NSP: 11	[8.1]:8, [8.2]:6, [8.3]:1, [8.5]:1 – [8.1] 27:36 [We hear a sound of a falling metal, a sound of rocks smashing] P1 "What is this?" P4 says no with her head P3 "A tsunami [inaudible]" [4,5,8] • [8.1] 27:40 [We hear a sound of rocks smashing] P2 "Working with boots or something" [3,8] • 33:59 P2 "What sounds have we got?" P4 "We've got snoring boring creepy cats another like boring" P2 "The cat one is good" P3 "Cats and babies" 34:13 P3 "And the... " P4 "Oh yeah, and the exciting one" [1,2,8] • [8.1] 34:19 [We hear a cat howling] [laughs] 34:30 P2 "Can you combine it [cat howling] with the snoring?" P3 "See you know there's a cat and a baby on here, on this, I think this sounds more like a combination between the cat and the baby" [laughs] P4 "Like a cow" P3 "Yeah" P2 "Every cat with snoring" [Identification of sounds, 8] • [8.2] 35:31 [We hear a sound of a kid laugh with an effect] P4 "So you make quite a good at doing mess in your head on how many permutations there are" P2 laughs • ([8.1], [8.2]) 35:52 [We hear a sound of a kid laugh with an effect] P2 "Ok so this is the sound of the parent making jokes out about himself" P3 "It looks an old man laughing" All laugh P3 "Or maybe a bit obsessed" P4 "Sounds like an exorcist stuff" P2 "So maybe we can try if it is snoring to have some fun" [2,3,6,8] • ([8.1], [8.2]) 36:22 P4 "This one is the creepy one, [we hear a sound of a gun shot]" 36:27 P2 "Kill the baby" 36:30 P4 "And it's so wrong" [1,2,8] • [8.2] 36:37 [We hear a sound of a car crash] P3 "Just shot the baby" P2 laughs [2,8] ([8.1], [8.2]) 36:50 P2 "Let's make the creepy cat" P3 "The creepy cat is going to shoot something" 37:04 P2 "Exactly, it's gonna shoot someone" P1 laughs • ([8.1], [8.3]) 37:19 [We hear the sound of a man laugh] P2 laughs P2 "Let's make a horror movie" [2,8] • ([8.1], [8.2]) 37:25 P2 "So we have the cat, we have the shooting and..." P4 "We have like the sony baby" P3 "We have a creepy background" P2 "Exactly we have the [inaudible] baby" P3 "It's walking down a park" P2 "And maybe someone is sleeping and gets shot in the sleep [we hear a sound of a gun shot], you can also add the snoring one" P3 laughs P4 "Yes, so there's a shot as long as I know" P2 "This is the snoring [we hear a sound of a long snort]" P4 "Somewhere yeah" [we hear a sound of a groan] P2 "This is the shot" [we hear a sound of a gun shot] P3 "It's like [inaudible]" P2 "And then burn it" P2 "Ok I think we left them not very good" [Identification sounds, 8] • [8.5] 40:01 P2 "So the sound comes mainly from this speaker, yeah? [pointing to his back right speaker]" P1 turns his head to his back left speaker, and P4 turns her head to her back right speaker and then to her back left speaker. There are no comments. •

	SP: 7	([8.1] 5, [8.2] 1, [8.3] 1): [8.3] 11:19 [We hear a sound of a car crash, a boom rumble, a pipe hiss] P2 "So let's make this positive" P1 & P3 laugh P2 "Happy now" [2,8] ● [8.1] 12:12 [We hear a sound of a cat howling and then a sound of a groan] P2 "This is the sleepy cat" [2,8] ● [8.2] 17:32 [We hear a sound of a tibetan chant, before there was a sound of a cat howling] P3 "The cat is praying [laughs]" [3,8] ● [8.1] 17:39 [We hear a sound of a tibetan chant, a ceramic impact, a fire crackle, a boom rumble] P3 "It's like noises from films" [2,8] ● [8.1] 19:21 [P1 is manipulating the 'explosions' white cube and we hear a sound of a car crash] P3 "Who's like the vampire car?" P2 "It's probably..." 19:25 P4 "It might be someone with [inaudible] anything that fill over [inaudible] cause it does steeping" P2 "I prefer the curtain" [4,5,8] ● [8.1] 20:57 [We hear a sound of a long snort] P2 "Ok this is the I think sleeping dog [pointing to the 'whispers' white cube with a sound of a long snort]"[5,8] ● [8.1] 23:23 [We hear a sound of a man laugh] P3 "It is so [adventurous] [laughs]" [2,8] ●
G8	NSP: 6	[8.1]:4, [8.2]:2, [8.3]:3 – [8.1] 08:56 [We hear a sound of a group whispering] P1 "whispering, underground noises [laughs]"[3,8] ● [8.1] 10:00 [We hear a sound of a tibetan chant] P3 "Like a story [inaudible]" ● [8.3] 11:05 P2 "We started making some music"[Realism and ambisionics, 9] ● ([8.1], [8.2]) 11:12 [We hear a sound of a boom rumble, and a car crash] P3 "Is that a thunderstorm?" [we hear a sound of a gun shot] P1 "Yeah or the shot & the thunderstorm" P3 "It's like things happening" [3,5,8] ● [8.2] 13:32 P2 "Uh the sounds are playing with us" 13:34 P3 "Yeah" 13:35 P2 "[inaudible] 6 noises" ● ([8.1], [8.3]) 18:19 P2 "I wouldn't call it music, I don't really know..." P4 laughs P1 "House" P2 "Some kind of order on it" P3 "That's what I thought you need order" P3 "But then John cage wouldn't say you need order, would he?" P2 "Yes, I suppose no, that's why" ●
	SP: 5	[8.1]:1, [8.2]:2, [8.3]:3 – [8.3] 26:47 P2 "I give you the bit [while lifting up and down rhythmically the 'whistles' white object with a sound of an air escaping hiss]" [2,8] ● [8.2] 28:37 [We hear a sound of a man laugh] P3 "He's just laughing at us" [Identification sounds, 8] ● [8.2] 29:13 [We hear a sound of a cat howling and an effect] P3 "Yes something else in there" [in response to P3's question "There is a second voice?"] P1 "It's like, so they are two cats now" [P3 adds another black cube next to a white cube with a sound of a cat howling] P1 "3 cats?" P1 "Wow, 3 cats!" P4 smiles P2 "More" P1 "4 cats" P3 laughs [3,6,8] ● ([8.1], [8.3]) 33:24 [P2, P3 and P4 are lifting up and down rhythmically a white cube each] P2 "[inaudible] a kind of percussion anyway" P3 "It's music now" [Realism and ambisionics, 9] ● [8.3] 34:56 P2 "[inaudible] sounds that you like, I am trying to build something out of it" ●

D.8.9 Immersive musical experience

TABLE D.12: Transcripts of *Immersive musical experience*

Group	Condition	Quotes
G1	NSP: 1	28:08 P2 "What are you trying now?" P1 "This is nice, it's a different way" [after pressing on and off a cube side with the sound of a tunnel short whistle at different parts of the table rhythmically] ●
	SP: 0	–
G2	NSP: 0	–
	SP: 1	36:00 [We hear a sound of a thunderstorm, a tom drum, a fire crackle, a kid laugh, and a purring cat] P2 "This sound is awesome" P4 "We can make some noises" P2 "We created a scene that's all we wanted to do, wasn't it?" P1 "That's the purpose, we have to combine them" P2 "We have our urban noises we've got a purring cat by a fire, and a baby, and some snoring, and a thunder, and a banging door; it seems realistic at least in terms of urban noises" [3,8,9] ●
G3	NSP: 0	–
	SP:1	07:22 P2 "Cool!" [P1 laughs] [2,9] ●
G4	NSP: 2	19:26 P3 "We are already close to a preview performance, aren't we?, we don't have that much time for" ● 20:04 P2 "Oh for a moment we had something nice" ●

	SP: 3	30:57 P2 "I think we have reached some musical plateau" ● 36:08 P3 "I am thrilled for the department of performance, with mandolines and string quartets [laughs] (...) some experimental music" P2 & P4 laugh P4 "That makes a good match for the [Barbican]" P3 "Oh yes" P2 "Oh dear dear" ● 39:10 [In response to P2's question "Who's got this?"] P3 "[These are all these things] in the middle" P2 "all together?" P3 "Yeah" P2 "That's very cool!" [2,9] ●
G5	NSP: 0	–
	SP: 10	11:00 P1 "That works [nodding his head following a rhythm lifting up and down a white cube with a sound of water splash]" [2,8,9] ● 12:30 P1 "I like what we've got, it's good!" 12:32 P2 "Yeah" P3 laughes ● 13:29 P1 and P2 are nodding her head following the rhythm ● 14:27 P1 "That's good! [while nodding his head following the rhythm]" ● 17:19 P1 & P3 are nodding their heads rhythmically ● 16:08 P1 "There are too many pictures, are there? they are mostly, fairly sound effected" P1 "I like that [nodding his head]" [6,9] ● 18:18 P1 & P2 are nodding their heads rhythmically ● 16:48 P1 "It was the piece for a real scene really, isn't? It was a mix of the London Olympic game ceremony" [laughs] [8,9] ● 20:23 P1 "We've arrived, we are doing Stockhausen now, we've arrived, that's definitely Stockhausen, take it back down, stepping in, boom" [laughs] [8,9] ● 20:49 [We hear a sound of a water splash with an effect] P1 "I like that [while turning a white cube left and right next to a black cube and nodding his head rhythmically]" P1 & P2 are nodding their heads [2,6,9] ●
G6	NSP: 1	13:35 P2 "Wow, good [while turning left and right a black cube next to a white cube]" [6,9] ●
	SP: 1	37:56 P3 "We should record a song like this [with] 2.6 or..." [8,9] ●
G7	NSP: 0	–
	SP: 0	–
G8	NSP: 1	11:05 P2 "We started making some music" [Realism and ambisionics, 9] ●
	SP: 2	25:51 P2 "Uh I quite like that [creating a rhythm by lifting up and down the 'whistles' white object with a sound of an air escaping hiss]" [2,9] ● 33:24 [P2, P3 and P4 are lifting up and down rhythmically a white cube each] P2 "[inaudible] a kind of percussion anyway" P3 "It's music now" [Realism and ambisionics, 9] ●