# Computer Musicking: Designing for Collaborative Digital Musical Interaction.

Fencott, Robin.

# Computer Musicking:

## *Designing for Collaborative Digital Musical Interaction*

**Robin Fencott**

Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2012

# Computer Musicking:

## *Designing for Collaborative Digital Musical Interaction*

**Robin Fencott**


### Abstract

This thesis is about the design of software which enables groups of people to make music together. Networked musical interaction has been an important aspect of Sound and Music Computing research since the early days, although collaborative music software has yet to gain mainstream popularity, and there is currently limited research on the design of such interfaces. This thesis draws on research from Computer Supported Cooperative Work (CSCW) to explore the design of systems for Collaborative Digital Musical Interaction (CDMI). A central focus of this research is the concept of Awareness: a person's understanding of what is happening, and of who is doing what. A novel software interface is developed and used over three experimental studies to investigate the effects different interface designs have on the way groups of musicians collaborate. Existing frameworks from CSCW are extended to accommodate the properties of music as an auditory medium, and theories of conventional musical interaction are used to elaborate on the nature of music making as a collaborative and social activity which is focused on process-oriented creativity. This research contributes to the fields of Human-Computer Interaction (HCI), Computer Supported Cooperative Work, and Sound and Music Computing through the identification of empirically derived design implications and recommendations for collaborative musical environments. These guidelines are demonstrated through the design of a hypothetical collaborative music system. This thesis also contributes towards the methodology for evaluating such systems, and considers the distinctions between CDMI and the forms of collaboration traditionally studied within CSCW.


Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2012

# Acknowledgements

I would like to thank my supervisors Dr. Nick Bryan-Kinns and Prof. Mark Plumbley for offering me the opportunity to undertake this research, and for supporting me throughout. Also my examiners Dr. Robin Laney and Dr. Tim Blackwell for taking the time to read this thesis, discuss my work and offer constructive criticism.

My gratitude to everyone who offered support and academic advice along the way. There are too many people to name in person, however in particular I would like to thank Prof. Peter McOwan, Prof. Pat Healey, Dr. Jennifer Sheridan, Dr. Andrew McPherson, Dr. Tony Stockman, Dr. Matthew Purver, Dr. Dan Stowell, Dr. Andrew Robertson, Dr. Oussama Metatla, Dr. Tim Murray-Browne, Dr. John Dack and Dr. Thor Magnusson.

Many thanks to all those who read draft chapters and papers: my supervisors, my colleagues within the Interaction, Media and Communication (IMC) and Centre for Digital Music (C4DM) research groups at Queen Mary University of London, and my parents Clive & Anne Fencott.

Thanks to all my family and friends for their support, encouragement and much needed distraction over the last four years. Clive & Anne Fencott, Lue Cuttiford, Sean Kennedy, Max Payn, David Scales, Bob Nicholson, Steven Atkinson, Matthew Jeffries, Sarah Atkinson, Matt Keane, Imre Torok, Sarah Hacking-Brian, Evguenni Penksik, and all my friends at Queen Mary University of London.

Finally, I would like to thank all the participants for taking part in my experiments. This work was funded by an Engineering and Physical Sciences Research Council (EPSRC) Doctoral Training Account Award (DTA) .

# Contents

# List of Figures

# List of Abbreviations

**CDMI** Collaborative Digital Musical Interaction

**CSCW** Computer Supported Cooperative Work

**DMI** Digital Musical Interaction

**GT** Grounded Theory

**GTM** Grounded Theory Methods

**GUI** Graphical User Interface

**HCI** Human-Computer Interaction

**NIME** New Interfaces for Musical Expression

**SDG** Single Display Groupware

**TA** Thematic Analysis

**WA** Workspace Awareness

**WYSIWIS** What You See is What I See

# Chapter 1

# Introduction

---

Music is a fundamental part of human expression (David J. Hargreaves and Meill, 2005, Makelberge, 2010, Small, 1998), and although not always a collaborative activity (Makelberge, 2010), the creation, performance and enjoyment of music is nevertheless highly social in nature; people often write music in creative partnerships (John-Steiner, 2000), perform to audiences as ensembles (Jordà, 2005b), and attend concerts together (Small, 1998). Acknowledging the pervasive nature of music within human life, Small (1998) uses the term Musicking to assert that music is not an object, but rather an activity and social ritual through which people explore their identity and relation to others. A literature survey suggests that the theory of Musicking has so far been overlooked within musical interaction research, although the underlying message of Musicking has been proposed as a tool for facilitating creativity in other group work contexts (Atau Tanaka, 2012).

According to the theory of Musicking, there is no fundamental difference between any of the ways an individual can engage with the rituals of music; listening to a personal stereo, playing in a rock band, dancing or attending a classical music concert should all be regarded equally as acts of Musicking. Small argues that all musicking is a social phenomenon, and no individual can detach themselves from the musical traditions of their culture. For many people, the balance of challenges, rewards and satisfaction associated with Musicking provides a route to the psychological state of Flow; a state in which a person forgets themselves, loses track of time and through single-minded immersion becomes fully engrossed in the activity at hand (Csikszentmihalyi, 1990). Likewise, playing in a musical ensemble can be a source of Group Flow, a situation in which a group of people all reach a flow state whist engaged in a joint activity (Sawyer, 2003). The theories of Flow and Group Flow neatly capture the sense of immersion and engagement musicians feel when 'jamming' or 'grooving' together (Swift et al., 2011).

*Computer Musicking*; the title of this thesis, reflects the notion that computer technology has permeated many aspects of Musicking. There are multiple technologies for listening to, storing and searching digital recordings of music, whilst the internet provides a social forum for discussion, discovery and sharing of musical recordings. There are also countless software applications for the representation, manipulation and creation of music, all based on a variety of musical paradigms (Gerzso, 1992) and interface metaphors (Duignan et al., 2004). Yet crucially, whilst music making is often a group activity, the majority of existing music applications (commercially available, within academic research or otherwise) are designed to support single user interaction (Bryan-Kinns, 2012, Bryan-Kinns and Hamilton, 2009, Fencott and Bryan-Kinns, 2010, 2012, 2013, Healey et al., 2005). The single user paradigm adopted by the majority of music software entails a view of musical interaction in which an individual interacts with a single system, or a group of individuals each interact with there own personal and discrete systems. There are very few examples of music software which let groups of musicians collectively and simultaneously interact with musical data or representations in real-time. Research into the design, development, use and evaluation of such interfaces is the central concern of this thesis.

Within academia there has been limited interest in the development of technology to support the collaborative creation of music using computers. The work that has been conducted has for the most part been concerned with technical innovations, rather than focusing on the evaluation of user interface designs, or studying the activities surrounding musical collaboration. Consequentially, there are a multitude of choices to be made when implementing interfaces for collaborative music making, yet little information about how these interface design choices will impact on the way groups of musicians work together. Through an analysis of thirty-five groups of three people using novel collaborative musical interfaces, this thesis explores a number of fundamental design concerns, and proposes a series of practical implications for future collaborative musical interfaces. This research was informed by a thorough literature survey of previous research into collaborative music environments, and studies of collaboration in non-musical domains such as workplace interaction. Later parts of this thesis acknowledge the differences between musical interaction and other forms of collaborative work, paying careful attention to the sonic, and time based nature of music.

## 1.1 Designing for Collaborative Digital Musical Interaction

Computers afford entirely new forms of musical interaction which are not possible with conventional instruments, and computers consequentially provide opportunities for new forms of musical interaction between people (Bown et al., 2009). Going futher, Makelberge (2010) frames a division in musical interface design research between the questions of *how* people make music, and *why* people make music, arguing that while the former has radically altered, the latter remains unchanged. Designing new systems for musical collaboration may therefore entail the creation of new relationships and interactions between musicians. For example the internet allows multiple people to jointly create music in real-time from remote locations, multi-player instruments let groups of musicians interact in complex and interdependent ways with a single interface, and where musicians are in the same place at the same time, networked instruments distributed across multiple computers can (as demonstrated by this thesis) provide an ensemble with new opportunities to share, conceal, collectively edit and jointly organise their musical contributions in real time.

This thesis proposes the term Collaborative Digital Musical Interaction (CDMI) to classify these new forms of technologically mediated musical interaction, and to help understand how to design systems to support them. The term CDMI bounds the concept in three ways. Firstly, it focuses on the collaborative aspects of musical interaction, where groups of people jointly participate in the creation of music. This represents a departure from previous discussions of digital musical interaction which largely focus on interfaces for single users (Gurevich and Fyans, 2011). Secondly, it emphasises the broad process of interacting musically, rather than concentrating on the constraints of a particular musical context such as performance, composition or improvisation. Finally, CDMI focuses on the use of digital technology to support and mediate collaboration, but does not commit to a specific device or platform. Subsequent sections of this thesis explore the relationship between CDMI and other forms of group collaboration, and in doing so reveal the challenges entailed in designing systems to support CDMI.

## 1.2 Awareness and Awareness Mechanisms

There is currently a paucity of empirical studies to investigate interface design concerns for collaborative digital music environments, and there are no established methodologies for conducting such research. This thesis therefore turns to the field of Computer Supported Co-operative Work (CSCW) (Carstensen and Schmidt, 1999) as a starting point for understanding group collaboration. In particular the concept of Awareness (Gutwin and Greenberg, 2002) forms an important conceptual foundation for the user studies conducted here. Within the context of collaboration, awareness refers to a person's understanding of the people around them and activities taking place in the shared workplace.

Within the field of CSCW, the terms 'co-located' is used to describe situations in which groups of people share the same physical environment whilst collaboration. In co-located collaboration where a group of people are working on a shared task in the same place and at the same time, awareness of the actions and presence of others is easy to maintain (Gutwin and Greenberg, 2002). In particular, the perceptual information gained through being able to see and hear other group members provides a wealth of information about their activities. Co-located collaborators may also be available for conversation about the task, whilst overhearing the discussions of other people in the room may provide further cues to what is happening (Heath et al., 2002b). Finally, the state of the physical workspace may present information about the task at hand and may offer opportunities for awareness to be gained about the status of tools and artefacts associated with the activity.

In contrast, where collaboration is mediated by computers and shared virtual workspaces, many of these sources of awareness information may be prohibited, causing collaboration to become more difficult. For instance a shared on-screen workspace may not provide the wealth of information available when people work together in the same physical space, and the interactions associated with a computer interface provide a fraction of the perceptual information made available through face-to-face collaboration (Gutwin and Greenberg, 2002). Therefore, when computers become involved in supporting collaboration, the interface must be designed such that awareness information is conveyed to support the participants in carrying out their activities. An understanding of awareness has implications for other aspects of collaboration, such as privacy (hiding and sharing contributions), territory, and the congruence of interfaces made available simultaneously to multiple users. Subsequent sections of this thesis more thoroughly unpack these concepts.

Interface components which support aid in the provision of awareness information in collaborative systems are known within the field of CSCW as *Awareness Mechanisms* (Dourish and Bellotti, 1992, Gutwin et al., 2004). Within CSCW research, Awareness Mechanisms have previously taken a variety of forms, including live video feeds of people's workspaces (Hudson and Smith, 1996), visual indicators of the location of other users' mouse positions in collaborative software, (Greenberg et al., 1996) and real-time sound effects to indicate the activities of others (Gutwin et al., 2011). However, to date there has been little research into Awareness Mechanisms within systems for collaborative musical interaction. One contribution this thesis makes, therefore, is to consider and investigate how previous research into awareness mechanisms from CSCW can apply to the design of musical interfaces.

Despite drawing on concepts and methodology from CSCW, this thesis also stresses that musical interaction is in some respects fundamentally different from the workplace collaboration typically discussed in CSCW literature. Consequently, ideas brought forward from related fields need to be modified, or applied with care. For instance in a workplace context it may be advisable to reduce the dependancies (known as 'coupling') between people required to complete a task, so as to to increase efficiency by allowing people to work in parallel. However in a real-time musical context where 'efficiency' is a less meaningful concept, an increase in coupling may be beneficial in promoting more tightly interwoven interaction between the musicians.

An understanding of awareness can benefit design for both remote interaction (e.g. via the internet) and real-time co-located musical interaction where musicians are working in the same place at the same moment in time. Although remote and distributed forms of musical interaction are an important and fascinating field of research, it is the co-located domain which this thesis deals with. This decision is based on several observations. Firstly, the most common means by which musicians collaborate has traditionally been co-located and real-time, yet beyond the timing synchronisation of single-user devices there is only limited technological support for co-located interaction. Moving forwards, the interfaces studied within this thesis extend beyond synchronisation by allowing users to modify each others musical contributions in real-time. Secondly, there are a number of existing research studies dealing with real-time and asynchronous musical interaction over the internet (see Chapter 2), yet there are fewer attempts to study co-located, computer mediated musical interaction using shared interfaces. Finally, as has been argued with reference to other forms of collaboration (Heath et al., 1995, 2002b), an understanding of co-located interaction is a prerequisite for designing software to support people working remotely.

## 1.3 Research Questions and Approach

Previous CSCW research has emphasised the role and importance of awareness and privacy as part of collaborative systems (see Sections 2.2 and 2.2.3). Within collaborative music research however, these concepts have yet to be fully explored. The overall question this thesis therefore addresses is *how do interface mechanisms which support awareness and privacy effect the use of collaborative software environments for music-making?*

In relation to this overarching concern, a number of more focused research questions are addressed by three experimental studies:

- *what effect does the provision of privacy have on the way individuals share and edit musical contributions?*
- *what effect do awareness mechanisms have on the way shared software environments are used?*
- *How does the way audio is delivered to co-located groups impact on collaboration?*
- *How do mechanisms for authorship and sound source awareness alter the way people engage in collaborative digital musical interaction?*

In approaching these questions the thesis re-examines existing concepts and models of collaboration through the lens of musical interaction. The concept of CDMI is proposed to bridge knowledge from these previously disparate research domains and provide insights into the design and study of new collaborative computer music tools. Design implications for CDMI are explored and supported by three controlled experiments whereby groups of musicians used software environments for real-time co-located musical interaction.

All the primary data presented in this thesis is based on studies of human participants, and a mixed of quantitative and qualitative methodologies have been employed to gather and analyse this data. Each study follows an experimental design in which multiple conditions are compared. The conditions in each study vary, however in all cases they represent different interface design choices and contain different features for the support of collaboration. Quantitative analysis has been performed on interaction logs of the participants' use of the software, thus enabling the identification of statistically significant differences in the way participants used the software when presented with different interface conditions. This controlled approach represents a departure from the frequently presented 'informal' studies of new digital musical interfaces.

Multiple choice questionnaire responses were also analysed for statistical significance to collect self-reported data such as preferences from the participants. In addition to these quantitative measures, all the experiment sessions were video recorded and manually transcribed for further analysis using qualitative approaches. Analysis of the participants conversations provided insight into the problems participants faced, the topics of discussion, and the ways in which they understood and interpreted the software interface. Finally, interviews were conducted to collect the participants' thoughts and suggestions about the software and the music they created.

The recruitment process for this research sought to find participants with musical and/or computing backgrounds, including playing traditional musical instruments, DJing, or composing music with computers and electronic instruments. Using participants with musical backgrounds and experience offered a more informed insight into the interfaces being evaluated than would have been possible using non-specialist users. This also enabled the software to contain specialist musical and technical terminology, and allowed the inclusion of features similar to those found in professional music-making applications. This approach represents a departure from previous research into musical collaboration between groups of novice users (Bryan-Kinns, 2004, Gurevich, 2006, Miletto et al., 2009, Weinberg, 2003), and contributes to an understanding of how to design applications for musical collaboration which are more advanced than simplistic 'sound toys' (Robson, 2001) or musical games (Pichlmair and Kayali, 2007). This is an important point to stress, as it has implications for both the generalisation of the findings presented, and the application of this work in a real-world context.

Technologically, the focus of this thesis is the design of software which transforms multiple generic computing devices such as laptop computers into a platform for musical collaboration. This is a practical decision intended to promote the design of software which lets people collaborate using existing equipment. It has been suggested that for musical collaboration, new technological developments such as interactive tables are superior to conventional computing devices (Xambó et al., 2011a), however such devices are not yet commonplace. That being said, many of the design implications and findings within this thesis may apply to the development of collaborative interfaces for other platforms and devices, such as tablet computers.

## 1.4 Contributions

The research questions posited in the previous section are answered by the three studies presented in Chapters 5, 6 and 7.

Chapter 5 contributes to the field by addressing the roles of privacy and awareness in a collaborative music environment. The study identifies several significant effects brought about by the inclusion of privacy. In particular, the findings suggest that when privacy is incorporated into the interface, users tend to use the privacy feature to develop and refine musical contributions in isolation from one another before sharing with the group. The study also investigates the effect of an awareness mechanism which allows users to gain awareness of who is doing what, at the expense of reducing the level of privacy. Results suggest that the awareness mechanism does not change the way the software was used, although it does influence the participants experiences whilst working together.

Chapter 6 investigates the effects of audio delivery in a collaborative music environment, by comparing three configurations of headphones and speakers for the presentation of public and private audio. The findings indicate that using a single device for audio delivery promotes a greater degree of reported awareness, and that presenting public and private audio exclusively through speakers causes participants to make significantly less use of the privacy feature. Chapter 6 also investigates territorial organisation of the interface by the participants, and identifies a significant difference between groups who use the shared workspace in a territorial fashion and groups who do not.

Chapter 7 investigates the distinction between awareness of who contributed which musical parts (Authorship Awareness), and awareness of which interface components are responsible for creating which sounds (Source Awareness). The study findings suggest that providing Authorship awareness causes users to work more individually, and leads to a reduction in the degree to which they are willing to edit musical contributions they did not create. The use of Thematic Analysis also provides insights into the way the interface was interpreted by the participants.

This thesis is immediately relevant to designers of new collaborative music applications, and to researchers evaluating existing collaborative systems for musical interaction. In addition to directly addressing the research questions posed in Section 1.3, this thesis makes a number of more general contributions to the field:

1. The application of CSCW design concepts within interfaces for musical collaboration

2. A range of empirical findings concerning the effect of different awareness mechanisms and collaborative interface designs on the processes of group musical interaction:

    - Chapter 5 stresses the importance of privacy mechanisms within shared musical interfaces by identifying several significant effects such features have on collaboration
    - Chapter 6 identifies a significant change in the use of a privacy feature when audio is delivered through speakers as opposed to headphones
    - Chapter 6 demonstrates that groups exhibiting territorial interaction within the shared interface are statistically less likely to edit each others' musical contributions
    - Chapter 7 argues that providing visual and auditory information about the authorship of musical contributions within a shared interface encouraged users to work more individually

3. A series of design implications, derived from controlled empirical studies of user interaction, and offering practical advice to designers of new interfaces for musical collaboration (Chapter 8)

4. The design of a hypothetical collaborative interface which embodies these design implications (Section 8.4)

5. A comparison of Collaborative Digital Musical Interaction to other forms of technologically supported collaboration, articulated through a consideration of the properties of sound and the constraints imposed by technological mediation (Chapter 3)

6. A novel combination of quantitative and qualitative methods (including Thematic Analysis) employed to study collaborative digital musical interaction (Chapters 4, 5, 6, 7)

This thesis places a strong emphasis on the properties of sound as a medium and the implications this has for computer based collaboration. Consequentially, the outcomes of this research have application in forms of interaction outside of collaborative digital musical interaction. In particular the findings may directly translate to the design of other forms of collaborative systems where sound and auditory interaction are of central importance. Contributions in this area are made through:

- A comparison of different audio delivery devices for collaboration between groups of three people jointly working with auditory media (Chapter 6)

- A comparison of different interface designs for presenting information about the authorship and origin of sounds in a shared interface (Chapter 7)

## 1.5 Additional work

In addition to the research outputs outlined above, the following practical work was conducted during this PhD:

- The iterative development of a collaborative musical interface, used as a vehicle for conducting the empirical studies presented in Chapters 5, 6 & 7
- Sensory Threads, a group mobile sensing and sonification experience created in collaboration with Proboscis, Birkbeck Colleges Pervasive Computing Lab, The Centre for Digital Music at Queen Mary (University of London), Mixed Reality Lab (University of Nottingham) and the School of Management at University of Southampton. See Fencott and Bryan-Kinns (2009).

## 1.6 Thesis Structure

Following this section, the thesis is structured as follows:

Chapter 2 provides a comprehensive literature review of related research from the fields of Computer Supported Cooperative Work (CSCW), and previous work on collaborative musical interfaces, including contributions from the New Interfaces For Musical Expression (NIME) conference series.

Chapter 3 proposes the term Collaborative Digital Musical Interaction (CDMI). Through considering the nature of sound, and the technological constraints digital mediation can impose on musical interaction, this chapter considers some of the key characterises of group musical interaction and frames a distinction between CDMI and other forms of computer supported collaboration, such as those typically studied in the field of CSCW research.

Chapter 4 presents the research methods employed within this thesis, and provides motivation for the software developed to conduct this research.

Chapters 5, 6 and 7 present and discuss the three empirical studies conducted for this research. Chapter 5 investigates the role of privacy in a collaborative music environment. Chapter 6 considers how audio delivery devices can influence the way individuals share their musical contributions. Chapter 7 explores how different sound identity and authorship awareness mechanisms effect the way musicians choose to view, edit and share in a collaborative environment.

Chapter 8 draws together the findings of the three studies, relates this to previous research, and proposes a number of design implications for the design of systems to support CDMI.

Chapter 9 summarises the findings of the three studies, re-states the contributions and briefly concludes the thesis with suggestions for future work.

## 1.7 Associated Publications and Presentations

*Publications*

Fencott and Bryan-Kinns (2013). “Computer Musicking: HCI, CSCW and Collaborative Digital Musical Interaction” *Music and Human-Computer Interaction, Springer Cultural Computing Series. Springer*. (In Press)

Fencott and Bryan-Kinns (2012) “Audio Delivery and Territoriality in Collaborative Digital Musical Interaction” *The 26th BCS Conference on Human-Computer Interaction, Birmingham, UK*.

Fencott and Bryan-Kinns (2010). “Hey Man, You’re Invading My Personal Space: Privacy and Awareness in Collaborative Music” *NIME: New Interfaces for Musical Expression, Sydney, Australia*.

Fencott and Bryan-Kinns (2009). “Sensory Threads: Sonifying imperceptible phenomena in the Wild” *6th Sound and Music Computing Conference, Porto, Portugal*.

*Presentations and Demonstrations*

18th April 2012. “SuperCollider as Audio Engine for Collaborative Music Applications”. SuperCollider Symposium 2012. London, UK.

27th-29th July 2011. “Computer Musicking”. InterFace 2011, UCL, London, UK.

8th February 2011. “Group Computer Music”. Invited talk, Middlesex University, Cat Hill Campus, UK.

3rd February 2011. ‘Collaborative Computer Music’. Invited presentation of research in relation to music technology for education and special needs. TeachMeet Music, Cornerstone House, Croydon, UK.

9th Dec 2010. Presentation of research at Culture Lab, Newcastle University, Newcastle upon Tyne, UK

16th-17th September 2010. “Collaboration, Privacy, Awareness, Music”. Presentation and participation in ShareIT Workshop on Exploring Collaboration with Shareable Interfaces, University of Sussex, Brighton, UK.

3nd June 2010. Paper presentation: “Computers in Support of Group Music-Making”. Sight, Sound, Space and Play Conference. Music, Technology and Innovation Research Centre of De Montfort University Leicester.

2nd June 2010. Session Chair: Composition and Analysis. Sight, Sound, Space and Play Conference. Music, Technology and Innovation Research Centre of De Montfort University Leicester.

11th May 2010. Presentation of collaborative software at the London SuperCollider meet-up group . a10Lab. Peckham, London, UK.

26 February 2010. Invited presentation on research within the Interactional Sound and Music group. Mixed Reality Lab, University of Nottingham, Nottingham, UK.

# Chapter 2

# Background

---

*This chapter surveys relevant work from within the fields of Human-Computer Interaction, Computer Supported Cooperative Work and previous research into collaborative music systems. Particular attention is given to the concept of Awareness, to and evaluation methodologies for musical interaction and collaborative systems.*

Chapter 1 recognised CSCW as an established research field concerned with the design of systems to support collaboration in the workplace, and acknowledged that there is a precedent for the development of technology to mediate group musical interaction. While it is argued within this thesis that workplace activities and musical interaction are in many ways distinct (see Chapter 3), there are many design concerns which are shared by both computer supported workplace collaboration and computer supported creative activities such as music making. For instance in both forms of activity it is important for group members to maintain awareness of each others actions, and issues such as ownership may play an important role in the unfolding interaction. These issues have frequently been addressed by researchers studying collaborative systems for the workplace, although a consideration of such design concerns is less commonly reported in literature surrounding computer supported musical collaboration. Literature on technology for musical collaboration has more often focused on issues such as documenting compositions and the development of novel controllers for musical interaction, reducing network latency and using the internet for the purpose of real-time jamming or performance.

This chapter begins by surveying literature from the fields of Human-Computer Interaction (HCI) and Computer Supported Cooperative Work (CSCW) to identify contributions which are relevant or influential to this thesis. Particular attention is given to work which concerns support for co-located collaboration, and research dealing with issues such as awareness, territoriality and privacy. The chapter then shifts focus to literature from the computer music community, dealing in particular with musical collaboration, New Interfaces for Musical Expression (NIME) and methods for the evaluation of new forms of musical interaction. While technological contributions within both CSCW and computer music research are acknowledged in this chapter, the primary focus is to survey and assess work which deals with the design for, and study of human interaction.

## 2.1 Computer Supported Cooperative Work

Computer Supported Cooperative Work (CSCW) is a specialised branch of Human-Computer Interaction (HCI) focusing on understanding the nature of group work and designing appropriate technology to support work-related collaboration between people (Bannon and Schmidt, 1991). CSCW can be viewed as an umbrella term for all HCI research concerning multiple users, however Hughes et al. (1991) stressed that all work, no matter how individual, occurs within a wider social context. CSCW therefore might usefully be regarded as a paradigm shift within the Computer Science community away from the view of HCI as interaction between people and computers, and towards a social view of work and interaction as a collective phenomenon. Some recurrent themes within CSCW literature are awareness, role formation, collaborative coupling, support for both private and public working, and the study of verbal and non-verbal communication during workplace activities. These issues are addressed in subsequent sections of this chapter.

Two related approaches to research within CSCW are the design and evaluation of collaborative and multi-user software termed ‘groupware’, and the study of collaboration in real-world work settings. Studies of real-world work activities are often performed to gain a deeper understanding of current working practices in a given context, with the intention of proposing design guidelines for technology to support established working arrangements, or to facilitate new forms of work practice such as remote online collaboration. For example the ClearBoard was an early attempt to support remote collaborative drawing and drafting activities using digital video to recreate aspects of face to face interaction (Ishii et al., 1993).

Although generally not concerned with musical interaction, CSCW is significant within the context of this thesis as it represents a substantial body of research into the nature of group-work and the design of technology to augment collaboration between people. Within this thesis, the research questions, study designs and evaluation methodologies were all informed by research from the CSCW community.

### 2.1.1 Examples of CSCW

To set the scene, and provide context for readers who are not familiar with CSCW research, this section briefly describes several systems and software implementations which have been contributed to the field.

*Clearboard*

The Clearboard (Ishii et al., 1993) is a video-mediated collaborative system which enables users working in geographically remote locations to engage in drawing activities. The Clearboard comprises of a drawing surface which also acts as a back-projected video display. The system is based on the metaphor of two people drawing together whist sitting on opposite sides of a glass window (see Figure 2.1b, and attempts to re-create many of the interactional properties teams of co-located people experience when working together on a whiteboard (see Figure 2.1a).


(a) Whiteboard interaction (b) Clearboard glass prototype (c) Clearboard System

Figure 2.1: The Clearboard system and prototypes. Images reproduced from http://tangible.media.mit.edu/project/clearboard/ Last accessed 5 October 2012.

The Clearboard uses video communication to transmit the marks a user makes on a drawing surface to the back-projection of another Clearboard in a different location. The transmission of marks is reciprocated by the second users' system, enabling the two users to jointly work on a drawing in real-time. In addition to transmitting the drawn marks, the system also transmits an image of each users upper body, allowing the Clearboard to show the gesture, gaze and the hand position of each user. By flipping the video image along the vertical axis the system also provides a consistent representation of left and right orientation for both users. Figure 2.1c shows the system in action, with a user drawing on the projection surface, whist another user is projected as a video link behind the drawn image. A demonstration video is also available at the MIT Media Lab website[1].

The key contribution of Clearboard is that users can easily maintain awareness of each others' gaze, as they can simultaneously view their partner's face and upper body, as well as the shared image they are working on. This feature of Clearboard distinguishes it from conventional video conferencing systems which do not provide information about where people are currently looking. However, one drawback to the clear board system is that it does not scale to supporting large numbers of users, as the images of multiple users would need to be overlaid on top of one another.

*ShrEdit*

ShrEdit (Dourish and Bellotti, 1992) is another early example of a CSCW system. ShrEdit allows groups of people working on separate computers to collaboratively edit a text document. ShrEdit allows users to work individually in private windows, as well as allowing users to work in publicly shared windows. Each user is represented with their own cursor, thus allowing concurrent editing of the same document. Changes to a document are displayed with low latency to all other users, however ShrEdit does not however provide information about the position of other users' mouse cursors, or their gaze on the screen.

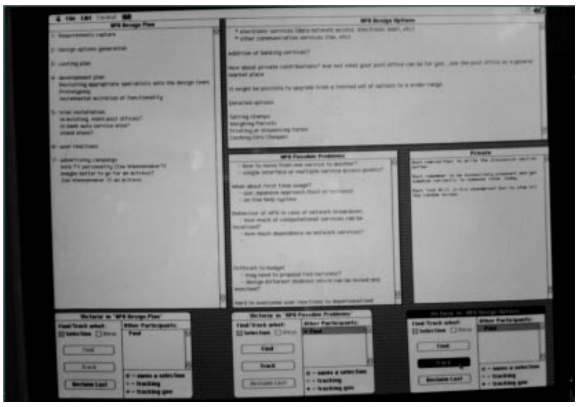


Figure 2.2: Screenshot of the ShrEdit interface, showing private and shared text editing windows. Reproduced from Dourish and Bellotti (1992)

[1]http://tangible.media.mit.edu/project/clearboard/ Last accessed 5 October 2012

*GroupKit Framework and Applications*

GroupKit is a framework that lets developers build desktop applications for synchronous and distributed computer-based conferencing and real-time long distance interaction (Roseman and Greenberg, 1996). The framework handles communication, remote procedure called between application instances, access to shared data, and GUI widgets for designing distributed applications. Gutwin and Greenberg (1999) present a number of studies using applications built with the GroupKit framework, for instance Pipeline, a desktop application which allows users working remotely on separate computers to collaboratively design and edit a complex plumbing network. Applications such as this allow multiple people to simultaneously edit a shared document and also offer means of communication such as video links or text chat tools. In the Pipeline system Gutwin and Greenberg (1999) explored the use of the 'Viewports' an awareness mechanism which shows where within a document other users were currently viewing and editing. Another study by Greenberg et al. (1996) focused on 'Telepointers', awareness mechanisms which display the mouse position of other users as a means of presenting more fine grained information about the activities of other collaborators. Figure 2.3 shows the two screenshots of the Pipeline interface side by side. The right hand interface featuring the Local viewport which indicates where the user is currently looking, whilst the left hand interface also features teleprinters and multiple viewports, to indicate the presence of other remote users.

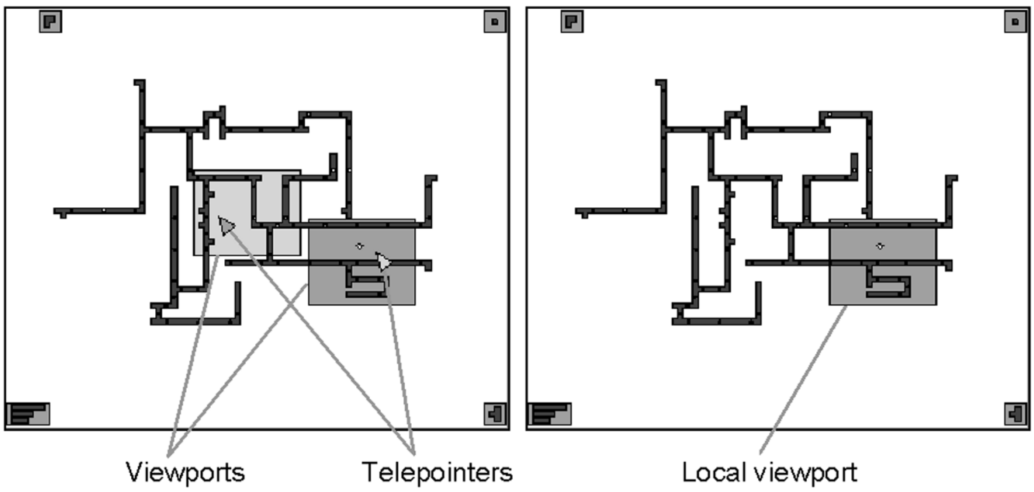



Figure 2.3: Screenshot of the Pipeline interface with Radar views. Reproduced from Gutwin and Greenberg (1999)

### 2.1.2 Methodology in CSCW

Within the remit of CSCW, a variety of methodological approaches have been used to study collaboration. Workplace studies have employed ethnographic techniques (Hughes et al., 1994) in a variety of environments, such as offices (Rouncefield et al., 1994), trading rooms (Heath et al., 1995), news rooms, operating theatres, control rooms (Heath et al., 2002b) and high risk environments which demand full concentration, and where people's activities are time-critical, highly interdependent and potentially life-threatening (Endsley, 1995, Gutwin and Greenberg, 2002). Other studies have looked at the complex interactions between people engaged in creative tasks and design activities (Healey et al., 2005, Healey and Peters, 2007). These ethnographic workplace studies typically attempt to characterise the interaction and group processes that take place through the presentation, analysis and detailed discussion of specific incidents within the observed interaction. For instance Heath et al. (1995) identified instances of individuals verbalising thoughts, activities and work-related information as a means of informally drawing co-workers' attention to potentially salient details about the task at hand. These verbalisations were named 'outlouds', and were identified in a number of different workplace environments.

Furniss and Blandford (2006) and Hutchins (1996) used Distributed Cognition frameworks to develop rich accounts of interaction in work settings. Distributed Cognition can be used to explain some of the ways in which groups of people are able to collaborate despite holding incomplete, disparate or conflicting mental representations of the situation or activity. Distributed Cognition also considers the role objects, artefacts and technologies play in shaping and scaffolding cognition and collaboration. Finally, Distributed Cognition has proved a useful tool for the study of musical interaction (Nabavian, 2009, Nabavian and Bryan-Kinns, 2006). Heath et al. (2002b) argued that an understanding of the ways in which groups of people engage in focused work activities can inform approaches to CSCW, however they stressed that the outcome of observations about the way people work should not necessarily be embodied in the design of new collaborative systems until there is a thorough understanding of how people engage in conventional activities. Similarly, Rouncefield et al. (1994) stressed that in addition to systems design to support workplace collaboration, CSCW is concerned with 'work redesign', and consequentially CSCW research requires a sensitivity to the (possibly unexpected) implications such redesigns may have on the target workplace.

Broadly speaking, an alternative to ethnographic or observational workplace studies of collaboration are laboratory studies which take place in artificial or controlled environments rather than genuine work settings. By definition, this type of study observes participants in a unnatural context, rather than in a genuine workplace environment. This approach may give more options for data collection, by for instance allowing the researcher to set up multiple video cameras, interrupt participants during interaction (Endsley, 1995), conduct interviews and collect questionnaire information. In many cases, such methods cannot be used within purely observational workplace studies, as they might interrupt the people under observation, and disrupt the natural flow of the activities the researcher is attempting to study. The use of questionnaires and interviews might also alert the participants to the motivations of the researcher and could potentially cause the participants to alter their behaviour in response to what they believe the researcher is seeking to discover.

Evaluation of groupware in CSCW research is often based on an experimental approach where groups perform a task or work activity using interfaces with different features or support for collaboration. For instance Gutwin and Greenberg (1999) investigated different on-screen awareness widgets for supporting coordination in a shared document editors. Similarly Dourish and Bellotti (1992) explored the role of privacy in shared document editing by observing how groups collaborate using different interface configurations, and Olson et al. (1993) measured performance in group activities with and without technological support. Such studies can offer additional flexibility for the researcher, for instance in determining the activities participants are engaged in, and the introduction of additional constraints on the participants, such as limiting the time allocated for a given task or restricting the available resources.

Typical observations and measures in laboratory studies are task completion time, quality of the solution, ease of collaboration and user satisfaction. Gutwin and Greenberg (1999) described these features as Product (the result or outcome of a task), Process (the activity of the group while developing their solution), and Satisfaction (feelings about the work and interaction with the system). These metrics are appropriate within a work context where the emphasis is on efficiently completing a task to a high standard, and where the outcome and performance of the task is quantifiable (Olson et al., 1993). However in less productivity oriented domains such as music the balance may need to be adjusted. This issue is explored further in Chapter 3.

Grounded Theory (Charmaz, 2006, Glaser and Strauss, 1967) and Grounded Theory derived techniques (Muller and Kogan, 2010) are increasingly common research methods within HCI and

CSCW (Furniss et al., 2011, Muller and Kogan, 2010). The definition of Grounded Theory is a contentious issue, due to the conflicting ideological perspectives of some of its' key proponents. Grounded Theory is based on a process of coding and continual comparison of coding with the data set, which is performed in tandem with data collection. The data collection and sampling process should also be informed by the on-going coding process. The method also requires copious note taking, memo writing and comparison of these written documents with previous codes and collected data. This bottom-up approach is intended to promote the discovery or construction of theories which are 'grounded' in data, in opposition to methods which 'forcing' data into a preexisting theories. Grounded Theory methods typically propose that underly theories 'emerge' from the data set (Glaser and Strauss, 1967), although more constructivist views acknowledge the active role a researcher plays in the generation of such theories (Charmaz, 2006). Within this thesis, Grounded Theory methods are applied in Chapter 6.

In the case of both lab based studies and real-world observations, studying collaborative groups presents many practical challenges. In ethnographic studies gaining access to existing workplaces may be difficult, and the presence of the researcher may have effects on the activities being studied (Hutchins, 1996). On the other hand, Rogers et al. (2009) discussed the logistical problem of recruiting groups for experimental sessions, and implications this has for both experimental designs and reliable statistical results. Large numbers of sample groups may improve the reliability of statistical results, however finding a large quantity of suitable participant groups may be difficult, or impossible given limited research budgets, project timeframes and access to potential participants. In an experimental context the researcher may also need to balance the controls and manipulations introduced through the experimental design against the similarity the activity bears to a real-life work setting.

## 2.2 Awareness

The concept of awareness is central to a great deal of CSCW research, and underlies the research questions addressed within this thesis. Awareness can be viewed as a person's knowledge of their environment (Gutwin and Greenberg, 2002), and can include the location and activities of people (Dourish and Bellotti, 1992), a memory of previous events, the state of systems or artefacts within the environment and predictions about what may happen in future events. Within CSCW, researchers have used the concept of Situation Awareness (SA) to describe the awareness requirements for people working in dynamic, high risk, and high information load situations (Endsley, 2000), or 'centres of coordination' (Heath et al., 2002b). Endsley (2000) defined Situation Awareness as 'knowing what is going on around you', but specified that SA applies to situations in which individuals need to maintain high levels of awareness as part of their current activity, as opposed to a situation where an individual may have some form of awareness, yet is not required to act upon it. Endsley (2000) describes three levels of Situation Awareness (SA):

**Level 1: Perception** The ability to gather and selectively attend to perception of cues and important information.

**Level 2: Comprehension** Interpreting and integrating information from multiple sources, to make sense of the current situation.

**Level 3: Projection** Anticipating future situations and making decisions.

Gutwin and Greenberg (2002) used the three levels of SA to formulate the *Workspace Awareness* (WA) framework as a means of describing awareness between small groups of people in physically shared workspaces. WA is defined as an 'up-to-the-moment understanding of another person's interaction with a shared workspace', and concerns the means by which collaborators gather and use information about each-other's activities while working on a joint task. Gutwin and Greenberg (2002) decomposed WA into the categories of *Who*, *What*, *Where*, *How*, and *When*. Table 2.1 gives examples of these categories.

The Workspace Awareness framework is based on studies of or co-located interaction, meaning where the collaborators are in the same physical location at the same moment in time. WA is also concerned with groups of collaborators that are focused on a specific work related activity in a physical workspace. Considering remote, technologically mediated interaction Gutwin and Greenberg (2002) stated that when using groupware, participants may also face difficulties at any of Endsley's three awareness levels due to the technological constraints of the software and technologically mediated communication. This idea is revisited in Section 3.3, where the constraints technology imposes of musical interaction are discussed in more detail.

It has been stated that awareness information an individual holds about their environment is not static, but is maintained over time during an individual's conduct and interaction (Heath et al., 2002b). Individuals do not simply possess awareness information, but they must gather and maintain it throughout the course of an activity (Gutwin and Greenberg, 2002, Heath et al., 2002b). Supporting this argument, Akoumianakis et al. (2008) emphasised that in order for an

| Who | |
|---|---|
| Presence | Is anyone in the workspace? |
| Identity | Who is participating? Who is that? |
| Authorship | Who is doing that? |
| What | |
| Action | What are they doing? |
| Intention | What goal is that action part of? |
| Artifact | What object are they working on? |
| Where | |
| Location | Where are they working? |
| Gaze | Where are they looking? |
| View | Where can they see? |
| Reach | Where can they reach? |
| How | |
| Action | How did that operation happen? |
| Artifact | How did this artefact come to be in this state? |
| When | |
| Event | When did that event happen? |
| Who (past) | Presence history Who was here, and when? |
| Where (past) | Location history Where has a person been? |
| What (past) | Action history What has a person been doing? |

Table 2.1: Categories of Workspace Awareness, from Gutwin and Greenberg (2002)

individual to make sense of what others are doing they must invoke their own knowledge of the common practice. As such, maintenance of awareness is not simply a result of 'being there' but instead requires explicit effort, and domain knowledge. Heath et al. (2002b) also claimed that awareness is not necessarily symmetrical between individuals, but rather that individual's awareness of others must continually be produced through social action, and that awareness does not arise simply because individuals have overlapping frames of activity. Within the WA framework, Gutwin and Greenberg (2002) identified a number of ways awareness information is gathered by individuals, including *Consequential Communication, Feedthrough and Intentional Communication*.

*Consequential communication* is information made available to others as a consequence of ones actions. For instance bodily movement, gestures and physical orientation associated with an activity may convey details to others about the activity at hand. Consequential communication is not designed explicitly to communicate with others, and is therefore unlike exaggerated body movements (Hornecker et al., 2008) or actions designed to alert others to one's own actions (Heath et al., 1995). However as many workplace activities involve bodily movement (hands, arms, head, etc.,) and the physical manipulation of artefacts, a wealth of information can be gained about others' activities through casual observation.

*Feedthrough* refers to information generated by inanimate artefacts in the work environment, such as tools. The term is derived from the notion of feedback, where an artefact communicates information about its state through, for instance, sounds or visual indicators. The concept of feedthrough extends this idea by suggesting that these sounds and visual indicators may also be picked-up by other people within the environment. In this way, feedthrough information allows a person to infer the activities of someone else without reference to their bodily action, gestures, posture and so on. Gutwin and Greenberg (2002) and Smith and Hudson (1995) state that feedthrough information is typically auditory, for instance Gutwin and Greenberg (2002) used the example of hearing the sound of scissors 'snipping' to demonstrate how people can gain information about what is happening in a shared workspace (i.e. the scissors are being used) via feedthrough. Feedthrough and consequential communication are closely related, but the terms are not interchangeable, as feedthrough becomes an aspect of consequential communication when associated with information gleaned through observation of a person's activities, rather than simply information which is made available in the workplace through sounds or other indicators.

Verbal communication and associated gesturing are acts of *intentional communication* which can be used to share awareness information in a workspace. Verbal communication can be used in planning, discussion of activities, delegation of roles and in seeking assistance or clarification. Through talking, awareness information may be conveyed not only to those who are being spoken to, but also to others in the workspace who may overhear. Heath et al. (2002b) observed that individuals often design their actions to be selectively noticeable by collaborators through the use of body language and verbalisation, whilst simultaneously allowing their collaborators to continue their own activities without disruption. For instance the production of 'outlouds' (Heath et al., 1995), a form of verbal commentary which can informally alert others to important information within the workspace. Along with verbal communication comes intentional gestures which can be used to communicate information such as shape, size and motion.

### 2.2.1 Awareness Mechanisms

Within groupware, Awareness Mechanisms are features which convey information about the activities of users, or the status of systems or processes within the interface. Although awareness has been a common feature in CSCW literature, the awareness mechanism studied in CSCW have been primarily visual, rather than auditory (Gutwin et al., 2011). Examples include multiple views onto a document (Dourish and Bellotti, 1992), telepointers to show the position of other users' mouse pointers (Gutwin and Penner, 2002), video channels (Olson et al., 1995), shared spaces (Ishii et al., 1993) and three-dimensional representations (Benford et al., 1994). Within CSCW research, the use of audio in collaborative interfaces is an under explored modality (Metatla, 2010), although it is acknowledged as a potentially fruitful area of future research (Gutwin et al., 2011).

Hudson and Smith (1996) state that awareness mechanisms serve the dual function of alerting others to an individual's activity whilst simultaneously reducing the level of privacy they have within a shared interface. As such, awareness mechanisms may have an effect on the way people choose to share, communicate or contribute in a collaborative environment. Rogers et al. (2009) stated that knowing ones actions are highly visible to the group via a shared interface may influence the degree to which people contribute or are willing to participate. Similarly, within electronic group brainstorming activities, the phenomenon of Evaluation Apprehension (EA) has been used to describe the negative influence awareness can have on participation, by potentially making people reluctant to contribute due to fear of retaliation or judgement by peers (Olson et al., 1993, Pinsonneault et al., 1999).

### 2.2.2 Studying Awareness

Support for awareness has been identified as a central aspect of groupwork (Gutwin and Greenberg, 2002, Heath et al., 2002b, Rogers et al., 2009), however awareness resides within the individual and is therefore difficult to measure or study (Endsley, 1995). Observation and video coding are common approachs to analysing awareness gathering in group interaction. The Workspace Awareness framework (Gutwin and Greenberg, 2002) is based on observations of groups in shared workspaces. Examples of video based studies of awareness include Healey and Peters (2007), Heath et al. (1995, 2002a,b) and Hornecker et al. (2008), Scott et al. (2004). Another common method of evaluating the effects of awareness support in software is through a comparison of task executions, where participants are presented with experimental conditions in which different amounts of awareness information are presented. For instance Gutwin and Greenberg (1999) measured factors such as task completion time, communication efficiency and perception of effort using two versions of a collaborative software application. However, studying the effects awareness mechanisms have on a task outcome is only effective if the outcome can be rated or quantified in terms of a measurable quality (Olson et al., 1993).

### 2.2.3 Mixed Focus in Collaborative Systems

A key interface metaphor in groupware is the What You See is What I See (WYSIWIS) abstraction. This metaphor is based on the notion of a shared physical workspace such as a chalkboard, which is in use by a group of co-located individuals who all have consistent visual access to the shared content. Systems which employ a strict WYSIWIS design adhere to the principle that multiple users working on separate devices all see an identical on-screen representation of the information held within the workspace. One disadvantage of this design is that it forces all contributions to be public, and prevents individuals from working in isolation from the rest of the group. In response to this, Stefik et al. (1987) discussed the importance of relaxing the WYSIWIS abstraction along four dimensions: space, time, population and congruence. These relaxations to the metaphor enable support for many of the activities afforded by pen-and-paper interaction in shared workspaces. For instance writing individual notes, deciding what to share with others, and breaking into smaller collaborative groups.

The term Mixed Focus refers to the movement between individual and coupled work during collaboration (Gutwin and Greenberg, 1998). Individual work refers to people working independently of one another, and conversely, coupled work refers to situations where two or more individuals engage in a highly interdependent activity (Tang et al., 2006). Dourish and Bellotti (1992) and Heath et al. (2002b) observed that within group-work, people frequently shift between individual activities and tightly coupled collaboration. Support for privacy and individual work in CSCW systems is typically through the inclusion of personal workspaces (Stefik et al., 1987) or other forms of access control. Shen et al. (2003) addressed the issue of privacy and personal information spaces on shared screen based computers by allowing personal data to be viewed on individual screens, whist shared data was made available through a tabletop display. They also stressed the importance of considering private and public data as a continuum, rather than in binary terms. However it is important to acknowledge that within technologically mediated collaboration there may be less well defined social conventions associated with privacy (Hudson and Smith, 1996), and as such it may be difficult to predict how features intended to support mixed focus interaction will be employed by users.

Personal and private workspaces in group text editing have been explored by Dourish and Bellotti (1992) and Olson et al. (1993). In these studies, participants used a shared text editor which provided private and shared editing spaces. It was noted that the private spaces were not used extensively during a real-time collaborative task, as participants preferred to work in the publicly shared workspace. The researchers proposed that this may have been in part due to the short time duration of the task, which forced participants to work quickly in order to finish on time. Their research also considered the emergence of roles in the interaction, and the function of awareness in supporting task completion. In relation to non-musical auditory collaboration, the importance of providing private and public workspaces was emphasised in Metatla (2010).

However, incongruent, or mixed focus WYSIWIS interfaces can become problematic from an awareness standpoint as users may be working in entirely different ares of a shared workspace, and may therefore have a reduced awareness of the activities of others. Greenberg (1996) therefore argued for the importance of incorporating awareness mechanisms to counteract confusion in relaxed congruence WYSIWIS systems. Within WYSIWIS systems design, researchers have explored the presentation of ownership information in groupware in the form of colour coding contributions (Gutwin and Greenberg, 1999), telepointers (Gutwin and Penner, 2002) to indicate where other users' mouse cursors are on the screen, and graphical overviews which present users with a simplified impression of what other users' screen layouts look like. Smith and Hudson (1995) and Gutwin et al. (2011) have also emphasised that audio can be a powerful means of supporting awareness in shared workspaces.

### 2.2.4 Territoriality

Within collaborative activities, territory is another important way awareness is mediated. Territorial activity by someone in a shared software environment can signal that an individual or group has claimed ownership or responsibility for objects, artefacts or spaces. In relation to collaborative three-dimensional virtual environments, Benford et al. (1994) noted that space is a central factor in establishing awareness, arguing that it is a form of awareness mediation and activity enabler that provides a means for constraining access, marking territory and designating locations for particular activities. Scott et al. (2004) discussed the role of physical space in supporting collaborative tabletop activities, noting in particular the ways in which different physical workspace areas are appropriated by people for different activities during collaboration. Territory can also be established through explicit communication between people, for instance agreement over tasks, space and the spatial layout of an activity.

Territory has been identified as an important feature of authorship awareness in collaborative document writing and other forms of creative collaborative activities (Thom-Santelli et al., 2009). In the case of online document editing within the Wikipedia community, Thom-Santelli et al. (2009) observed that existing software functionality was appropriated by expert users to signify ownership, despite guidelines against this behaviour. This observation is interesting as it suggests the importance and prevalence of territorial activities in collaboration, even in situations where the written rules and cultural norms warn against the idea of ownership. Finally, territory can also emerge through activities associated with the collaborative environment itself, for instance some activities lend themselves to natural spatial arrangements (Tse et al., 2004).

### 2.2.5 Sharable Interfaces and Co-Located Collaboration

Heath et al. (2002b) stated that computer workstations are designed for one user at a time interaction, rather than simultaneous use by multiple people. As a consequence, collaborators may often have restricted and incongruent access to information about the activities of their colleagues, and the single user nature of the technology may make it difficult to informally gather awareness information. As an example it may not be possible for someone to determine 'at a glance' the details a colleague may be entering via a keyboard, or to casually take note of what is currently displayed on another person's computer screen.

This section discusses multi-user Sharable Interfaces for co-located interaction in collaborative work contexts. Such interfaces begin to address the problems presented by single user devices such as desktop computers. Rogers et al. (2009) drew a distinction between 'early groupware technologies' designed to support work within groups of geographically distributed individuals, and 'Shareable Interfaces' designed to support physically co-located people in tasks which require co-presence. Under the general umbrella of Sharable Interfaces a variety of collaborative shared-screen based systems have been developed, including multi-touch surfaces (Hornecker et al., 2008, Rogers et al., 2009), large public displays (Brignall and Rogers, 2003), and multi-user extensions to conventional desktop computers (Stewart et al., 1999). Given the focus of this thesis on co-located interaction, the more current research on sharable interfaces is especially relevant.

A central distinction between groupware and sharable interfaces is that while a distributed groupware application may be required to re-create, simulate or convey the functions of co-present verbal and non-verbal interaction (e.g., gaze, physical gesture, orientation) via software mechanisms such as telepointers (Gutwin and Penner, 2002), Shareable Interfaces designed for co-located interaction do not need to provide explicit support for these aspects of interaction. Instead, Sharable Interfaces can attempt to leverage the affordances of co-located interaction to support collaboration which is more efficient, equal, free-flowing (Rogers et al., 2009) and serendipitous (Shen et al., 2003) than conventional single user interaction paradigms such as laptops and computer workstations. Similarly Stewart et al. (1999) defined Single Display Groupware (SDG) as a model of co-present collaboration where people use private interface channels such as personal keyboards and mice to interact with a shared screen. Stewart et al. (1999) argued that the shared nature of these interfaces could enhance the collaborative experience by reducing social barriers between users, eliminating conflict over shared input channels, and encouraging more communication between users as the shared nature of the interface requires communication in order to resolve conflict. Negative aspects of SDG were identified, notably that new forms of conflict may arise through different users performing simultaneous incompatible actions, and that task completion time may be increased as SDG does not permit a single strong willed user to direct the collaboration through a monopolisation of the input channels. Finally Stewart et al. (1999) noted that the parallel nature of the interface might encourage users to work in isolation, rather than communicating or contributing to the primary group task.

Sharable Interfaces are most frequently evaluated through comparative studies of novel and conventional computer interfaces (Rogers et al., 2009, Stewart et al., 1999) or through comparison to 'pen-and-paper' based approaches which do not incorporate digital technologies. As there are few instances of novel Shareable Interfaces being used in real-world scenarios the majority of studies in the literature take place in contrived or artificial situations. Observation, video, interaction log analysis and interviews are common techniques for studying interaction around Sharable Interfaces (Rogers et al., 2009). For instance Ryall et al. (2006) discussed observations of people interacting with multi-touch screens in public places, noting in general that people did not associated these interfaces with conventional computer systems, although during initial encounters with the device, groups of people under-exploited their capacity for simultaneous multi-user interaction.

Using a generalised 'Shared Information Spaces' metaphor Rogers et al. (2009) presented a detailed analysis of groups performing a collaborative task using three different physical interface metaphors, a conventional laptop computer, a touch-surface and physical-digital interface which combines digital data representation with tangible control. Their analysis considered the variance in the amount and type of conversational and interactional turn taking, turn negotiation, joint action and verbal and participatory equality. In addition they discussed the arrangement of physical resources and the ways in which location of collaborators and devices was observed to contribute to the structure of the interaction. Their study suggests that allowing users to have multiple points of 'entry' or interaction within a shared interface can be beneficial for creative activities such as design tasks, although the nature of the interface can have an effect on the equality of contributions between group members and the degree to which groups work in a parallel fashion. Similarly Scott et al. (2004) observed territorial behaviour in tabletop activities, discussing in particular spatial organisation, and the means by which groups utilise space to work, share and store items. Tse et al. (2004) studied the way people spatially separate an SDG interface to avoid overlapping or interfering with each-other's work. They argued that spatial partitioning happened consistently between groups of people, and occurred naturally without verbal negotiation. They also argued that participants naturally avoided spatially interfering with each other in a shared workspace, however the exact patterns of spatial division were determined by factors such as the underlying structure of the task and the seating arrangement of the participants around the shared screen.

## 2.3 Auditory Interaction and Collaboration

Auditory interaction is under-explored research areas with HCI and CSCW (Metatla, 2010), and consequently there is a paucity of research addressing the properties and affordances of audio as a medium for interaction and collaboration. This section surveys research concerning collaborative sonic interaction design, and related issues such as the delivery of sound to users in support of activities which are primarily visual in nature.

### 2.3.1 Audio Delivery

Gutwin and Greenberg (2002) acknowledged that during face-to-face collaboration in shared workspaces many forms of feedthrough awareness are auditory (see Section 2.2 for a discussion of the term 'feedthrough') , however there are few examples of research exploring the affordances of audio for collaboration in software interfaces. Metatla (2010) explored audio-only collaborative diagram editing in support of visually impaired users, using task based measures of evaluation. Within the context of workspace awareness Gutwin et al. (2011) identified distinctions between synthetic and real-world sonic feedback, and argued that using sound in distributed groupware could aid users in gathering awareness, especially if they are working in different areas of a shared workspace.

In a single user context Kallinen and Ravaja (2007) investigated the psychophysiological effects on individuals of listening to news broadcasts on headphones or speakers, reporting that headphones cause people to become more immersed in the task at hand and less conscious of their surroundings. Nelson and Nilsson (1990) reported similar results in a single user simulated driving activity. These studies show that the devices used to deliver audio can encourage a variety of different interactions and can influence the way people engage with both the task and their surroundings. However perhaps due to the focus on visual interfaces within HCI research, there is very limited information on this phenomenon within HCI or CSCW research. Within the domain of music and music technology, Mulder (2010) studied spatial audio and the perceptual implications of sound amplification, surveying relevant work, and discussing the effects of musical instrument amplification in a performance context, although this research was primarily conducted to inform artistic outcomes. Pfadenhauer (2009) described the idiosyncratic way in which DJs hold or wear headphones while cueing or pre-listening to music during a performance, although this is a small part of a much wider ethnographic study into DJ culture.

Limited research has been directed towards studying the interactional or collaborative affordances of devices for presenting audio. For instance there is currently contradictory information on the effect headphones have on human interaction and engagement in both single user and multi-party contexts. Morris et al. (2004) compared the use of shared speakers and shared speakers plus individual headphones to deliver auditory feedback in a collaborative multi-touch system, demonstrating that group's task strategies changed when users were given a personal audio channel presented through a single in-ear bud. With the personal audio channel participants tended to adopt a divide and conquer approach to complete a collaborative decision making task whereas when audio was presented using a single speaker shared by the entire group, the participants tended to work more sequentially. Morris et al. (2004) also observed that using single in-ear bud style headphones did not impede group communication. Similarly, Metatla (2010) compared the effects of speakers and headphones on participants' ability to maintain Workspace Awareness (Gutwin and Greenberg, 2002) using a collaborative diagram editor.

Within the domain of collaborative music there is also a limited amount of information on the implications of audio delivery devices. Laney et al. (2010) noted the use of headphones to play personal audio in a multi-touch musical interface, however there is no previous research which performs a direct comparison of headphone and speaker based audio presentation for public and private audio within the context of collaborative music-making. Blaine and Perkis (2000) also compared headphones and speakers through informal user testing, and suggested that headphones caused their participants to be less communicative and feel more isolated from the group. They also argued that spatial delivery of audio might help individuals identify their own contributions, although their results showed that non-musicians still encountered problems identifying the changes they were making to the system. Their findings were based on visual measures such as body movement, gestures, and 'general levels of excitation', rather than changes in interaction caused by the form of audio delivery.

### 2.3.2 Summary of CSCW

Computer Supported Cooperative Work is a broad and well established field. At the time of writing, the ACM Digital Library returned 11,406 unique results for a search on 'CSCW'. This section has highlighted the discussions and issues within CSCW which are considered most relevant to the development of this thesis. In particular, the concept of Awareness plays a central role in the three studies presented in Chapters 5, 6 and 7. This section has also highlighted the concepts of congruence, privacy, and territory, and discussed some of the distinctions between remote collaborative systems, and systems designed for co-located interaction. The following section surveys the current state of research concerning collaborative music systems.

## 2.4 Collaborative Musical Interfaces

A recurrent theme within the Computer Music community has been the creation of physical interfaces and software environments for collaborative musical interaction. Although some acoustic instruments can be played by multiple people (e.g., the piano) (Makelberge, 2010), the ability for computer based instruments to run parallel processes and provide multiple channels of simultaneous control makes support for multiple users a logical developmental progression (Jordà, 2005b). Arguably much of the work to date has focused on building novel software applications, such as presented by Bryan-Kinns (2004), Gurevich (2006), Miletto et al. (2006) and hardware interfaces (Fels and Vogt, 2002, Jordà et al., 2007, Weinberg et al., 2002), software infrastructures (Burk, 2000, Wyse and Mitani, 2009) and collective interactive sonic experiences (Fencott and Bryan-Kinns, 2009, Wozniewski et al., 2008).

Although these technological innovations are undoubtedly a vital stage in expansion of computer supported group musical interaction as a research domain, to date there has been less research directed towards the study of human interaction in software mediated musical collaboration. The previous section has discussed how concepts such as awareness, territory and privacy have been investigated by CSCW researchers looking to develop collaborative workplace tools. Focus on these currently under explored issues within musical interaction may result in a more clear picture of how to design software to support real-time musical collaboration, and shed light on the shortcomings of existing collaborative music software. In a related problem, it is also important to stress that there currently exists no established experimental or methodological paradigm for investigating these features within creative domains such as music. A key problem, as will be demonstrated in Chapter 3, is that musical interaction is characteristically less task-constrained and product-orientated than the workplace activities typically studied in CSCW literature.

### 2.4.1 Classifications of Collaborative Digital Musical Interaction

There have been various attempts to categorise systems for collaborative digital musical interaction. Such classifications provide a means of comparing the properties and affordances of different systems, and may also suggest insights into new design or evaluation techniques. Weinberg (2005) presented a theoretical framework for classifying types of interconnected musical networks, based on their system architecture, network topology and style of interaction. This framework considers the types of network configuration, the structure of the activities undertaken using the network, and the types of parameters made available to users, however it does not consider what information should be presented to users in order to support interaction, collaboration or engagement. Barbosa (2006) presented another technology centric classification, the Networked Music Classification Space, as a means of identifying musical CSCW applications based on the axes of local/remote and synchronous/asynchronous. Although this framework represents a thorough classification of 'CSCW systems for musical applications', the primarily technology-centric view of musical collaboration does not account for the issues related to interface design issues such as awareness, or how people might use such interfaces in practice.

Blaine and Fels (2003) proposed a detailed classification of collective musical interfaces which is more focused on user interaction. This classification identifies ten dimensions of comparison: Focus, Location, Media, Scalability, Player Interaction, Musical Range, Physical Interface, Directed Interaction, Pathway to Expert Performance and Level of Physicality. These are summarised in Table 2.2. Within this thesis, the most relevant dimensions of Blane and Fels' framework are those of *Media* and *Player Interaction*, both of which are manipulated in the experimental studies presented in subsequent chapters. Other dimensions, such as *Scalability*, *Location*, and *Focus* are discussed later as pertinent avenues of future work. Similarly Jordà (2005b) identified *number of users* and *user number flexibility* as crucial parameters for a multi-user musical instruments, noting that some multi-user instruments can also be played by a single person. Jordà (2005b) also established interdependencies between users as an important consideration, suggesting that environments which do not allow users to influence each-others contributions might not be usefully regarded as multi-user.

| **Feature** | **Description** |
|---|---|
| Focus | if the interaction is purely between the musicians, or if it also includes an audience. |
| Location | does the interaction happen in a specific place (e.g., in the case of an interactive installation) |
| Media | Is the interaction entirely auditory, or is there a visual component. |
| Scalability | The size or scale of the system (e.g., three people using a shared interface, or hundreds via the internet) |
| Player Interaction | The type of interaction musicians engage in (e.g., a multi-touch tabletop, a computer interface) |
| Musical Range | The range of notes available to musicians. |
| Physical Interface | The kinds of sensors or physical devices used to create music. |
| Directed Interaction | Is the interaction directed by an experienced user or conductor, or is it entirely open and free. |
| Pathway to Expert Performance | Does the interface/instrument allow users to develop virtuosic skills? |
| Level of Physicality | How physical is the interface (e.g., a full body suit, a keyboard, or brain-wave sensing) |

Table 2.2: Blane and Fels' 'Contexts of collaborative musical experiences'

### 2.4.2 Design Recommendations for Collaborative Musical Interaction

A number of researchers have contributed design guidelines for musical interfaces, although there is currently only a limited number of design recommendations specifically for multi-user systems. Primarily concerning single user interaction, Cook (2001) proposed a set of influential design principles for hardware controllers, such as the idea of designing interfaces for performances and compositions, rather than as generic tools. More recently, Murray-Browne et al. (2011) outlined a new set of design principles to inform the design of novel performance interfaces, however these do not address design requirements for multi-party or collaborative interaction. Similarly Overholt (2009) proposed a design space for musical interfaces, and discussed seven design principles for designing musical interface technology, however this framework primarily addresses the issue of expressivity, and does not consider in great detail the roles or challenges of multiple users in a musical interface.

Barbosa (2006) considered the importance of reducing latency within internet based musical collaboration, and although this work draws on previous CSCW research as a point of contextual reference, it does not advance the understanding of issues such as awareness within collaborative musical interaction. Weinberg (2003, 2005) and Weinberg et al. (2002) proposed a number of ways in which shared musical experiences for novice users can be developed, however this work does not easily translate into implications for the design of systems for more experienced users.

Based on a survey of literature related to computer based musical interaction and tabletop collaboration, Xambó et al. (2011b) used the popular 'Design Pattern' format derived from Alexander (1978) to propose four 'Candidate Design Patterns' for collaborative multi-touch environments. These Design Patterns were designated *candidate* as they were based on literature searches rather than primary data such as user studies. As the patterns are not based on observed data, they address hypothetical design situations and are left intentionally vague so therefore act as guidelines, rather than claiming to be concrete solutions to known design challenges. The patterns are: *shared and personal spaces, learning and fun, map of actions, divide and conquer*. Whilst these patterns do address some of the issues raised previously in this chapter, such as the need to support mixed focus interaction (discussed by Xambó et al. (2011b) as the notion of divide and conquer), the patterns are very generalised, and could potentially apply to non-musical collaborative interaction design.

### 2.4.3 Co-Located Musical Interaction

During the twentieth century, composers such as John Cage and Karlheinz Stockhausen began to write pieces which called for more tightly interwoven interactions between groups of performers and the processes of sound creation using electronic instruments and devices (Jordà, 2005b). Cage's *Imaginary Landscapes* is often cited as the first instance of a networked musical performance (Renaud et al., 2007). In this piece performers were required to jointly manipulate the volume and frequency controls on a radio tuner to create an aleatoric composition. Similarly Stockhausen's *Mikrophonie* compositions (Maconie, 1989) set out precise instructions for a group of performers to collaboratively explore and electronically manipulate the sonic properties of a large gong-like instrument.

Such pieces can be described as examples of Live Electronics, where elements of live performance are combined with electronic sounds, which may be generated in real-time, or may be played back from a pre-recorded medium such as tape. Stroppa (1999) discusses many topics surrounding live electronics, including the relationship between the real-time generation of audio and the relation this has to the potential for musical expressivity in performance. Davies (2001) notes that the practice of improvised performance incorporating live electronics emerged in part due to a lack of formalised notation for devices such as filters and oscillators, and also due to more experimental approaches to music-making becoming popular during the 1960s. The development of Live Electronics in the mid 20th century was an important step forward within the field of electronic and experimental music, and paved the way for the use of computer software within group musical performance.

The League of Automated Composers and the Hub (Gresham-Lancaster, 1998) were seminal proponents of networked musical performance using computers (Barbosa, 2003, Hutchins, 2011, Jordà, 2005b, Weinberg, 2005, Wyse and Mitani, 2009). Members of these early computer music groups wrote music generating programs which communicated via serial interfaces and bespoke messaging protocols. These two groups were among the first documented attempts to design for computer mediated group musical interaction. Similarly Interface (Bahn and Trueman, 2001) explored duet musical performance with computers using novel physical interfaces to place an emphasis on 'group interaction' rather than interaction with the computer, and in 1995 Sensor-Band explored collaborative musical interaction via interconnected physical interfaces in their SoundNet composition (Jordà, 2005b).

Protocols such as MIDI and Open Sound Control (Wright et al., 2003) have facilitated a number of Laptop orchestra and Mobile Phone Orchestra projects which have explored co-located technologically mediated musical interaction using mobile computing devices. Key examples are the Stanford Laptop Orchestra (Slork) (Wang et al., 2009) and Mobile Phone Orchestra (MoPho) (Oh et al., 2010, Wang et al., 2008), the Princeton Laptop Orchestra (PLOrk) (Trueman et al., 2006) and The Carnegie Mellon Laptop Orchestra (CMLO) (Dannenberg et al., 2007). Returning to the classification of collaborative music environments presented by Blaine and Fels (2003), it is important to stress that the concept of a laptop orchestras tends to suggest a formalised performance context featuring live musical interaction before an audience. Publications related to these projects often consider the technical implementation issues of creating these distributed musical ensembles, discuss the technological affordances of the mobile devices which are employed, and document in great detail the musical compositions and performance created for the ensembles to play. However, perhaps due to their performance centric conception of musical interaction, literature related to laptop orchestras has devoted less time to discussion of the human interaction between the musicians themselves. An exception is the work of Trueman et al. (2006), who discussed the use of gesture, hand signals and cue cards in co-ordinating the Plork as a solution to reducing network traffic. Hutchins (2011) also stressed the importance of gaze and listening as part of laptop ensemble performance. In considering the MoPho repertoire, Wang et al. (2008) noted the spatial location of players as influencing the group interaction, and also created a distinction between compositions which are conducted, scored or improvised. Trueman et al. (2006) also considered the role of the conductor, and noted that the computer network can play many different roles in facilitating communication and co-ordination between musicians, ranging from coarse communication such as text messages to low level control over musical timing.

In attempting to emulate the physical properties of acoustic instruments in traditional ensembles, the importance of coupling the performer with a localisable sound source has been a central feature in the aesthetics of laptop and mobile phone orchestras (Oh et al., 2010, Trueman et al., 2006, Wang et al., 2009). Oh et al. (2010) considered several different methods of attaching portable speakers to performers in a mobile phone ensemble, discussing in particular practical and logistic issues, as well as the association of agency (in this case touch based interaction on mobile phones) with sound sources. However while devoting a great deal of time to the engineering and technical aspects of creating personalised speaker setups for networked musicians, these bodies of work have not evaluated the effects of localised audio on the interaction between collaborating musicians. Furthermore, Merritt et al. (2010) observed that the two groups of laptop musicians studied in their evaluation of a visual awareness mechanism for co-located musical interaction rejected the idea of personal speaker channels, suggesting that the way audio is presented is a subjective issue amongst electro-acoustic musicians.

There is a tendency for laptop orchestra projects to be focused on the development of software for specific performances and compositions, rather than generalised collaborative musical environments. This demonstrates an interesting distinction between some networked music research, and CSCW systems such as shared document editors which are designed for general purpose tasks. Such collaborative systems do not lend themselves to the production of specific works or outcomes, whilst musical environments created for the realisation of a specific work may lead their users towards more constricted outcomes, and are by definition limited in their scope outside of their intended purpose. Systems which have been designed for the realisation of a particular composition have been described as 'composed instruments' (Bown et al., 2009, Murray-Browne et al., 2011), as the interface becomes a central component in the composition itself. Such interfaces may also be imbued with assumptions about musical theory and representation (Magnusson, 2006, 2009). In comparison, whilst a general purpose tool such as a shared document editor will contain assumptions about the kinds of documents which it can be used to make, it does not lead users towards to creation of a specific textual work.

Another trend in computer music performance is 'Live Coding' (Collins et al., 2003), a performance practice which involves writing or modifying music generating computer code in front of an audience. Although Live Coding has gained publicity in recent years, Collins et al. (2003) stated that the early computer music ensemble the HUB (discussed in Section 2.4.3) employed live-coding style practices while developing their compositions. While some live-coders perform solo, there are many who choose to perform as a group, and specialised collaborative live-coding environments have been developed, such as *Co-Audicle* (Wang et al., 2005, 2006), a tool that allows groups to share and jointly edit pieces of code across networked computers. Collaborative Live Coding is an interesting case when considering awareness, as compared to playing acoustic instruments, the activity of live-coding (typing code at a laptop) contains very few gestures or physical indications of what a musician is doing or what sound they are creating, however for those with an understanding of the languages, programs and structures in use, being able to see the code as it is typed and modified may provide detailed and nuanced insights into the activities of a collaborator.

Aside from Laptop Orchestras there are other projects based around musical interaction using networked computer workstations. Wyse and Mitani (2009) presented a generalised software infrastructure for networking arbitrary software and devices, but offered no evaluation of this technology. Rebelo and Renaud (2006) presented the Frequencyliator, a software system which attempts to facilitate co-located or remote musical collaboration between musicians using laptop computers. The Frequencyliator does so by imposing constraints on the musical structures available to musicians. These structures include rhythmic pulse, performative roles, and the ability to anticipate events. However Rebelo and Renaud (2006) did not include any detailed information about how musicians used the system during composition or performance.

### 2.4.4 Live Algorithms and Collaboration With Computers

'Live Algorithms' is a field of computer music research concerning human interaction computer applications which participate in creative performance through autonomous behaviours, impulsive responses and generative processes. Although this thesis does not directly concern the design of such systems, it is important to acknowledge this field as an example of computer supported collaboration, where the human player(s) engage in collaboration with the computer as though it were another member of the ensemble. An example of a live algorithm is the B-Keeper drum tracking system Robertson et al. (2008), which follows the tempo and rhythm of a human drummer. Another example is the Swarm algorithm proposed by Blackwell and Bentley (2002), and demonstrated improvising in real-time with a human vocalist. Finally, Band-in-a-Box [2] is an example of a commercially available application which promotes collaboration between people and computers through generative musical processes.

### 2.4.5 New Interfaces for Musical Expression

The availability of low cost sensor and micro-controller technologies, advanced computer game controllers (Kiefer et al., 2008) and high power personal computers has facilitated a new generation of collaborative musical interfaces which represent a departure from traditional screen, keyboard and mouse based interaction. The New Interfaces for Musical Expression conference

[2] http://www.pgmusic.com/ Last Accessed 2 October 2012

series[3] commonly features such interfaces. By way of comparison to the musical networks discussed later, in Section 2.4.7, many of these new interfaces feature a single shared device, present users with musical controllers built from unconventional or modified physical objects, or do away with the physical interface entirely. As well as challenging the conventions of the standard computer interface, some projects falling into this category place an emphasis on co-located interaction where intimacy and co-dependant connections between participants are created. For instance the Tooka (Fels et al., 2004) is a two player wind instrument which requires two musicians to share and jointly manipulate the pressure of a volume of air contained within a tube. Successful musical performances or interactions with the Tooka require participants to form a repertoire of tightly coupled, intimate and inter-dependant gestures, while maintaining ‘mutual awareness’ (Fels, 2004).

The BeatBug Network (Weinberg, 2003, Weinberg et al., 2002) is a co-located physical interface designed explicitly for multi-party interaction. Using this interface, musicians or novice users manipulate and share short drum loops using a novel sensor based hand-held controller. The drum patterns are modified by a stochastic process, before being transmitted to other players in the group.The stochastic nature of the music making process and the aleatoric methods used to share and exchange musical material between participants limits the BeatBug Network as a tool for serious musical collaboration, however it must be stressed that the system is designed primarily for non-musicians, and as a tool for workshop activities with groups of children.

### 2.4.6 Touchscreen and Tangible Interfaces for Collaborative Music

Just as touch screens have become a major focus for Shareable Interfaces in CSCW research (see Section 2.2.5), touch surfaces for musical interaction, performance and composition are a rapid field of expansion (Xambó et al., 2011b), with touchscreen mobile telephones and tablet computers becoming commonplace tools for musicians (Oh et al., 2010). At a larger physical scale, interactive table interfaces such as the reacTable (Jordà et al., 2005) have captured a great deal of public attention. The reacTable allows musicians to collaboratively patch together sound generators and processors by manipulating and arranging small physical objects on a rear-projected tabletop interface. These physical objects represent different sound generators and processors; with position, rotation, and proximity to one-another mapped to various synthesis parameters. This physical interaction metaphor is intended to provide both an intuitive experience for musi-

[3] http://www.nime.org/

cians and a compelling spectacle for an audience by providing a visual connection between the performer's physical gestures and the music being created. Similar tangible object based interfaces for music-making and musical interaction include BlockJam (Newton-Dunn et al., 2003) and Audiopad (Patten et al., 2002).

Purely touch-based music environments have also been developed, including those presented by Davidson and Han (2005), Fencott (2008), Iwai (1999), Xambó et al. (2011). Davidson and Han (2005) stressed that purely touch based interaction provides more flexibility than interfaces based on tangible objects, as touch-only allows for dynamic re-contextualisation of the graphical interface. Furthermore the interface components are not limited to the physical size constraints of the physical objects, and the screen does not need to be horizontal. However while noting the potential for the multiple-user interaction, Davidson and Han (2005) focuses on technological concerns, rather than social or collaborative aspects of multi-touch technology.

Regardless of whether these systems use direct touch or tangible object based interaction, a key feature of interactive surface interfaces is the support for multiple points of interaction by one or more simultaneous users (Davidson and Han, 2005, Klügel et al., 2011). This calls not only for consideration of musical interaction metaphors, but also a sensitivity to the nature of human collaboration and in particular issues of awareness, territory, sharing and privacy, as discussed in Section 2.1. Acknowledging this fact, Klügel et al. (2011) drew extensively on studies of collaboration to inform the design of a collaborative multi-touch music interface. Using CSCW literature, the paper identified a number of constraints and requirements for collaborative interaction; *group awareness*, *group articulation* and *tailorability*. A key contribution this paper makes is to propose that non-spatial interaction metaphors which operate independently of the orientation or position of interface elements are preferable in a collaborative context to mapping strategies which rely on the spatial position of objects, as the non-spatial designs can more effectively support the dynamic coupling of users. They also acknowledge the role of awareness in collaboration and suggest that interfaces should support tailorability by allowing adoption and use in personal and unintended ways.

Aside from the interactive surface interfaces based on conventional paradigms of computer music software (on-screen oscillators, musical keyboards, sliders and so on), multi-touch interfaces are ideal for placement in public contexts where accessibility and immediacy are central concerns (Benko et al., 2006). Examples of interactive surfaces designed for playful engagement or public exhibition include the interactive cellular automata presented by Fencott (2008) the Scambi Surface Sequencer (Fencott and Dack, 2011), and Composition on the Table (Iwai, 1999), all of which leverage the potential of interactive surfaces to support direct intuitive engagement with sonic and visual materials in a manner which is distinct from conventional music-making techniques and tools.

Xambó et al. (2011a) argues that such interfaces are better suited to collaboration than both single user computers and interconnected musical networks, as the 'interaction with digital data is more embodied and social'. While this statement may have merit, Xambó et al. (2011a) does not support this claim though a direct comparison of the collaborative affordances of touch tables and other interfaces for musical collaboration. Although the work in this thesis does not deal directly with touch-screens, tangible interaction or tabletop displays, these projects are included due to the similarities they share with other aspects of the work studied in this thesis. In particular touch-screen interfaces support real-time co-located interaction, and multi-touch musical interfaces allow multiple users to simultaneously interact with, and contribute to a shared musical or sonic product. The spatial and territorial aspects of table interfaces are also reflected in the interface designs used for the studies presented in later parts of this thesis.

Another crucial consideration is the physical orientation and configuration of the shared interface, and how this supports group interaction. The reacTable was deliberately designed as a circular table, so as to create a situation where there is no 'head position or leading voice, and with no privileged points-of-view or points-of-control' (Jordà et al., 2007). While not discussed in his publications, photographs of Davidson's interface show it in a diagonal orientation, rather than as a horizontal table-top, and the multi-touch platform is described as a 'drafting table' (which is typically tilted, rather than flat). This has implications for how a group of people might interact with and orient around the interface. A slanted table clearly privileges those stood facing the front, and many of the graphical widgets (e.g., the on-screen musical keyboard) privilege a user who is facing them in the correct orientation, and although the widgets could potentially be rotated and re-sized for use by someone standing at a different side of the drafting table, they may not promote the same degree of collective interaction as the reacTable pucks which are independent of orientation. Finally, the physical nature and arrangement of the reacTable pucks presents all users with a persistent account of the system state and makes changes immediately visible. On other hand, on-screen widgets could potentially become minimised or obscure each other, therefore hiding certain information about the system.

### 2.4.7 Geographically Distributed Musical Interaction

While long-distance real time musical collaboration had been demonstrated by the Hub in the 1980s (Gresham-Lancaster, 1998, Renaud et al., 2007), the increasing uptake of the internet made larger scale collaboration over vast geographic distance a feasible (Alexandraki and Kalantzis, 2007, Renaud et al., 2007) and a popular avenue for exploration (Jordà, 2005b). Barbosa (2003), Renaud et al. (2007) and Mills (2010) presented surveys of remote internet based collaboration. Examples of geographically distributed musical interaction include Daisyphone (Bryan-Kinns and Hamilton, 2009), Jam2Jam (Brown, 2010) and WebDrum (Burk, 2000). The Daisyphone is a desktop computer or mobile phone based application for informal musical collaboration over the internet. Similarly, WebDrum is an internet based drum sequencer written in Java to allow remote collaborators to create simple drum loops. Jam2Jam is a collaborative music environment which allows groups of people to create music by manipulating algorithmic musical units in a process Brown (2010) defined as *Network Jamming*. In an education context Akoumianakis et al. (2008) presented a prototype distributed musical collaboration and learning environment to facilitate on-line and off-line activities which provides support for mutual awareness.

While the impact of latency (the time taken for information to pass between computers) in a local network can be reduced to low levels, when data is travelling via the internet latency becomes much more noticeable, and as such must be taken into account. Furthermore, real-time audio streaming has been identified as one of the most intensive networking applications (Alexandraki and Kalantzis, 2007). One way to lower latency is to use low bandwidth protocols such as MIDI to transmit musical information (Bryan-Kinns, 2004, Burk, 2000), while more recently high-speed internet connections have become capable of reliably transmitting compressed audio data with minimal delay (Renaud et al., 2007). Barbosa (2003) stressed that rather than viewing latency as a technological limitation, it should be celebrated as an inherent property of distributed online musical collaboration.

There are also a number of online collaborative systems and services with a commercial or non-research origin. Some of these these programs are attempts to create collaborative environments which function in a similar way to popular single user music composition applications typically referred to as Digital Audio Workstations (DAW). DAWs typically allow sounds and musical elements to be placed in sequence along a time-line, and also feature a range of audio editing tools, software synthesisers and audio processing effects such as reverberations, distortion and filtering. Examples of single user DAWs include Cubase[4], the popular GarageBand application [5] and Pro Tools[6].

Commercially available collaborative applications include ResRocket[7] (Barbosa, 2003), Digital Musician[8] eJamming[9] and NinJam[10]. These services typically allow users to compose music together in real time or asynchronously, or provide a means for musicians to play or rehearse music in real-time using streaming audio. Where musicians play together in real-time they may be provided with a video or audio channel for communication. Renaud et al. (2007) presented a more in-depth technical analysis of many such systems, placing emphasis on approaches to dealing with network latency. By contrasting novice and experienced users, James and Stanton (2011) draw on CSCW literature to discuss the affordances of asynchronous and synchronous systems for distributed musical collaboration. The paper stated that systems for asynchronous collaboration may be advantageous for non-musicians who need more time to formulate their musical contributions, while synchronous systems may be more suitable for the immediate testing of ideas and negotiation over the musical direction of a composition.

Asynchronous composition systems such as Digital Musician enable musicians to collaborate over wide geographic distances by means of a centralised server which holds a shared record of the composition. Some services, such as Digital Musician, also provide a social networking infrastructure for musicians to connect with each other, initiate collaborations and share their work. Other examples of collaborative Digital Audio Workstations (DAW) include 'Ohm Studio'[11](in Beta testing at time of writing), and the open source LNX Studio[12]. At the time of writing, the forthcoming BitWig DAW advertised a collaborative function [13].

### 2.4.8 Awareness Research in Collaborative Music

Awareness mechanisms were a concern for the CODES project (Miletto et al., 2006), an asynchronous web-based 'music prototyping' system for novice and non-musician users. CODES

[4]http://www.steinberg.net/en/products/cubase/start.html Accessed 30 September 2012.
[5]http://www.apple.com/ilife/garageband/ Accessed 30 September 2012.
[6]http://www.avid.com/us/products/pro-tools-software Accessed 30 September 2012.
[7]http://www.jamwith.us/about_us/rocket_history.shtml Accessed 7 June 2012
[8]http://www.digitalmusician.net/index.php Accessed 7 June 2012.
[9]http://www.ejamming.com/ Accessed 7 June 2012.
[10]http://www.ninjam.com/ Accessed 7 June 2012.
[11]http://www.ohmstudio.com/ Accessed 22nd May 2012
[12]http://lnxstudio.sourceforge.net/ Accessed 22nd May 2012
[13]http://bitwig.com/bitwig_studio.php Accessed 22nd May 2012

addresses the issue of awareness by providing users with information about the modifications, past actions and motivations of other uses. The results presented by Miletto et al. (2009) and Pimenta et al. (2011) using HCI methods such as think-aloud observations and task execution challenges followed by questionnaires, heuristics and likert scales suggest a general approval for the system, which is 'intuitive and easy to use'. The study did not, however, address in detail the collaborative features of the system, and consequently the findings of a more detailed evaluation would be beneficial to the wider research community.

Merritt et al. (2010) developed and evaluated an audio visualisation system to support laptop based group musical improvisation. The system uses audio visualisation techniques such as sonograms to provide musicians with real-time awareness information about the sounds being generated by other members of the laptop ensemble. A sonogram is a visualisation of sound in which time is represented along the x-axis and frequency is represented from low to high along the y-axis. Colour is plotted on the graph to represent the amplitude of frequencies at particular points in time. Sonograms therefore provide a highly accurate visual representation of complex sounds. Figure 2.4 shows the sonogram of a piece of electronic music.

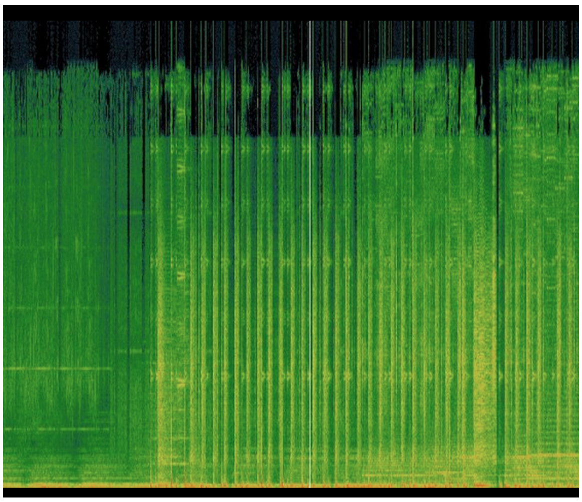


Figure 2.4: A piece of electronic music, represented as a sonogram.

Merritt et al. (2010) used a participatory approach in the development of their visualisation system. In this approach groups of laptop musicians were observed in performance and interviewed about their needs and requirements. The software was then developed through several prototype iterations, and evaluated through contextual interviews with the musicians. However, their work did not make an explicit connection to issues of awareness in CSCW research.

Gurevich (2006) investigated the issue of privacy in JamSpace, a music environment in which users have access to a hierarchy of workspaces, each showing an increasing level of information to fellow collaborators. Although the JamSpace project investigated levels of privacy within multi-party music-making, it did not focus specifically on the issue of inter-participant awareness, and also differs from the focus of this thesis, as it considers privacy at the level of the individual, rather than at the lower level of control over specific musical contributions within a collaboration.

Gates et al. (2006) used the Workspace Awareness framework (Gutwin and Greenberg, 2002) to describe the kinds of awareness information DJs utilise when playing in night clubs, and describes in detail how this information is gathered. They suggested that the primary means employed by DJs was observation of their audience, through which the DJ is able to monitor the level of audience engagement, and discover which records, or types of music the audience is most likely to respond positively to. Their paper proposed a number of design implications for creating new awareness gathering technologies for DJs, stressing that such awareness mechanisms should not add to the DJ's cognitive load, not interfere with their working practice, and not allow too much direct audience involvement. Pfadenhauer (2009) also employed ethnographic techniques to discuss the role and activities of DJs in nightclub environments although this work did not lead to a discussion of technology or redesign implications.

### 2.4.9 Interfaces for Non-Musicians

Another recent trend is the development of technology which allows non-musicians to engage in musical interaction. These systems and experiences are often playful and more approachable than traditional musical instruments. A characteristic of such interfaces is a shallow learning curve or 'low entry fee' (Fels, 2004) which enables the user to learn the interface quickly, but at the expense of offering limited challenges. Work which falls into this category includes 'sound toys' (Robson, 2001), which are music producing objects that do not resemble musical instruments. Overholt (2009) claimed that interfaces which do not present a challenge to their users may be more similar to interactive installations which are designed for use in a single visit for a limited amount of time. Weinberg (2003) and Weinberg et al. (2002) presented several collaborative musical systems designed for the 'lighter novice-oriented' end of the user axis, and which were perceived by users more as 'toys' than 'serious tools'. Blaine and Perkis (2000) proposed an electronic drumming circle to support and encourage novice users in creating music collaboratively, whilst the Daisyphone (Bryan-Kinns, 2004) is an online collaborative jamming system for novice users.

Music based based computer games are currently a popular genre (Pichlmair and Kayali, 2007). Examples include Guitar Hero[14], and Dance Dance Revolution[15]. Guitar Hero is a musical game (Overholt, 2009) which employs a physical controller that visually resembles a conventional electric guitar, however playing Guitar Hero is quite unlike playing a guitar as the player is required to press buttons in a pre-determined sequence displayed on a screen. It is also interesting to note that multi-player interaction in Guitar Hero is in part competitive and offers little room for individual or group creativity. Dark Circus (Hadrup et al., 2004) is another example of collective sound mediated gameplay. Dark Circus is a team based strategy game in which players are placed in a darkened room and use sound (as opposed to vision) as the primary medium for co-ordination.

Museum and gallery based interactive musical exhibits which seek to engage groups of non-musicians through casual, low-impact activities can also be included under the heading if interfaces for non-musicians. Examples include Jam-O-World (Blaine and Forlines, 2002) and WorldBeat (Borchers, 1997), both of which allow participants to create music using easily learnable interfaces. Finally, there are a number of mobile and pervasive experiences which engage groups of people in digitally mediated musical activities on the city streets. Examples include the work of Wozniewski et al. (2008), Sensory Threads (Fencott and Bryan-Kinns, 2009) and Malleable Mobile Music (Tanaka, 2004). Within this category, Sensory Threads relates most closely to this thesis as it concerns design issues for co-located group interaction and considers the provision of awareness information about other people's activities within a collective, interactive data sonification.

While designing for non-musicians is an interesting avenue of research, with wide-reaching applications (e.g., entertainment, music therapy and education), this thesis focuses on studying features which might be appropriate for more professional level music tools, and for people with existing musical skills and domain knowledge.

## 2.5 Evaluation of New Interfaces for Musical Expression

Within CSCW there are established methodologies for performing evaluation and user testing. However evaluation is a problematic issue for research focused on domains such as music, where

[14]http://www.guitarhero.com/ Last Accessed 7 June 2012
[15]http://www.ddrgame.com/ Last Accessed 7 June 2012

the target activities does not have clearly defined goals and outcomes, and where highly subjective interpretations of the interaction or experience make a strong contribution to the perceived success or quality of the interaction design. The problem of evaluation is further compounded in domains characterised by high levels of creativity and spontaneity, and where there exists multiple (or infinite) solutions to a single problem or task. Terry et al. (2004) used the notion of the 'ill-defined problem' to describe this situation, emphasising the role multiple solutions play in creative design processes. This is discussed further in Chapter 3.

A current drawback to research from the NIME community is that these new musical interfaces are often used by a small number of people, perhaps only those involved in its development. This means there is currently a limited pool of information about how such systems should be designed, and how they might be used. The NIME community therefore requires more developed evaluation methodologies to explore musicians' and audiences' experiences with these interfaces. Furthermore, the use of bespoke hardware and custom electronics may make the interfaces difficult for other musicians to re-create, use and re-appropriate for their own needs. Without the widespread adoption or use of these new interfaces it is difficult to assess their potential to change the way music is played, composed or performed.

There have been various attempts to explore and classify engagement with interactive experiences. Within the context of public performance with interactive technology, the Wittingness Framework (Sheridan et al., 2007) considers the transitions between bystander, participant and performer, taking into account the development of skills, and the acquisition of knowledge about the interactive system. Morrison et al. (2007) proposed the 'Lens of lucid engagement' to evaluate interactive artworks by observing how they are used and interpreted by participants. Brignall and Rogers (2003) discussed how people gather around, express interest in, and ultimately engage with large interactive public displays. Similarly, Hornecker and Stifter (2006) presented observations of visitors using interactive exhibits in a science museum. Research from the Creativity and Cognition Studios (CCS)[16] has attempted to explore more detailed levels of interaction with digital artworks, often through detailed video observation and by conducting interviews with participants after they have experienced the work. Costello and Edmonds (2007) asked participants to indicate if they experienced various emotions while interacting with an exhibit, while Bilda et al. (2008) proposed a model of interaction which described how people explored and understood an

[16] http://www.creativityandcognition.com/ Accessed 22 May 2012

interactive system. Freeman (2008) used a similar approach, by employing short questionnaires with audience members at an interactive performance, while Morrison et al. (2008) proposed a multi-model approach, combining discussions, interviews and observations. A key criticism of these questionnaire and interview approaches when used in isolation is their post-hoc nature, which is based on the subjective memories of the participants.

Quantitative questionnaire based approaches have also been adopted. These approaches include collecting subjective ratings using questionnaires and Likert scales (Hsu and Sosnick, 2009, Laney et al., 2010, Unehara and Onisawa, 2003). Here, participants rate factors such as the quality of the music created, their level of enjoyment, or the degree to which they have control over the musical output. However, Likert scale measures have previously been identified as problematic for the evaluation of musical interfaces (Stowell et al., 2009) because of the subjective quality of the ratings, and the arbitrary nature of the scale itself.

Wanderley and Orio (2002) proposed a quantitative HCI derived ‘musical task’ approach to evaluation, through the comparison of interfaces based on their support for executing tasks such as playing notes, performing gestures (e.g., trills, vibrato), and playing musical phrases. Kiefer et al. (2008) used a variation of this approach to compare several musical interfaces. While Hsu and Sosnick (2009) noted that the methodology proposed by Wanderley and Orio (2002) could be useful in the evaluation of input devices, Stowell et al. (2009) argued that by focusing on the execution of simple musical tasks, the HCI approach potentially ignores the creative and emotional aspects of music-making. Furthermore, considering musical interaction, it can be argued that assessing the ability of participants to execute a task does not give a very clear indication of how effective an interface is. For instance Candy (2007) noted that musicians often work within constraints which might appear to hinder their performance, but which in practice provide a fertile ground for experimentation and creativity. Such constraints could include obeying a set of compositional rules or working within a musical framework such as the Western tonal tradition. Conventional instruments also apply limitations on the performer, such as the range of notes which can be produced, the speed at which the instrument can be played, and the number of timbres which are available, yet even instruments which appear to be highly constrained have remained popular for hundreds of years. It could therefore argue that the success of a musical interface cannot be measured simply on the basis of how well it supports the execution of particular activities, but should perhaps consider less easily measurable parameters, such as how expressive it is (Dobrian and Koppelman, 2006), or how creative it makes the musician feel. Finally, when thinking about collaborative musical interaction, factors such as time taken to reach a conclusion or solution may be less important because musical interaction is often loosely structured (Nabavian, 2009), and may not have a specific goal or easily delineated outcome (Holland, 1999).

Bryan-Kinns (2012) drew on CSCW research to inform the design of controlled studies investigating distributed group music-making using a novel collaborative application called the Daisyphone (Bryan-Kinns et al., 2007). The Daisyphone studies focused on the concept of *Mutual Engagement* as a means of understanding group creativity on the internet. This was explored through the controlled manipulation of simple awareness features within the interface to determine their impact on the way participants use the software. These studies used questionnaires, and interaction measures such as the number of contributions made, the amount of editing and degree of co-editing, as well as spatial arrangements of the circular sequencer interface. The results presented by Bryan-Kinns and Hamilton (2009) suggested that allowing musicians to annotate the shared workspace increased the level of mutual engagement between people, as did providing colour coded indications of who was responsible for which contribution. However there are a number of limitations to the Daisyphone studies. Firstly, the studies are all based on groups of non-musician users, rather than people with established musical backgrounds. Secondly, the Daisyphone does not resemble the interfaces typically used in conventional music software, and thirdly, the Daisyphone offers users a very restricted range of pitches and sounds with which to compose. These issues make it difficult to transfer the study findings into design guidelines for more complex collaborative software environments.

Truman (2011a) used similar interaction log metrics to evaluate learning in a group music composition activity. Her approach combined quantitative interaction log analysis and questionnaire data to explore the effect interface support for several aspects of group creativity has on learning in a collaborative music environment. This analysis was performed in relation to a three part framework to characterise group creativity in terms of personal and social aspects of *preparation*, *generation* and *evaluation*. However, as with the Daisyphone, the interface used by Truman (2011a) was very simplistic, and designed for children. This limitation makes it difficult to draw firm conclusions about how the study findings might be applied to the design of interfaces for adults, or experienced musician users.

Analysis of interviews and dialogue has featured in the evaluation of musical interfaces. Wilkie et al. (2010) proposed a 'conceptual metaphor based evaluation' to analyse the language musicians used to describe aspects of music, and discover the metaphors which underpinned their understanding of musical concepts. In this approach, a group of musicians answered questions about the written score for a piece of music. The discussion was video recorded, transcribed, and coded to identify the metaphors used by the musicians. Existing music applications were then analysed for evidence of these metaphors in their interaction design. Wilkie et al. (2010) argued that this approach can help designers identify usability issues, design tradeoffs and interaction design improvements. One critique of this approach is that the metaphors identified were derived from discussion of a written musical score, rather than audio or musical interaction. This is problematic because a score is a set of instructions for the realisation of music, rather than the music itself (Small, 1998). Furthermore, a score contains its own set of representations metaphors (e.g., time as a horizontal continuum, pitch as vertical position), which may have influenced or directed the musicians to use metaphors appropriate to score based music, rather than encouraging them to discuss the metaphors they personally use when composing, performing or interacting with music software.

Discourse Analysis was suggested by Stowell et al. (2009, 2008) as a means of identifying similarities and common themes within participants descriptions of a musical interface for the purpose of understanding their conceptualisation of it. For similar reasons, a number of researchers have used Grounded Theory methodology (Muller and Kogan, 2010) to evaluate musical controllers and interfaces (Kiefer, 2010, Laney et al., 2010,?, Michael Gurevich and Marquez-Borbon, 2010, Morrison et al., 2007, Xambó et al., 2011). Using Grounded Theory methodology, the investigator uses coding techniques to categorise their data; usually video recordings or interviews transcripts. The researcher then attempts to construct (Charmaz, 2006) or discover (Glaser and Strauss, 1967) a model or theory to describe what is taking place. Unlike other research methods, the aim of Grounded Theory is not to support or refute an experimental hypothesis, but rather to arrive at a theory of the interaction derived from the data itself (i.e., is 'grounded' in the data) (Matavire and Brown, 2008). In evaluating a multi-touch music application Xambó et al. (2011) and Laney et al. (2010) argued a combination of Grounded Theory and ethnographic content analysis (Altheide, 1987) can provide a more balanced approach than relying purely on Grounded Theory methods.

A number of researchers have used video observation techniques to study co-located musical interaction. Drawing on theories of non-verbal multi-party interaction, Healey et al. (2005) described the way co-located musicians use space and gesture to co-ordinate their activities. Nabavian (2009) used the distributed cognition framework previously applied to the observation of work teams (Furniss and Blandford, 2006, Hutchins, 1996) to study skilled musicians over repeated rehearsal sessions, engaged in joint musical composition in music studios. Swift et al. (2011) discussed the evaluation of group musical interaction with reference to Sawyer (2003), and used the concept of 'groove' to explain the sensations musicians typically experience and report when playing together. However their position paper is inconclusive, and does not attempt to offer any firm recommendations for the evaluation of collaborative musical environments. Gratier (2008) used a combination of video observation and other measures such as analysis of musical waveforms to compare the recorded musical sound signals with the posture and orientation of musicians engaged in musical improvisation. However, the detailed video observation techniques employed in these examples may not be effective tools for studying all forms of Collaborative Digital Musical Interaction. For instance where musicians are located remotely, issues such as gaze and orientation may become less useful features to study. Also, where people interact via generic computing devices such as laptop computers, physical gestures and posture may not provide as rich a source of information about the activities of the musicians under observation. For this reason, the studies presented in this thesis use a combination of video observation and automated logging to capturee a more detailed picture of the interactions (see Chapter 4).

## 2.6 Summary

This chapter has presented a thorough survey of existing collaborative digital musical interfaces, and has drawn a number of important concepts from within CSCW research. However, despite a shared focus on designing for group interaction, few researchers have previously made connections between the domains of CSCW and collaborative music. Within this chapter existing points of confluence between the two areas have been highlighted, although there is still a great deal of scope for computer music researchers to re-appropriate knowledge and understanding about the design and evaluation of collaborative work systems (James and Stanton, 2011). In particular, this thesis argues that an understanding of issues such as awareness, congruence, and privacy are especially relevant for the design of new collaborative music environments for co-located interaction. The following chapter explores these concepts in more detail, considering in particular the nature of musical interaction and the constraint imposed on it by technological mediation.

# Chapter 3

# Collaborative Digital Musical Interaction

---

*This chapter uses reference to previous theories of musical interaction and other forms of collaboration to define the term Collaborative Digital Musical Interaction (CDMI). CDMI is considered in relation to the properties of sound as a medium, and in turn this leads to the formulation of a number of design challenges for systems to support CDMI. CDMI is also discussed in Fencott and Bryan-Kinns (2013).*

It has been stated that digital technology has significantly altered the nature of musical interaction by introducing new ways for music to be created, heard and shared (Bown et al., 2009, Gurevich and Fyans, 2011). It has also been suggested that people conceive of acoustic instruments in entirely different ways to digital instruments (Gurevich and Fyans, 2011, Magnusson and Mendieta, 2007), and Bown et al. (2009) argued that the language used to discuss musical interaction is frequently tied to inadequate metaphors of acoustic instrumentation. In response to this, the term Digital Musical Interactions (DMI) has been used to describe digital technologies for musical performance, the relation these technologies have with conventional acoustic instruments, and the implications for spectators (Gurevich and Fyans, 2011, Marquez-Borbon et al., 2011, Michael Gurevich and Marquez-Borbon, 2010). However, the term DMI as defined by (Gurevich and Fyans, 2011) appears to be focused primarily on the interaction between a user/musician and interface/instrument, and between the user/musician and the audience, DMI does not consider in any detail the concept of collaborative musical interactions between groups of people making music together, or the ways in which a digital technology might mediate these interactions.

Just as digital instruments can change the way musicians make music, technology can alter the way groups of musicians work together. This thesis proposes the term Collaborative Digital Musical Interaction (CDMI) to describe the phenomenon of technologically supported musical interaction between groups of people. This chapter draws together theories of collaboration and musical interaction to define the concept of CDMI and situate it within the context of previous research. This provides a means by which group music making can be compared to the primarily visual and spatial forms of collaboration that have typically been studied within CSCW.

While musical interaction is not subject to the pressures and dangers of control rooms, operating theatres or other centres of high concentration studied in CSCW literature (see Section 2.1), for the musicians involved the interaction may exhibit similar attributes to those identified through workplaces studies. For instance real-time musical interaction is typically time-critical, and musicians are especially sensitive to timing accuracy on many different musical time scales. Group musical interaction can be highly interdependent, with musicians using and broadcasting cues for changes and transitions, adapting their own contributions in response to what they hear from others in the group (Sawyer, 2003), and using non-verbal communication to help each other recover from mistakes (Gratier, 2008).

This section considers these phenomenon in more detail, with a view to highlighting the implications of designing technology to support CDMI. Emphasis within this chapter is placed on the properties of group music-making as a collaborative activity, the open-ended nature of music-making, the modality of sound and the implications this has for collaboration. This chapter also uses the theory of Communication Constraints (Clark and Brennan, 1991) to further examine how group music-making might be influenced, supported or otherwise altered by technological mediation.

As stated in Chapter 1, this thesis is not concerned with producing a definition of music, and does not comment on the meanings of music, performance, or other such concepts. Such definitions are often overlapping and ambiguous (Cross, 2005), and are not easily translated into practical implications for the design of new musical interfaces or interactions. Neither does this thesis attempt to separate or delineate between different varieties of musical interaction (Bowers, 2002), comment on musical style, or use theories of music (such as Western Tonality) to understand or analyse interaction (e.g., (Gratier, 2008)). Finally, and for the purposes of clarity, this chapter makes no distinction between sound, noise or what might be regarded as 'musical sounds'. Such distinctions are based on a social and cultural interpretations (David J. Hargreaves and Meill, 2005) and bear little significance in a discussion of interface design for collaborative features such as awareness mechanisms. Despite sidestepping these problematic and fuzzy distinctions, Section 3.2 of this chapter does consider the properties of sound as the medium and focus of collaboration, giving particular attention to the ways in which collaboration around a primarily sonic media distinguishes CDMI from other more visual forms of collaboration. The following section considers previous attempts to characterise musical interaction as a group activity.

## 3.1 Musical Interaction

Music has been described in terms of an information theoretic model of source and receiver (Clarke, 2006, Cross, 2005). In this model two or more people share a common set of musical knowledge (described by Akoumianakis et al. (2008) as common ground), which is used for the encoding, transmission and decoding of musical 'communications'. However this view neglects the role of interaction in co-present music making by ignoring the multitude of non-musical communication which takes place during musical interaction (described in Section 3.1.2). For this reason CDMI is rooted in the notion that musical interaction is a collaborative activity which is characterised by rich multi-modal, verbal and non-verbal interchanges, as well as communication through musical contributions which serve both aesthetic and communicative roles (Bryan-Kinns and Hamilton, 2009). Consequentially, computer systems to support musical interaction should consider not only the transmission of musical data, but also the conveyance of extra-musical information related to ownership, activity and awareness.

### 3.1.1 Music and Conversation

Real-time group musical interaction has often been described as similar to a face-to-face conversation (Bryan-Kinns and Hamilton, 2009, Gratier, 2008, Healey et al., 2005, Small, 1998). Frequently cited commonalities include the time-based nature of the interaction, the exchange and development of ideas, the notion of turn taking during interaction, the use of non-verbal communication such as posture, gesture, and the presence of embodied humour. Sawyer (2003) observed that improvising musicians frequently use conversation as a metaphor for describing the process of group improvisation. Healey et al. (2005) identified a turn-taking process used by musicians in improvised music to introduce new musical themes and allow different individuals to solo on their instruments. Similarly, Sawyer (2003) stressed that improvisational music making shares many properties with everyday conversation, including emergence, contingency on the contributions of others, and a reliance on intersubjectivity. However, despite similarities between musical interaction and conversation, there are also some important distinctions. For instance it has been argued that in musical improvisation there is no turn-taking as all musicians perform simultaneously (Sawyer, 2003), and musical contributions serve simultaneously as both communicative acts between musicians, and as an aesthetic product (Bryan-Kinns and Hamilton, 2009, Gratier, 2008). Section 3.1.1 discusses these distinctions with reference to constraints on grounding in communication (Clarke, 2006).

In studying these phenomena, Sawyer (2003) used the concept of interactional synchrony to describe a situation in which performers are attuned to each-others musical gestures and are able to cohesively modify their own contributions during the on-going interaction. Interactional Synchrony was also studied in video of musicians playing acoustic instruments, for example (Gratier, 2008) compared physical posture, orientation, melodic contours, musical timing and gaze, using mother-child interaction to characterise interactional synchrony between improvising jazz musicians. However, due to the technological nature of CDMI, interactional synchrony may not be as useful an approach as it has proved to be in the observation of acoustic instrumental playing, due to the way in which computer based instruments and interfaces can remove the embodied and physical aspects of instrumental playing.

### 3.1.2 Non-Verbal Communication in Musical Interaction

It has frequently been observed that musicians co-ordinate using a mixture of verbal and non-verbal communication. When playing acoustic instruments, musicians provide a rich source of consequential communication (see Section 2.2) through the gestures and movements associated with using their instruments. For instance setting up, adjusting, playing and tuning all provide non-verbal indicators of activity to co-present musicians. Talking, visual cues, gaze, gesture and bodily orientation around shared interaction spaces (Healey et al., 2005) all contribute in various ways to an ensemble's ability to co-ordinate whilst playing, and to their ability to rehearse, compose and perform. These co-ordination mechanisms are used to, for instance, introduce new musical contributions or signifying engagement with the group. Furthermore, the sounds they produce with their instruments and the way they orient around each other within a shared physical interaction space provide information about their engagement within the collaboration.

A caveat to this is that visual cues such as gaze, orientation and posture are only relevant to sighted musicians. Visually impairment can therefore cause problems for blind musicians playing together, especially in cases where the performance deviates from a score, and where the performers need to use visual cues to recover from mistakes (Díaz et al., 2009). Díaz et al. (2009) proposes a technological solution to this, using motion capture and haptic feedback to provide visually impaired musicians with cues from the conductor's baton. This haptic awareness mechanism goes some way to addressing the issue of score following in an orchestral situation, however is very little additional literature concerning the ways in which visually impaired musicians coordinate whilst playing together in improvised contexts.

Collaborative interfaces for musical expression are also channels of human-human communication (Fels, 2004). However, in a computer-mediated context the non-verbal forms of communication, coordination and awareness gathering may be impeded, as the role and actions of the human body may play a less central role in the means of sound production (Bahn, 2001). For example co-located musicians sitting at laptop computers may have a less well defined physical interaction space to orient around, whilst the conventional posture of the laptop musician tends to focus his/her gaze on the screen, rather than the other musicians (Rebelo and Renaud, 2006). In a co-located context, sources of consequential communication may be reduced to include actions such as removing headphones, altering posture or turning away from the computer screen.

As noted above, physical space is an important collaborative resource for groups of musicians (Healey et al., 2005). However the role of this physical interaction space may be reduced when musicians are seated at computers using a shared software interface as the musicians may be less free to move while still able to reach their computer. Compared to acoustic instruments, the abstracted nature of the shared software interface may provide less opportunity for musicians to use physical gestures as a means of communication. In particular, the movements and gestures associated with sound production may have been abstracted away by generic input devices and on-screen instruments (Magnusson, 2009). While acoustic instruments often require specialised physical gestures which are directly associated to sound production (Hunt et al., 2000), a computer musician may interact simply by moving the mouse, pressing keys or using generalised reconfigurable inputs devices such as arrays of buttons or banks of sliders. The meaning of these actions is therefore highly dependant on others' understanding of the music-making environment and the mappings between controls and sounds. Merritt et al. (2010) observed that even ensembles of skilled electroacoustic musicians accustomed to rehearsing and performing together encounter difficulties determining who is responsible for making which sounds, and that visual indicators such as level meters are often used by the musicians to aid their understanding of the unfolding musical improvisation.

### 3.1.3 Group Creativity

Although this thesis is not directly concerned with the subject or study of creativity, it is important to acknowledge this as a central feature of musical interaction, as group musical collaboration has the potential to be spontaneous, unplanned and highly creative (Small, 1998). Truman (2011a,b) described musical creativity using a three stage cyclical model of preparation, generation and evaluation. Sawyer (2003) described the emergent qualities of group creativity, referring to the idea that the products of group creativity can in some ways become more than the sum of their parts, due to the combined actions and input of a group of individuals. Similarly, group creativity has been described in terms of Mutual Engagement, the point at which people 'creatively spark together' (Bryan-Kinns and Hamilton, 2009). Bowers (2002) stressed the role that coincidence, accident and chance play in shaping the outcome of musical interaction. Sawyer (2003) stated that 'In group creativity, interaction between performers is immediate, durationally constrained to the moment of creation, and is mediated by musical or verbal signs. The process of group creativity is coincident with the moment of reception and interpretation by other participants' , referring to the idea that the musicians are simultaneously contributing musical material and responding to the contributions of others. For many people, group creativity is a source of Flow (Bryan-Kinns and Hamilton, 2009, Csikszentmihalyi, 1990, Sawyer, 2003), a psychological state characterised by single-minded focus and attention on the activity at hand, and which goes in tandem with positive emotions such as joy. Extending the concept of Flow, (Sawyer, 2003) coined the term Group Flow to account for situations where multiple people are simultaneously engaged in flow whilst working in a joint creative activity.

An exciting potential for collaborative musical software is the possibility to leverage the spontaneous and emergent properties of group creativity in ways which are not possible either with acoustic instruments or existing single user music technologies. For instance the studies presented in the rest of this thesis explore the possibilities of interfaces which allow musicians to spontaneously edit the musical contributions made by other members of the ensemble, in ways which would not be possible if the musicians were playing conventional acoustic instruments.

### 3.1.4 Open-Ended, Problem Seeking

Musical collaboration is often loosely structured (Nabavian and Bryan-Kinns, 2006), and may not have a specific goal or clearly delineated outcome. Similarly, the activity of making music has been described as being process-oriented, meaning that the product *is* the process (Makelberge, 2010, Sawyer, 2003). In other words, people may engage in music-making as an activity for its own sake, and may be more concerned with exploring the medium, discovering new ideas and finding a creative 'problem' to resolve than with producing a finished work. During the process of creation, the musicians may not know what the final musical result will sound like. For this reason music has been described as a problem-seeking activity (Holland, 1999) where there are no clear goals, no criteria for testing correct solutions and no set of comprehensive or well-defined methods for reaching an end point or solution. Furthermore, judgments about the quality of the interaction may be influenced by the focus on process. For instance musicians engaged in musical collaboration may have enjoyed or otherwise benefitted from the process, whilst at the same time being dissatisfied with the final result. The subjective nature of music can also promote high degrees of disagreement between collaborators, as noted by Young and Colman (1979) in studies of string quartets.

With reference to the evaluation methodologies discussed in Chapter 2, the open-ended nature of music makes certain approaches to evaluation problematic. In particular, task-based evaluations which focus on the successful execution of a task within a given time window do not provide a useful basis for the comparison of musical interfaces. Such measures are concerned with the generation of products, whereas musical interaction is often much more process orientated. Within the context of a CSCW evaluation, participants may have a well-defined task with a clear point of conclusion (e.g., designing a newspaper layout), however with music, the 'task' is ill-defined and the point at which it might be regarded as complete is entirely subjective, and may never be reached. Furthermore, the emergent and unpredictable aspects of group musical interaction, (including inspiration, group flow and group creativity) could mean that the same group of musicians might work different over the course of different sessions together.

### 3.1.5 Mixed Focus

The concept of mixed focus collaboration was introduced in Section 2.2.3 to describe the movement between individual and group level collaboration (Gutwin and Greenberg, 1998). Group musical interaction can be regarded as an instance of mixed-focus collaboration as it features both individual activities and group level interaction. The mixed focus nature of musical interaction has been identified previously in a number of contexts. Studying a group of co-located improvising musicians, Healey et al. (2005) describes the way musicians create and maintain a physical 'interaction space' within the room, which they then orient around to mark their engagement with the rest of the group. The mixed focus nature of musical interaction is also an aspect of the creative process. Talking more generally about musical creativity, (Sawyer, 2003) elaborated on the cyclical, process-oriented aspect of creativity, with reference to the visual artist Picasso, stressing that artists frequently modify their work during the creation process, that inspiration does not always proceed execution, and not all acts of creativity are shared with the world (Sawyer, 2003).

Concerning real-time interaction, the benefits of incorporating privacy features into collaborative music systems have been identified in previous research. The JamSpace environment described by Gurevich (2006) allows users to regulate other users' access to their compositions via the internet. Laney et al. (2010) identified privacy as an important aspect of real-time musical interaction, although their argument is grounded in literature, rather than primary evidence. Fencott and Bryan-Kinns (2010) proposed that privacy features are used extensively when made available, as a mechanism for formulating and testing contributions before they are shared. Similarly, Truman (2011a,b) identified the importance of social and personal level 'preparation', 'generation' and 'evaluation' in the creative process. The concepts of personal and social generation were then used as the basis for a user study investigating the effects of incorporating interface support for these activities within a musical composition environment. However, whilst the creative process was described as cyclical (Truman, 2011a), the interface mechanisms implemented to support the preparation phase of the process were only offered at the start of the interaction, rather than as a feature which could be returned to throughout.

## 3.2 Properties of Sound

The concept of musicking is based on the notion that music does not exist as an object (Small, 1998). Putting to one side the (doubtlessly important) social and cultural aspects of musicking, it is essential to consider the physical and perceptual aspects of sound and music. Although there are many traditions for which notations and graphical representations are a central aspect of musical understanding (Small, 1998, Thiebaut, 2010), music is fundamentally an aural experience. The aural nature of music represents a key distinction between CDMI and research in CSCW, which typically investigates activities that are primarily visual and spatial, such as producing textual documents, drawings and diagrams. Compared to the large amount of research concerned with support for multi-user interaction with visual artefacts such as textual documents, there are relatively few examples of research focused on the design of systems to support the collaborative production of auditory media. This section considers the properties of sound as a medium, compares this to other domains for collaboration, and identifies the challenges sound poses for designers of collaborative musical systems. Although this section has been written with CDMI in mind, it has implications beyond the domain of music, and may be especially relevant for designers of systems to support other forms of auditory collaboration.

### 3.2.1 Temporal Qualities of Sound

Sound is intrinsically time-based; it is important to remember that all sounds, no matter how short, are perceived as a function of time (Roads, 2004). Sound is also ephemeral, and leaves no persistent trace. This means that unless a recording is made, there is no permanent record of a sonic event. This is quite different to visual media such as drawing and written documents, where symbols and marks remains in place after they have been created. Systems such as written notation have been developed to account for this. Written notation instructs a musician in the re-creation of a particular sequence of sounds, whilst time-line based sequencers allow sounds to be organised in time using graphical representations. In real-time musical interaction, however, there is no way to revise what has been previously contributed.

Human memory does however provide a means by which musicians can recall ideas between rehearsal and writing sessions Nabavian (2009), and memory also provides a mechanism for musicians to remember, recall and modify the musical contributions made by other members of an ensemble during improvisation Sawyer (2003). However, these instances of recall are unlike the marks which can be made on a piece of paper or the words typed into an electronic document, as they are reproductions of past event, rather than persistent traces or records.

The time-based nature of musical interaction means anticipation and predictions of future events are important. This has been emphasised especially in the case of improvised styles such as jazz (Sawyer, 2003). In cultures of musical improvisation such as jazz, shared knowledge of musical styles can facilitate the predication of musical developments such as chord changes (Sawyer, 2003, Small, 1998). These predications may also be informed by musicians' familiarity with the playing styles of their collaborators. However, the prediction of future events is excluded in the Workspace Awareness framework (Gutwin and Greenberg, 2002), as it is argued that designers are unlikely to be able to support these aspects of collaboration. Similarly, (Renaud et al., 2007) stated that the ability for musicians to anticipate each-other's actions is difficult to support in a networked environment.

### 3.2.2 Spatial Qualities of Sound

Working with sound is unlike working with visual materials, as sound is pervasive and ephemeral and has entirely different spatial qualities to visual media. Unlike visual media such as drawing, sound does not linger or leave a persistent spatial impression. Sound is not directly malleable, and computer based interaction with sound is usually via a graphical or tactile interface. Where graphical interfaces are employed within software as a means of controlling the generation and manipulations of sounds, the interaction can be regarded as multi-modal. This is because the means of interaction operates in a separate modality to the output. By means of comparison, in a collaborative diagram editing context users are often able to directly manipulate a visual representation of the artefact they are producing.

The theory of communicative grounding (Clark and Brennan, 1991) notes that 'indicative gestures' such as looking, pointing and touching are important means by which interlocutors arrive at an understanding that they are both referring to the same object. The spatial properties of visual media have been observed to contribute to the management of tasks and conversation (Healey and Peters, 2007), and are an important aspect of establishing joint attention in collaborative domains such as document editing (Cohen, 2008). However, the ephemeral nature of sound means groups of people are unable to point at, gesture or orient around a sound in the same way as they can when working with visual artefacts.

In CDMI this modality split between visual control and sonic output may complicate the design of collaborative interface features such as awareness mechanisms which rely on conveying information about the identity, location, source or ownership of a sound or musical contribution. For example telepointers have been used in collaborative environments to show where another user's mouse currently resides as an indicator of their current (and possibly previous) actions and attention within a graphical document (Gutwin and Penner, 2002). However such visual information may not provide users with as much information about what another user is currently attending to in a collaborative music environment due to the disconnection between graphical representation and fleeting sonic output.

Coughlan and Johnson (2007) stated that musicians use many forms of representation to convey ideas and refer to aspects of the music, including playing their instruments, vocalising, gesture and verbal communication. The studies in this thesis also show that people use verbal imitation, descriptions, and reference to graphical representations to draw attention to aspects of the shared music. This illustrates the idea that a musical gesture is both an act of communication and an aesthetic product in its own right (Bryan-Kinns and Hamilton, 2009, Gratier, 2008). Coughlan and Johnson (2007) argued that an understanding of how musicians represent and convey ideas is crucial to the design of new musical interfaces and software environments, while Nabavian (2009) noted that musicians often successfully collaborate while holding entirely different cognitive representations of the music they are co-creating.

### 3.2.3 Disruption

One property of the phenomenon of auditory masking is for loud sounds to conceal quiet sounds (Bregman, 1990). This means the generation of sound may cause verbal communication to be problematic during musical interaction as excessive sound volume may interfere with normal conversation. Equipment such as headphones may also cause verbal communication to be difficult. In addition to sound level in the environment, individuals may have difficulty speaking whilst concentrating on the activities associated with playing their instruments or controlling interfaces, and the positioning of individuals may also restrict conversation. Where technologies such as computers are involved musicians may also be tied to a specific location, and furniture, stands and other items may make it difficult for musicians to move around.

## 3.3 Communication Constraints for Musical Interaction

This section uses the Communication Constraints proposed by Clark and Brennan (1991) to understand how technological mediation can influence musical interaction. These constraints were previously used to consider forms of mediated human-human communication. In CDMI, technological mediation imposes constraints on the ways in which musicians can work together. Although this thesis does not attempt to make strict distinctions between different types of musical activities, the influence of these constraints are important to acknowledge. Clark and Brennan (1991) defined eight constraints which can be used to compare communication modalities. These constraints, discussed in the following sub-sections, refer to the properties of the situation or context itself (i.e. co-located interaction), rather than the constraints imposed by an instrument or interface, or creative constraints a musician may adhere to as part of their artistic practice (Candy, 2007, Michael Gurevich and Marquez-Borbon, 2010). This discussion serves to highlight the role technology places in shaping, supporting and prohibiting certain forms of coordination and communication between groups of people whilst engaging in music-making. The discussion also stresses a number of distinctions between CDMI and other forms of group collaboration, and between CDMI and face to face verbal conversation.

### 3.3.1 Copresence

Clark and Brennan (1991) uses the term copresence to refers to situations where people occupy the same physical space and are therefore able to interact face-to-face. Within this thesis copresence has previously been discussed in relation to the ethnographic studies of groupwork (Furniss and Blandford, 2006, Heath et al., 1995, 2002b, Hutchins, 1996), and observational studies of conventional musical interaction (Gratier, 2008, Healey et al., 2005, Nabavian, 2009, Nabavian and Bryan-Kinns, 2006, Sawyer, 2003). Copresence is a natural context for verbal communication, and copresence also affords non-verbal communication in the form of gestures, gaze, pointing, and orientation. Early research within CSCW attempted to recreate the experience of co-located interaction between geographically separated individuals using live video links (Ishii et al., 1993, Olson et al., 1995). However, it was soon realised that even with multiple channels of video, such systems provide a limited approximation of some aspects of co-located interaction (Hollan and Stornetta, 1992). For instance, gaze, eye contact and the ability to establishing reference to artefacts through pointing, are all potentially compromised by video mediated communication. Groupware systems also attempt to recreate certain aspects of co-located interaction. The Workspace Awareness framework (Gutwin and Greenberg, 2002) used throughout this thesis is based on observations of co-located interaction, and contributes to an understanding of how awareness features can be mediated by technology to support remote collaboration.

Before the advent of technologies such as the internet, musical interactions between people would have relied more heavily on copresence. However, digital technology allows for a variety of remote interactions, such as file sharing, audio streaming, digital video-mediation and online composition systems. In attempting to classify systems to support musical interaction between people, Barbosa (2003) referred to copresence using the terms *local* and *remote*. This thesis focuses on co-located (local) musical interaction between people, mediated via a shared software interface. From a design perspective, this approach alleviates the need for an interface to support communication and interaction which is naturally afforded by co-location, although compared to acoustic instrument playing, the use of screen-based interfaces for musical interaction may affect the degree to which users communication their activities as a consequence of their physical movements.

### 3.3.2 Visibility

The constraint of visibility refers to the ability for people to see each other whist interacting. Face-to-face interaction would normally presuppose visibility, and remote video mediated communication also supports visibility, despite the participants not being co-located. Visibility might normally be the case for co-located musical interaction between small groups of musicians (Fencott and Bryan-Kinns, 2010, Gratier, 2008, Healey et al., 2005, Nabavian, 2009, Nabavian and Bryan-Kinns, 2006), although it has been noted that visibility is not necessarily essential for grounding in musical improvisation (Gratier, 2008). Visibility may not be assured in larger ensembles of co-located musicians, such as orchestras, where musicians are spread out over a large area.

### 3.3.3 Audibility

The audibility communication constraint refers to the ability for people to hear each other during communication. This is assumed in verbal face-to-face communication, and is essential for telephone conversation, but is not necessarily a feature of other forms of mediated communication, for instance e-mail, video-only communication, or text-messaging. The sonic nature of music presupposes the need for audibility for collaborative musical interaction, and the ability to hear others is one of the main ways in which musicians co-ordinate their actions. This thesis explores a number of aspects of the Audibility constraint, particularly in relation to the ways in which interfaces for collaboration can be designed to disrupt the congruence of audio presentation between collaborators.

### 3.3.4 Cotemporality

The cotemporality constraint refers to receiving a contribution at the same moment that it is being produced by someone else. For instance face-to-face or telephone conversations can be regarded as cotemporal. However in contrast, communication formats such as e-mail can disrupt cotemporality by allowing messages to arrive out of sequence. Most real-time musical interaction relies on contemporality, not only for successful interaction between musicians, but also for the creation of music which is aesthetically pleasing (i.e., in time and in tune). Reflecting this, a fundamental concern for distributed technologies is network latency, which can heavily disrupt the cotemporality of musical interaction (Barbosa, 2003). Within the research presented in this thesis, contemporality is assured by several means. Firstly, colocating the participants allows face-to-face speech, secondly, the software environment was deployed on a high-speed local network (to minimise network latency), and the software (see Section 4.2) used a centralised mechanism to ensure that all musical events were delivered cotemporarily to all participants.

### 3.3.5 Simultaneity

Simultaneity refers to people being able to send and receive at once. In face-to-face and telephone conversation, interlocutors are able to speak at the same time. Simultaneity would usually be assumed where participants each have individual instruments or mechanisms for producing sound, and indeed it has been stressed that in activities such as group improvisation, musicians generally contribute simultaneously, rather than taking turns (Sawyer, 2003), although some group members may be taking a more prominent role such as a lead player. The issue of simutaneity becomes blurred when considering multi-user instruments which require two or more people to *jointly* control or co-author the sonic output of an instrument. Examples include the Tooka (Fels and Vogt, 2002), and experimental compositions such as Stockhausen's *Mikrophonie* composition (Maconie, 1989). In examples such as these, whilst the musicians are interacting in simultaneity their actions are causing the interface/instrument to produce a single musical outcome, rather than two distinct and simultaneous musical voices.

### 3.3.6 Sequentiality

Forms of communication such as e-mail allow contributions to go out of sequence, possibly leading to confusion or misunderstandings. However, as with face-to-face conversation, contributions in real-time musical interaction cannot fall out of sequence. In real-time musical interaction, sequentiality is also assured. Remote interaction and asynchronous collaborative activities may alter the sequence of musical contributions, however these activities are not addressed within this thesis.

### 3.3.7 Reviewability

Communication media such as letters and e-mail allow their contents to be reviewed after transmission, while speech is evanescent (Clark and Brennan, 1991) and fades away as soon as it has been uttered. Similarly, real-time musical interaction with conventional acoustic instruments does not include reviewability, as the sounds of acoustic instruments are ephemeral, and pass by in the same way as vocal utterances. Recording technologies allow speech and music to be reviewed after the event, although this arguably challenges their status as real-time activities.

Other forms of musical interactions may allow for reviewability, for example non-realtime composition systems may enable musicians to revise their contributions after they have been submitted. Real-time musical interaction which is based on the editing of repetitive looping phrases, such as in the Daisyphone (Bryan-Kinns, 2004), may also support reviewability in the sense that a musical phrase can be altered after it has been entered into the system. Similarly, (Gratier, 2008) stated that musical repetition is another essential part of musical grounding for improvising musicians.

### 3.3.8 Revisability

Forms of communication such as e-mail and letters allow messages to be revised before being transmitted. In some forms of musical interaction musicians may also be able to revise their contributions before they are shared with others. For example people composing music over a long period of time may work on ideas in private before bringing them together and sharing or combining them with the contributions of other people. Similarly, in real-time interaction, people may temporarily break away from the group or play ideas more quietly in order to revise and formulate a contribution in private (Healey et al., 2005, Nabavian, 2009). This idea has previously been discussed with reference to the notion of mixed focus collaboration. Previous work on cotemporal and co-located musical interaction demonstrates that the inclusion of revisability, when presented in the form of private workspaces was favoured by collaborators during a group composition activity (Fencott and Bryan-Kinns, 2010).

### 3.3.9 Summary

This discussion of the communicative constraints imposed by technological mediation in musical interaction has performed several functions. Firstly, it has highlighted a number of ways in which music differs from other forms of communication. This is an essential distinction to make, as it has implications not only for understanding the ways in which people perceive sound, but also for understanding how interfaces must be designed so as to support musical collaboration. Crucially, group musical interaction in real-time relies upon audibility and sequentiality, and although some coordination mechanisms may rely on visibility, this may not always be available. Secondly, this discussion identifies a number of dimensions along which technological support for musical interaction can influence the kinds of activities a system is likely to support. A breakdown in simultaneity will be a serious issue for real-time interaction (a known and well attended problem for online collaboration), whilst constraints such as revisability are open for support in software interfaces. Finally, these two considerations serve to delimit the domain of study addressed in the remainder of this thesis, which, as outlined previously is real-time co-located interaction between musicians, supported through a shared software interface.

## 3.4 Designing for CDMI

Previous research has considered theoretical perspectives on the design of musical interfaces. Gerzso (1992) stated that in all cases the systems embody a particular approach to making music. Such approaches which arrive with assumptions about the nature of music, musical representation, and interaction. For instance a time-line based interface such as found in applications such as Cubase[1] considers music in terms of musical events which are arranged in a pre-defined temporal order. Magnusson (2006) goes further by using semiotic theory to describe the design of music software in terms of as a process of combining signs and interface metaphors with compositional ideology, to create a system which allows users to express their ideas and communicate

[1] http://www.steinberg.net/en/products/cubase/start.html Accessed 30 September 2012.

their ideas, whist at the same time opening up the possibility for new ideas and ways of thinking to emerge. Magnusson (2006) describes three 'interaction paradigms' within musical computer interfaces, the 'computer-as-tool, computer-as-partner, and computer-as-medium'. However, less emphasis is placed within this work on the role of computers in a collaborative context where multiple people are simultaneously creating music together.

As previously discussed in this thesis, there are no clear boundaries between different kinds of musical activities, musicians approach musical creativity in different ways, and the activity of making music has no clearly defined end point or criteria for evaluation. Furthermore, it is essential to realise that the tools, instruments and interfaces employed in the process of making music influence the musical outcomes that are realised, and also affect the ways in which multiple people interact during the process. Musicians also have a tendency to re-appropriate technologies for new purposes (Cascone, 2000), making it difficult to predict how features within an interface will be used, and making it difficult to reason about how people will interact with each other. These properties of musical interaction make it difficult to formulate a design specification for a musical interface, and designers are consequentially placed in a position where they too must explore and experiment to arrive at interesting solutions. Put another way, Cook (2001) stated that musical interface construction perhaps necessarily 'proceeds as more art than science'.

Much like the activity of music-making itself, the challenges of designing technology to support musical interaction feature many of the traits of a 'wicked problem' (Fitzpatrick et al., 1998, Ritchey, 2005). The following points, adapted from Ritchey (2005), demonstrate the aspects of designing for CDMI which can be regarded as 'wicked':

- the open-ended and ambiguous nature of musical interaction makes it difficult to formulate an exact specification for the software
- There is no stopping rule in the development process, features can be added or extended indefinitely
- Judgements about the software (i.e., evaluations) indicate 'better or worse', rather than true-or-false
- There is no immediate and no ultimate test of a solution
- There is an infinite set of potential solutions
- The challenges of designing for CDMI are unique

## 3.5 Summary

This chapter has outlined Collaborative Digital Musical Interaction (CDMI). Whist acknowledging and building upon previous descriptions of Digital Musical Interactions, CDMI focuses on the interaction surrounding group music-making which is mediated by technology. Particular facets of CDMI which have been explored in this chapter include the ways in which sound as a medium distinguishes musical collaboration from other forms of collaboration, and the ways in which technology makes CDMI different from other forms of musical interaction. This chapter has also broken down CDMI in terms of the communication constraints defined by Clark and Brennan (1991), so as to understand the role technology can play in altering the ways in which musical interaction can proceed. The following chapter briefly discusses the research methods adopted within this thesis, and introduces the collaborative music software developed in order to carry out the empirical studies.

# Chapter 4

# Methodology

---

*The first section of this chapter discusses the motivations for employing lab-based user studies. The second section (Section 4.2) justifies the reasons for developing a piece of bespoke software for use within these studies, and presents some general details about the implementation of this software. This chapter serves as a general introduction to these issues; more specific details of the methods and software environments used in the studies are given in their associated chapters (Chapters 5, 6, 7).*

## 4.1 Methodological Approach

The literature review identified several approaches to studying group work and collaborative interaction. Broadly speaking, the two approaches most commonly seen in CSCW research are naturalistic observational studies of people engaging in work activities, and lab based experiments. Observational studies typically use ethnographic techniques (Heath et al., 1995, 2002b), or frameworks such as Distributed Cognition (Furniss and Blandford, 2006, Hutchins, 1996). This approach has been used in the study of groups of musicians writing and rehearsing in rehearsal rooms (Nabavian, 2009), and also ensembles of electro-acoustic musicians collaborating in real-time (Merritt et al., 2010). However, this methodological approach requires access to a group of subjects who can be observed whilst engaging in their collaborative activities over a sustained period of time.

Although it has been suggested that the evaluation of collaborative musical interfaces should occur in a setting which is as close as possible to a 'real context' (Xambó et al., 2011a), a number of issues discouraged the use of purely observational methods within this thesis. Firstly, the forms of shared interface designs this thesis set out to investigate are not commonly used by groups of collaborating musicians, making it difficult to find suitable groups of experienced users to observe. This also makes it difficult to identify an existing 'context' within which to perform an evaluation. Secondly, should such a group of musicians be located, gaining sustained access to them for the purpose of conducting the observations may not be easy. Thirdly, a purely observational approach, by definition, does not allow for the use of experimental manipulations or interventions, meaning it would be difficult to perform explicit comparisons between different interface designs. The scarcity of established groups of users to observe also excludes the user of more design orientated methodologies such as participatory design.

For these reasons, the studies presented in this thesis use a controlled experimental approach, whereby groups of musicians are observed in a laboratory context, and presented with a number of different interface variations. The decision to use short controlled studies was also positively motivated by several factors which mitigate against the unnatural context of experimental studies. In particular, such an approach bypasses the need to locate existing groups of musicians, and makes it possible to introduce experimental interventions such as changing features within the interface. Lab based studies also provide a greater degree of control over the study context and the data collection process. For instance they provide the opportunity for detailed logging of user interaction which would not be possible through an observation of people using software which was either commercially available or constructed by the musicians themselves. Secondly, the controlled nature of the studies enables statistical comparisons to be made between groups of participants exposed to different experimental conditions.

The use of short controlled studies also makes sense given the infancy of research within the field of collaborative music. The limited knowledge currently available about how people use shared interfaces for music-making means that there is still a lot to be learnt from small controlled studies, as demonstrated by the contributions of this thesis. In particular, rather than focusing on a limited number of people, the use of controlled studies allowed for a large number of participant groups to be observed, and enabled a range of different interface designs and features to be compared. As systems for CDMI become more widely adopted, the use of purely observational naturalistic studies will surely become more beneficial. This is acknowledged as future work.

Finally, the lack of commercially available collaborative music software for real-time interaction required the development of a new piece of software as a vehicle for conducting the research, although the timescale of this PhD restricted the amount of time that could be spent developing software for the purpose of running experiments. The construction and testing of software is an involved process, and designing and developing a compact prototype interface which is suitable for conducting short user studies is considerably less complex than creating a large, fully featured application which would be suitable for use in multiple sessions by experienced users over a long period of time. Given the paucity of information about the design of collaborative musical software, developing a large-scale application would arguably have been a premature step for this research. Without identifying and understanding some of the primary issues for a CDMI system, the development of an application for use in a large scale observational study would largely have been based on guesswork. The contributions this thesis makes can in this sense be regarded as an incremental step towards the design of a more featured application which could be used in a large scale study. An overview of the software developed for this thesis is described in Section 4.2, whilst the details of the interfaces used in each study are presented in their respective chapters.

### 4.1.1 Study Methodology

During experiment sessions groups of musicians make music using a collaborative music environment. The groups are presented with different software interface designs. Observations are taken to understand how this impacts on the way the groups of individuals use the software, their approach to organising collaborative activities and their reported preferences. This approach is inspired by CSCW studies such as that by Gutwin and Greenberg (1999), where various interface features are presented to assess the usability of groupware software. Quantitatively measurable features within interaction using automatically collected interaction log data from the software and multiple choice questionnaires. Video observation and group discussions were also used to gather qualitative data. This methodology is informed by the approach taken in some forms of CSCW research, although the additional qualitative measures extend the traditional CSCW approach to account for some of the distinct properties of CDMI.

This research focuses on studying aspects of group interaction surrounding CDMI, rather than attempting to address issues such as expressivity, which are highly subjective and difficult to operationalise. Similarly, the analysis performed for this research did not attempt to rate, judge or review the quality or musicality of the music created by any of the study participant groups. Such analysis would involve subjective judgements about the quality of the music, and this was seen as peripheral to understanding *how* the interfaces presented to participants supported or interfered with music-making processes, and how they influenced the participants' experiences whilst using them. Furthermore, the interfaces presented to participants featured various forms of incongruence in the way audio was presented, and in the forms of access the interface provided users with (e.g., relating to privacy and sharing of contributions). This incongruence makes it difficult to delimit what might be regarded as the *final* music the participants created during the experimental sessions, further complicating any hypothetical processes of critique by external judges.

### 4.1.2 Participants and Participant Groups

A decision was made to study triadic groups of participants (groups of three) in all studies. Interaction in triadic groups has been identified as distinct from interaction in dyads (two people). For example as part of conversation, gaze has limited value in a triadic group as it is only possible to look at one person at once (Battersby and Healey, 2008). Considering musical collaboration, issues such as ownership become more uncertain and ambiguous in triadic groups, as it less obvious who contributed which sounds, or who performed which actions. In a dyad, such information would be implicit for both group members. Studying triads also represents a compromise between the logistic challenge of recruiting anticipants and forming groups of more than two people.

E-mail mailing lists were the primary means by which participants were recruited. All three studies sought to recruit people who identified themselves as musicians, people with an interest in music technology or people with prior experience using computer music software. As discussed previously, the design of collaborative systems for novice users has been a concern for a number of researchers (Weinberg, 2003), however this places considerably different demands on the system design. The decision to use music terminology and interface metaphors similar to those found in conventional music software sets this research apart from collaborative music research aimed at supporting audiences, non-musicians, novice users, children or education (Bryan-Kinns, 2004, Freeman, 2008, Truman, 2011a, Weinberg et al., 2002). The value of conducting user studies of musical interfaces using novice users has not been fully assessed in previous research, although there is evidence that musicians and non-musicians perceive music in different ways (Koelsch et al., 1999), and the decision to recruit participants with at least a basic knowledge of music and music technology is also supported by James and Stanton (2011), who demonstrate a language barrier caused by lack of musical domain knowledge in novice users

While many researchers have identified social and musical rapport as an important aspect of group musical interaction (Sawyer, 2003, Small, 1998) several factors motivated us to study groups of people who have not previously worked together. Firstly, established groups will arrive with a history of shared experiences to draw upon, a shared musical repertoire and established working strategies. These working strategies and means of communication may be obtuse and difficult to interpret or study. Secondly, studying musicians who have previously worked together may introduce bias between groups, as not all groups will have an equal level of experience. Thirdly, the group's musical repertoire, group musical knowledge and established working strategies may be stronger than or resilient to the effects brought about by the experimental conditions under investigation. Studying groups of individuals who are not familiar with each other introduces control over differences in the level of group experience, as all participants will have an equal level of familiarity with one another. Although it is important to acknowledge that participants will need to build a social and musical rapport and although it is less common for musicians to play with people they have not previously worked with, this is not an entirely unnatural situation. Finally, recruiting groups of strangers simplifies the process of recruitment and allows us to use a larger sample of participants.

## 4.2 Collaborative Music Software

The collaborative music-making software used for this thesis was designed and implemented specifically for the purpose of performing the interaction studies presented in subsequent chapters. This section justifies the motivations for developing a bespoke application rather than using or adapting a pre-existing environment. The differing functionality requirements and research objectives for each of the three studies presented in this thesis necessitated changes to the functionality and appearance of the software. To avoid repetition, this section provides an overview of the software, with a focus on the high-level design motivations rather than specific interface features. The individual study chapters contain more detailed descriptions of the interfaces presented to participants in each instance (see Sections 5.1, 6.2, 7.4). This section then details the underlying architecture and iterative implementation process adopted throughout the research process.

The collaborative music software allows multiple people working on separate desktop or laptop computers to create music together via a local network connection. The software is based on the relaxed WYSIWIS metaphor of a shared workspace in which multiple virtual instruments can be created and manipulated simultaneously by all users taking part in the session (Greenberg, 1996, Stefik et al., 1987). The software does not map spatial layout to any musical parameters, and unlike digital audio workstations, does not feature a timeline for the arrangement of musical ideas at a compositional level, and does not incorporate extensive mechanisms for scheduling future events on a compositional or 'macroform' time-scale (Roads, 2004). From a research perspective this matters, as an analysis of interface designs which map on-screen spatial positioning to musical parameters may conflate musical arrangement with territorial behaviour. Klügel et al. (2011) suggested that avoiding spatial metaphors and musical mappings is beneficial in tabletop style interaction as it supports dynamic coupling between participants, however there is limited research on the issue of spatial mappings in collaborative distributed desktop environments. While this limits the applicability of some findings related to spatial layout, it is important to stress that many music applications do not rely on the timeline metaphor. Examples of music software which do not map spatial position to time include patching environments such as Plogue Bidule[1], Cycling 74's Max/MSP[2], Pure Data[3]and Audio Mulch[4], modular synthesis environments such as Native Instruments Reaktor[5], step-sequencer based applications such as Propellerheads ReBirth[6], interactive table systems such as the reacTable (Jordà et al., 2007), and live performance applications such as Ableton Live[7].

### 4.2.1 Software as Research Tool

Creating a bespoke application to perform the research presented in this thesis was necessary for a number of reasons. As stated in the background chapter 2, the prevailing interaction paradigm for music software is that of one user at a time interaction, and there are very few existing applications which allow groups of people to engage in complex real-time interaction over a local

[1] http://www.plogue.com/products/bidule/
[2] http://cycling74.com/
[3] http://puredata.info/
[4] http://www.audiomulch.com/
[5] http://www.native-instruments.com/en/products/producer/reaktor-55/
[6] www.rebirthmuseum.com/
[7] http://www.ableton.com/live

network connection. Consequentially, a primary motivation for developing bespoke software was born of necessity, as at the time the research was carried out, no readily available software existed to fully support the forms of interaction and user interface features under investigation. In particular, while some applications support real-time multi-user interaction with virtual software instruments, they do not necessarily feature multiple workspaces to support varying levels of privacy (investigated in Chapter 5), incorporate ways of altering which devices are used to present public and personal audio channels (see Chapter 6), or provide awareness mechanisms to highlight the contributions made by different users within the session (Chapter 7). Furthermore, one goal of this research is to arrive at guidelines for designing new tools to support musicians during the process of co-located group musical interaction. As such, a further requirement was for the research subjects to work with a software environment that bears resemblance to genuine music making applications, rather than a sound toy (Robson, 2001) or simplified interface for non-musician novice users (Bryan-Kinns, 2004, Weinberg, 2003, 2005, Weinberg et al., 2002). Using bespoke software also provides the researcher with control over the capabilities of the software so as to engender a suitable learning curve for participants given the limited time constraints of a short experiment session. This is important as it enables participants to gain a working knowledge of the software within a short amount of time, while at once allowing certain more advance features to be incorporated at the discretion of the researcher.

Developing bespoke software also presents a number of advantages over performing evaluation techniques on existing software. Specifically, such an approach enables the researcher to base an application *around* the research questions under investigation. In this case, using bespoke software provides complete control over every aspect of the software's functionality, appearance and behaviour, making it possible to implement multiple interface designs using the same underlying software framework. This supports the exploration of interface and interaction possibilities that are not present in existing applications, and is crucial for facilitating the experimental manipulation of specific features and performing interventions such as removing interface components or altering the behaviour of certain aspects of the interface. Finally, presenting research participants with a novel interface allows the researcher to control for familiarity with the interface, by ensuring that non of the research participants will have used it prior to the experiment taking place.

Having access to the internal workings of the software presents the opportunity to incorporate automated data capture mechanisms such as time stamped event logging, or caching of internal memory states throughout the course of an interaction. In practice, this allows the software to automatically capture both user interaction and the software responses to this interaction in a format dictated by the researcher, for the purpose of conducting fine grained analysis of how the software was used.

### 4.2.2 Implementation

The software implementation was informed by the Client-Server architecture of existing web-based collaborative systems such as Web Drum (Burk, 2000) and Daisyphone (Bryan-Kinns, 2004) and follows the model-view-controller paradigm. In practice, this required the development of two interrelated applications, a central Server, and a Client application incorporating a user interface. The graphical interface, underlying server-side code and client-side graphical interface was implemented in Java. Java was chosen for its simplified network layer abstractions, rich graphical interface libraries and capacity to support object-oriented design principles.

The Server performs back-end tasks such as establishing connection to all Client applications, maintaining a consistent representation of the shared interface components and generating correctly timed musical messages for the rendering of audio. For this purpose, the Server contains a *Module Database* to represent the state of the music software, including a record of which Music Modules are currently active within the workspace, where they are located on the interface, which client caused them to be created, and the state of their musical parameters. The Module Database is also used to generate timed messages which trigger the generation of music on the Client workstations, based on the currently active Music Modules. The generated audio can therefore be regarded as a sonification of the internal state of the Server.

The Client application is the user-facing graphical interface, containing the shared workspace used to create and edit music modules. To streamline development, a set of Java Swing interface elements were extended to encapsulate the transmission of user interaction modifications via a TCP/IP socket connection. In the Study One implementation the Music Module interfaces were generated on-the-fly based on commands generated from the Server, while in the Study Two and Study Three implementations which featured more complex Music Module interfaces, the Music Module GUIs were specified as separate classes, and instantiated dynamically within the Client and Server via a *Factory* class.

To guarantee the successful transmission of all control messages between Server and Client, the Server and Client applications communicate Server via TCP/IP sockets, with a separate thread running on the Server to handle message transmission for each connected Client. A custom messaging protocol was developed using Enumerated types to encapsulate message transmission between Server and Client applications and ensure the correct formation of network messages.

The network messages were designed to be human-readable, and are logged in text files as one form of interaction data logging. In addition, in the software used for studies Two and Three, each change made to the Client graphical interfaces or Server side Module Database triggered the status of the entire Module Database contents to be logged in a separate text file, alongside a timestamp. This extra information facilitates accurate reconstruction and analysis of the user interact at any given moment in the log files.

The audio synthesis algorithms were written in SuperCollider (McCartney, 2002). A SuperCollider Server instance runs on each client computer to render audio when directed by the Server application. This approach to audio rendering takes advantage of the advanced synthesis capabilities of SuperCollider, specifically the low latency audio server, pre-existing unit generators for synthesis and effects, simplified sample playback, and internal bus architecture for audio routing. Choosing to use SuperCollider also bypassed the latency issues and complexities of sound synthesis within Java. To ensure low latency across multiple Clients running on separate computer workstations, messages to trigger audio events are sent from the Server directly to SuperCollider instances using Open Sound Control (Wright et al., 2003) over UDP. While exact latencies were not measured, informal testing showed there to be no noticeable timing discrepancies between audio events triggered on multiple computers when using a local network. This approach would not provide stable timing over the internet, however as the intention of this work was to focus on co-located interaction this solution was entirely adequate.

### 4.2.3 Development Iterations

The software was used in three experimental studies, presented in Chapters 5, 6, and 7. The software was developed in iterations between each study. Details of the specific software features are presented in the chapters associated with each study. Table 4.1 presents some metrics to illustrate the stages of software development throughout the research.

| Iteration | Lines of Code | Number of Source Files |
|---|---|---|
| Study One | 3360 | 51 |
| Study Two | 7555 | 49 |
| Study Three | 8876 | 53 |

Table 4.1: Iterations of software development through the research

## 4.3 Summary

In summary, the implementation process resulted in the iterative development of a robust research tool for conducting the studies presented in this thesis, where the research questions and experimental designs drove and informed the software design. This approach is therefore in line with other researchers who develop musical interfaces specifically to support research (Marquez-Borbon et al., 2011), and is in opposition to approaches which focus on initial technological development of software or hardware interfaces, before applying posteriori evaluation techniques to these previously developed artefacts of new musical technology.

The studies within this thesis use a controlled experimental approach. Although there are good reasons to use observational and naturalistic approaches to studying musical interaction, these approaches are also problematic. For the study of CDMI, a key disadvantage is that there are few existing users of collaborative music applications, making it difficult to find groups of genuine users to observe in their natural contexts. Furthermore, at this stage, there are lots of benefits to using more controlled approaches, such as the ability to observe the effects systematic changes in a user interface have on the way groups of people interact and collaborate. To conduct the studies presented within this thesis, a software environment was developed. Developing software for the explicit purpose of performing research is a relatively new approach within the study of musical interaction, although it presents a number of advantages to the researcher, including complete control over the functionality, and the option of creating interaction logs to record all salient details of its use. The three chapters to follow present detailed studies using this software.

# Chapter 5

# Study One: Privacy and Awareness

---

*The first study of this thesis explored how the inclusion of a Private workspace and an awareness mechanism affected the way groups of users interacted whilst engaged in CDMI. Results suggest that the Private workspace was used extensively to support mixed focus collaboration, whilst the addition of an awareness mechanism changed the subjective quality of the interaction. Portions of this work were presented in Fencott and Bryan-Kinns (2010) and Fencott and Bryan-Kinns (2013). This study design proposed in this chapter was cited as the basis for future work by Müller-Rakow and Fuchs (2012).*

This chapter presents a study to investigate the effects varying the amount of privacy and awareness information provided by a software interface has on the way groups of people engage in the process of collaborative music making. This study addressed the research question *how do interface features designed to support privacy and awareness alter the way groups of co-located musicians engage in collaborative musical interaction?* For instance allowing users to work in isolation from one another may encourage less focus on the shared aspects of the music, or might disrupt the users ability to contribute cohesive musical parts. Allowing collaborators to work in private may also mean each collaborator is listening to a mixture of shared material and individual contributions. Given a situation where individuals are listening to heterogeneous versions of the music it may become more difficult for collaborators to co-ordinate their actions or maintain awareness of each others' activities. A software interface could potentially provide additional awareness mechanisms to compensate for situations where musicians are working individually, and in turn such mechanisms may also impact on the way the individuals within a group structure their interaction or choose to share musical material. Portions of this work were presented in Fencott and Bryan-Kinns (2010, 2013).

In this study nine groups of three musicians composed music together using three interface designs, each providing different levels of privacy and awareness information. The recruitment process is described in Section 5.7. The first interface offered a public workspace where all musical contributions are visible and audible to other users. The second interface introduced a Private workspace for each user. The Private workspace could be used to create, edit and audition musical contributions in isolation from other group members. The third interface design granted users read-only access to each others' Private spaces, therefore potentially providing users with increased awareness of one another's activities while at the same time reducing the level of individual privacy offered by the interface. The order of interface presentation was randomised between groups. Later sections of this chapter give more detail about the experimental conditions, hypothesis, variables and measures.

These interface designs were informed by previous CSCW research into interface support for privacy during collaboration (Dourish and Bellotti, 1992, Olson et al., 1993), and studies of large public displays (Vogel and Balakrishnan, 2004). Within the field of digital musical interaction, similar arrangements of individual and shared workspaces have been incorporated into collaborative music environments. For instance Laney et al. (2010) describe a multi-touch collaborative music environment which enables users to choose between playing audio privately through headphones or publicly through speakers. whilst (Gurevich, 2006) provides a range of privacy options within the internet based JamSpace environment. However the study presented in this chapter goes further than previous work by systematically investigating the effects of manipulating privacy and awareness mechanisms in a shared music making environment, and by studying the ways in which shared and individual workspaces contribute to the processes of collaboration and coordination during computer supported musical interaction.

## 5.1 Collaborative Music Software

A collaborative music making environment was developed to run this study. The software is a Client-Server based system which enables multiple Client applications to connect simultaneously to a central Server. Each user within the system uses a Client application, which comprises of the graphical interface and audio rendering. The Server maintains consistency between all connected Client applications and manages a database of the music being created. Section 4.2 describes the underlying technical and implementation details for the collaborative music software developed for use in all three of the studies in this thesis. This section describes the specific music-making capabilities, interface features, controls and collaborative mechanisms implemented for the study described in this chapter.

### 5.1.1 Creating Music

The music software allows musicians to create 'electronica' style music which relies primarily on synthesised sounds, making it suitable for generation using computer software. Users of the software make music by deploying Music Modules within an on-screen workspace in the Client application. The modules provided in the software can be used to create percussive parts such as bass drums, snare drums and high-hat rhythms, as well as melodies, bass lines and ambient textures. Modules are represented as windows containing sliders to control their musical parameters. Each module offers control over its volume and stereo pan position (see Figure 5.1), and also includes a number of module specific parameters. The music modules are algorithmic in nature, meaning they generate musical material automatically in response to the settings of various high-level parameters. While the algorithmic music modules do not allow fine-grained control over specific notes or musical patterns, altering their musical parameters provides the users with a wide range of musical possibilities.

This algorithmic 'music module' design is advantageous for two reasons. Firstly it presents a shallow learning curve and makes it possible to train participants in the use the software in a short amount of time, and secondly, the simplified music production environment enables more research emphasis to be placed on aspects of collaboration; a view echoed by Laney et al. (2010), who presented users with samples of pre-composed musical parts. Despite the software being designed to present users with a shallow learning curve, it is important to stress that the software is intended for musicians rather than novice users, and as such requires a certain degree of domain knowledge, for instance an understanding of musical terminology, and music technology related terms. The details and properties of each music module are presented below:

**Bass Drum** The Bass Drum module creates a low frequency percussion sound which occurs on

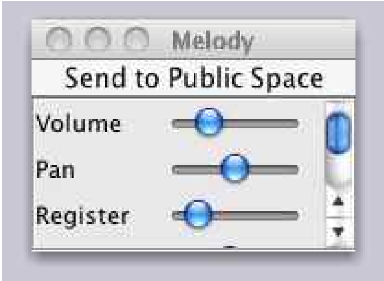



Figure 5.1: A 'Melody' music module.

the 4th beat divisions of the bar, according to the *Steady* parameter. The user can also set the pitch and tone of the Bass Drum sound.

**High-Hat** The High Hat module creates filtered white noise pulses on 2nd, 4th or 8th divisions of the bar, according to the *Speed* parameter. The timbre is also configurable via a tone control.

**Snare Drum** The Snare Drum module creates a percussion sound which may occur on the 2nd or 4th beat of the bar. The probability of the sound being triggered is set using the *Steady* parameter. The timbre and pitch of the sound are also modifiable. A final parameter is *Jitter*, which causes additional, low volume instances of the Snare Drum sound to occur on 16th divisions of the bar.

**Pulse** The Pulse module creates a single musical tone, which can be placed at any position within the bar, according to the *phase* parameter. The frequency and tone and the amount of modulation or *Wobble* applied to the sound can also be configured using the module parameters.

**Bass Pulse** The Bass Pulse creates a low frequency tone at the beginning of the bar. The timbre and pitch and attack of the Bass Pulse can be set by the user.

**Melody** The Melody module creates a melody by choosing notes from a pre-defined set of pitch classes. The *Variety* parameter sets how many of these pre-defined pitches are used, where a low variety causes only a few notes to be used, and a high variety setting allows the algorithm to introduce a wider variety of notes into the melody. The *Register* parameter selects which octave the melody occupies, while the tone of the synthesis voice is also configurable via a tone slider.

**Texture** The Texture module creates a continuous synthetic tone, whose timbre and modulation can be set.

### 5.1.2 Workspaces and Interface Variations

The study used three interface designs, each of which presented the user with different forms of graphical workspace. The workspaces are on-screen areas to place and manipulate the music modules. In all interface designs the participants share a Public Workspace which is consistently reflected across all Client computers, meaning the coordinate position of the music module is identical for all Clients, and changes to music module parameters are immediately updated for all Clients. Music modules present in the public workspace are editable, visible and audible by any collaborator within the software environment. Figure 5.2 shows the interface with only the Public workspace is enabled. The Public space only condition adheres to a strict WYSIWIS approach, where all users see (and hear) an identical representation of the shared interface (Stefik et al., 1987).

The second and third conditions of the study add a Private workspace for each connected computer. When the Private workspace is included in the interface it appears below the Public workspace (see Figures 5.3 and 5.4) and offer each user with an additional area to deploy music modules. Music modules within a user's Private space are only editable by that user. Modules can be freely transferred between the Public and Private spaces using a button located at the top of each module window (see Figure 5.1). A module can only exist in a single space at any one time and it is not possible to directly transfer modules between two participants' Private spaces. To achieve this modules must be transferred via the Public space. Figure 5.3 illustrates the interface with the Private Space enabled.

The additional inclusion of a 'View' awareness mechanism in the third condition allows users to see and audition each other's Private workspaces, thus reducing the level of privacy offered by the Private Space. When the 'Views' option is enabled users can use a tabbed window pane to select either their own Private space or the Private space of one of their collaborators. Figure 5.4 illustrates the interface with the Private Space and Views tab awareness mechanism.

The interface variations (Public only, Public+Private, and Public+Private+Awareness) form the basis of the experimental conditions outlined in Section 5.3, and are selected through command-line options at runtime. Figure 5.5 shows the interface with the Public, Private and Awareness features enabled.

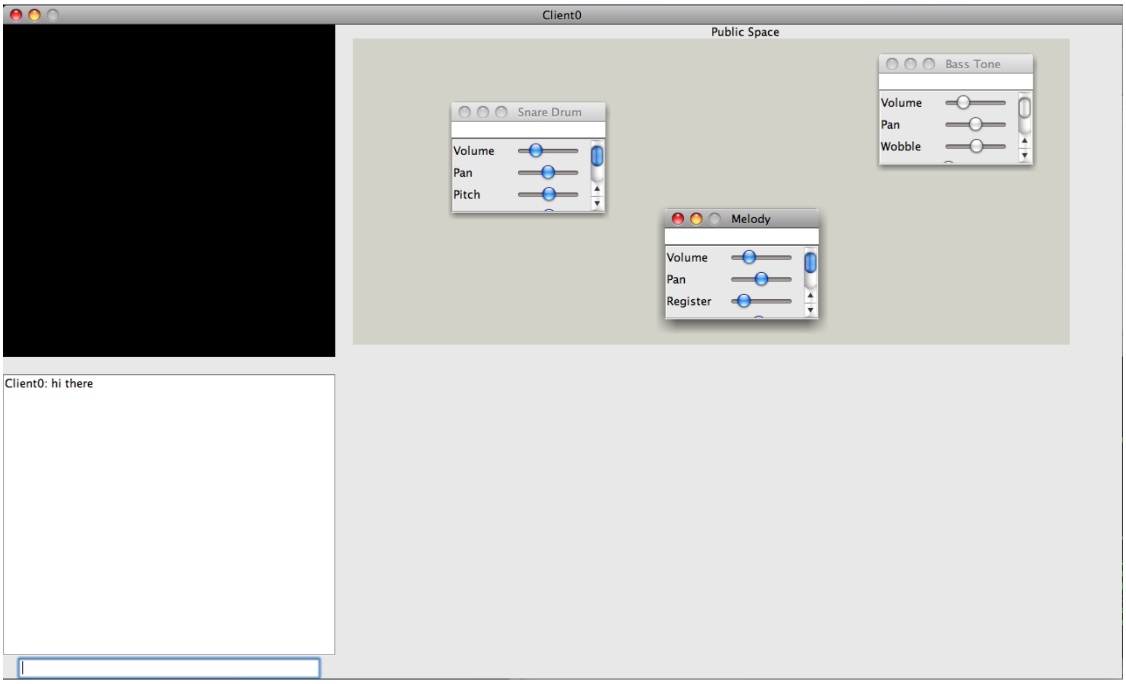



Figure 5.2: Screenshot of Study One collaborative music software, with C0 Public Only condition enabled.

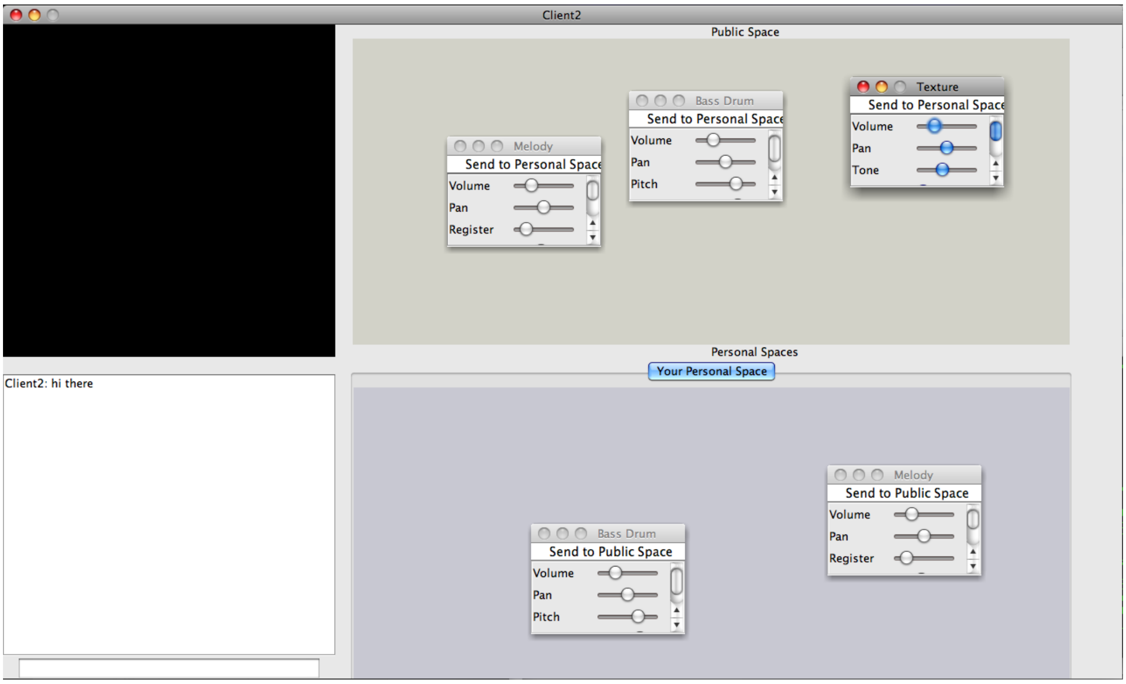



Figure 5.3: Screenshot of Study One collaborative music software, with C1 Public plus Private Workspace condition enabled.

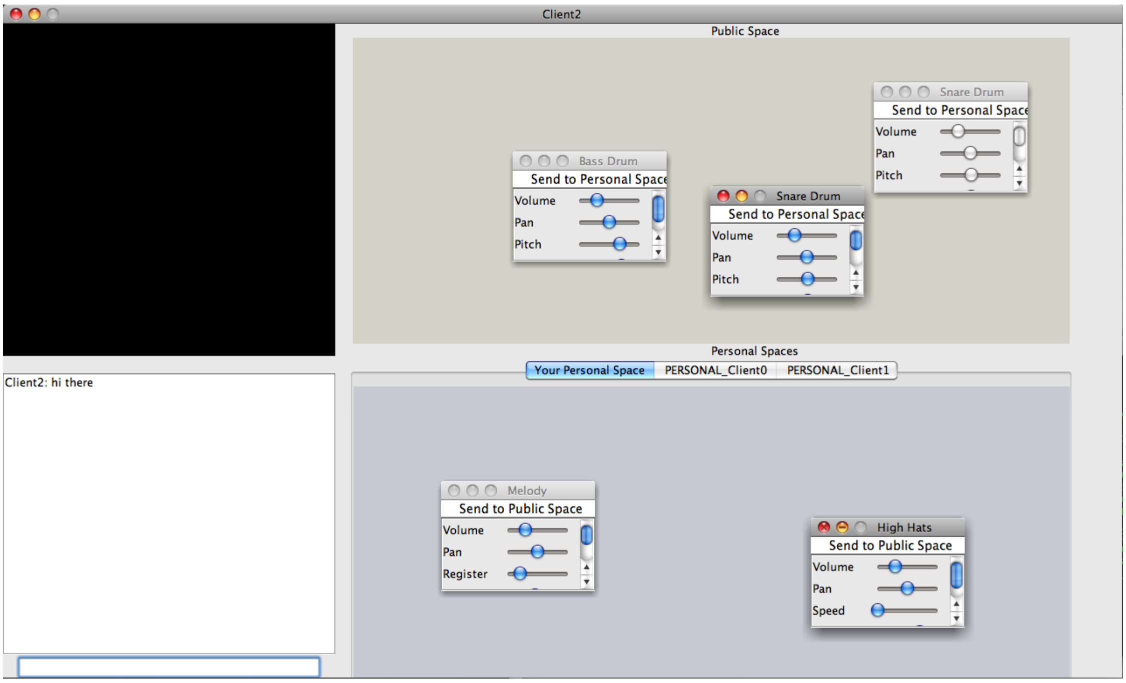



Figure 5.4: Screenshot of Study One collaborative music software, with C2 Public + Private + Awareness condition enabled.

### 5.1.3 Chat Tool

The software provides participants with a text-chat tool for communication. The chat tool is located to the left of the screen. Chat messages are typed into a small text field, and sent by pressing the return key. Messages appear immediately on all Client screens, although author names are anonymised in the chat tool, for example appearing as 'Client 1'. These anonymised Client names are based on the order Client workstations connected to the Server application. The text chat tool was situated on the left hand side of the interface (Figure 5.5).

### 5.1.4 Awareness information presented via the software interface

The software provides several forms of awareness to users. The public workspace displays the modules currently shared across all users, and therefore provides an indicator of when modules are created or deleted. Where Private workspaces are available to users, the interface does not visually distinguish between modules which are created in the Public space or modules which have been created in a Private space and subsequently transferred into the Public Space. Similarly, for a given user, the interface does not visually distinguish between modules which have been deleted or modules which have been moved into another users' Private space.

Within the Public workspace, the consistent window positioning across all connected Clients provides awareness of when another user re-positions a module. The movement of Music Module sliders is also immediately reflected to all Clients, providing an instantaneous indicator of which modules are currently being edited, however the software does not indicate *who* is performing the edit operation, and the software does not provide any indication of who is responsible for which musical contribution. A separate study examines Authorship awareness, and is presented in Chapter 7.

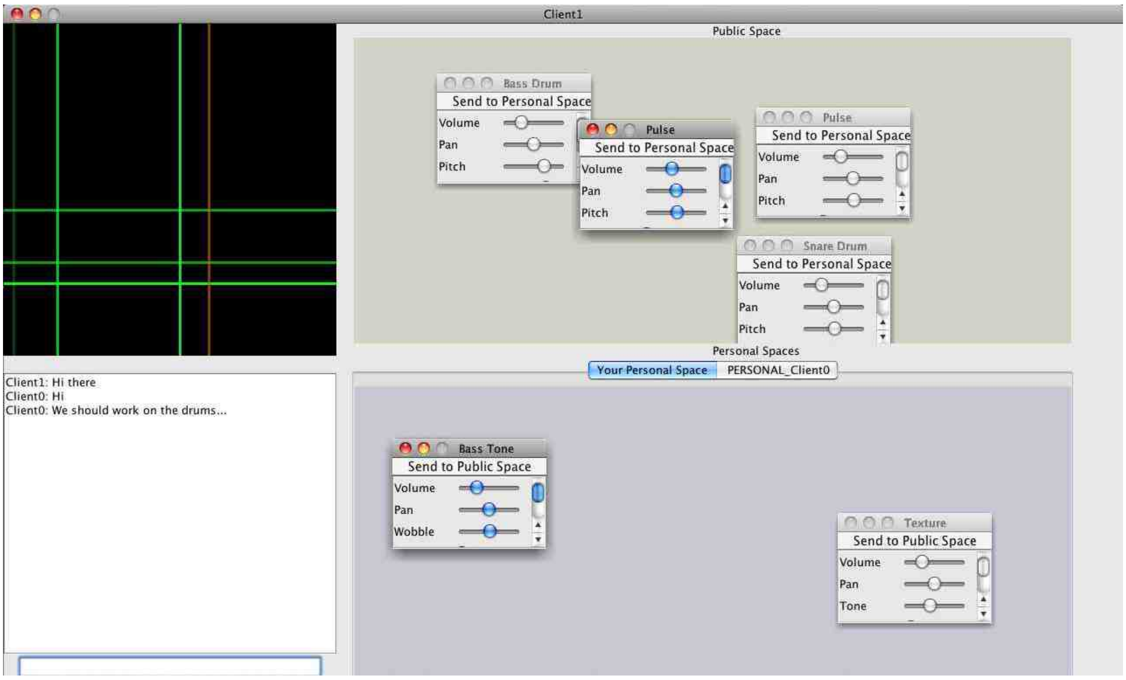



Figure 5.5: Screenshot of Study One collaborative music software. Video-window top left, chat window bottom left, Public Space top centre, Private Space bottom. 'View' tab selector centre.

### 5.1.5 Data Logging

Data about usage is logged by the Server and stored in three text files, a *Supercollider Log*, a *Client Log* and a *Server Log*. The SuperCollider log stores all synthesis messages sent by the Server and can be used to re-synthesise the music created by the software, however it contains no information about the state of the GUI. The Client Log stores all messages received by the Server (i.e., sent by users), and includes all messages related to GUI operations performed by the users, including the creation and deletion of music modules, editing, chat messages and changes to the on-screen co-ordinate position of modules. The Server Log contains all non-music related messages sent from the Server to connected Clients. Log messages include updates to the position of module sliders in the GUI and text chat message. All log events are time-stamped using the

Server computer's system clock.

## 5.2 Hypotheses

This study addressed the research question *how do interface features designed to support privacy and awareness alter the way groups of co-located musicians engage in collaborative musical interaction?* This research question was used to derive the following hypotheses:

- H1 The inclusion of privacy will cause measurable differences in the way the software is used.
- H2 Increased awareness will promote more collaborative use of the software, as evidenced through a greater degree of collective editing in the Public space

These hypotheses are not intended to attach value to any particular outcome. For example although hypothesis H2 posits that there will be a higher incidence of joint or collective editing (where group members edit each-others contributions) when awareness has been increased, the wording of the hypothesis is not intended to favour outcome over any other. This standpoint is also reflected in the design implications presented in Chapter 8, where it is stressed that many of the implications should be viewed as design tradeoffs between the effects of presenting certain kinds of awareness information within interfaces for CDMI.

## 5.3 Independent variables

The study design uses three experimental conditions, each of which provides different opportunities for participants to share and conceal their musical contributions from one another while collaborating. The conditions directly manipulate the types of workspaces made available to the participants. The conditions are *Public Only* only, *Public+Private*, and *Public+Private+Awareness*, and are labelled C0, C1 and C2. The study used a within-subjects design where each participant group was exposed to all three conditions. The condition ordering was randomised between experimental sessions to counteract ordering effects. The experimental conditions were as follows:

**C0: Public Only.** In C0 all music modules, and music module editing is constantly visible and audible to all participants via a single shared Public workspace. All participants see an identical copy of the graphical workspace, and any changes are immediately updated across all computer workstations.

**C1: Public+Private.** In C1 participants have the Public Space, as in C0, however participants are each also provided with a Private space. Each participant can see, hear and edit music modules contained in their Private workspace, but they do not have any access to the Private workspaces owned by other participants. Music modules can be created in either the Public or Private space, and can be freely transferred between spaces.

**C2: Public+Private+Awareness.** C2 builds on C1 by allowing participants to view and audition the contents of each other's Private spaces, but not edit music modules within them. As in C1, musical contributions can be freely transferred back and forth between a participant's Private Space and the Public Space. At any one time participants are able to view the Public Space and one Private Space, which can be their own, or the Private space of another user.

## 5.4 Dependant variables

The hypotheses from Section 5.2 were operationalised into the following dependent variables:

**Interaction with the software** Measures of interaction with the software are drawn from interaction log analysis collected by the software. This includes features such as the number of music modules created, the amount of music module editing, instances of editing modules created by someone else (co-editing), proportion of modules in each workspace and time spent viewing other user's Private spaces.

**Individual preferences** Measures of preference are collected via questionnaires presented to the participants after they have experienced the three interface conditions. Group discussions in the final stage of the experiment session also explore preferences, although the nature of focus group style discussions may influence the preferences reported by participants.

**Group strategies and discussion** Video observation was used to study the groups' working strategies and analyse verbal discussion during the musical interaction.

The rest of this section describes in more detail the data collection methods employed, and the justifications for choosing these methods.

### 5.4.1 Questionnaire Data

Demographic information was collected from participants via a pre-test multiple choice questionnaire. This questionnaire was in place to collect general details about participants', such as age and occupation, and information about musical experience, musical proficiency and knowledge of other collaborative computer software.

A post-test multiple choice questionnaire was administered to gather opinions and impressions about the experimental conditions. The questionnaire was based on the Mutual Engagement Questionnaire (Bryan-Kinns, 2012, Bryan-Kinns and Hamilton, 2009), which requires participants to select which experimental condition applies most strongly to a list of statements. Participants were presented with the statements described below:

**The best music** This question is designed to uncover preferences for the music created by the group, and potentially uncover relationships between the preferred music and the interface condition and the order the conditions were presented in.

**I felt most involved with the group** This question is designed to uncover participants' feelings of involvement, and how this may have been affected by the experimental conditions, as well as the influence of time spent during the session and their familiarity with the music software.

**I enjoyed myself the most** This question is designed to gauge the effects of interface conditions and ordering on reported enjoyment.

**I felt out of control** This question is designed to uncover the effect the interface designs had on participants' reported feelings of control.

**I understood what was going on** This question is designed to probe participants' self reported sense of awareness of what was happening during the interaction.

**I worked mostly on my own** This question is designed to uncover the effect of the interface designs on participants' reported feelings of working individually.

**We worked most effectively** This question is designed to gauge how well participants felt they worked together, and how this may have been influenced by the interface conditions, ordering effects and time spent using the software.

**Other people ignored my contributions** This question is designed to uncover the effect of the interface designs on participants' reported feelings of working individually.

**The interface was most complex** This question is designed to uncover participants' interpretation of the interface complexity, and is included partially to determine if participants identified any differences in the interface designs presented to them.

**I knew what other people were doing** This question is designed to probe participants' self reported sense of awareness of others in the group.

**I felt satisfied with the result** This question is in place to gauge the participants' satisfaction with the music they created, and the relation this has to the interface conditions, ordering effects and the time spent using the software.

**We edited the music together** This question is designed to gauge the degree to which the interface conditions contributed to participants' feelings of group cohesion.

### 5.4.2 Interaction Logs

The Server component of the music software logs all user interaction, including module creations, module deletions, parameter modifications, module position changes, button presses, slider movements and text-chat entries. Each event log entry contains a time-stamp and an author name. Programs were developed in Java to autonomously reconstruct the user interaction, providing a means to parse the logs and extract various details of how the software was used by the participants. Data from these scripts was saved to a comma separated value (CSV) formatted text file for further analysis using a spreadsheet. The following features were identified as useful to study:

**Module Creations** A record of each module creation during the interaction. Module creation log entries include the time the creation occurred, the type of module created, the user who created it, which workspace it was created in, and the coordinate position within the workspace. The number of music modules created by an individual provides an overall estimate of the amount of ideas contributed during the interaction, and can be broken down by workspace location to indicate which workspace an individual was using to create their contributions. The spatial location of creations can be used to indicate territory and patterns of spatial usage within the interface, while the temporal information can be used to plot the frequency of creations over time.

**Music Module Parameter Editing** Parameter editing includes the manipulation of Music Module sliders and changes to the coordinate position of a module within a workspace. An Edit log entry includes the Client performing the edit, a timestamp, the parameter being edited, the new value and the module it is associated with. Using additional knowledge about the state of the interface at that time the edit occurred it is also possible to determine the coordinate location of the module being edited, which can contribute to identifying patterns of spatial usage within the interface.

The continuous nature of the GUI sliders means a single slider movement will incur multiple log entries, which is also affected by the speed of slider motion. One implication of this is that different participants may generate different quantities of log entries depending upon their style of slider interaction. To account for this, a time-based filter was applied to the logs to distinguish between separate edit operations using a one-second time delay to differentiate between the end of one edit operation and the start of a subsequent edit operation.

Two types of music module parameter editing were identified, *Self-Edits* and *Co-Edits*. A self-edit occurs when a participant edits a module they initially created. A co-edit occurs when a participant edits a module created by a different user. A measure of Co-editing can indicate the degree to which music modules are shared between users.

**Module Deletions** A record of each module deletion during the interaction. Module deletion log entries include the time the deletion occurred, the identify of the module deleted, and the user who performed the deletion. Using additional information from the log file it is possible to determine which workspace the module was in when deleted, and its coordinate position within the workspace.

**Transfer** Instances of module transfers between workspaces are logged. The log entry for a transfer includes the module being transferred and the space it was transferred to. Transfer information can be used to characterise the use of the workspaces, by for instance identifying if modules were more frequently transferred in or out of a workspace. Additionally the time modules spend in each workspace can be inferred from information about module transfers, creations and deletions.

**View Activity** The amount of time each user viewed their collaborators' Private spaces can be extracted from the log files. This information can be used to characterise users' use of the Views awareness mechanism.

**Spatial Use and Division of the GUI** The study software records information about the spatial use of the GUI, including the spatial movement of music modules and the position music modules were created at. This spatial information can be presented as visualisations of module locations and positions for creations, deletions and editing over the course of the interaction. Colour coding can be used to distinguish between the three participants. Within this study, analysis of the spatial interface arrangements users adopted was considered exploratory, however Study Two (Chapter 6) presents a more thorough analysis of the way the shared screen space was used as a collaborative resource.

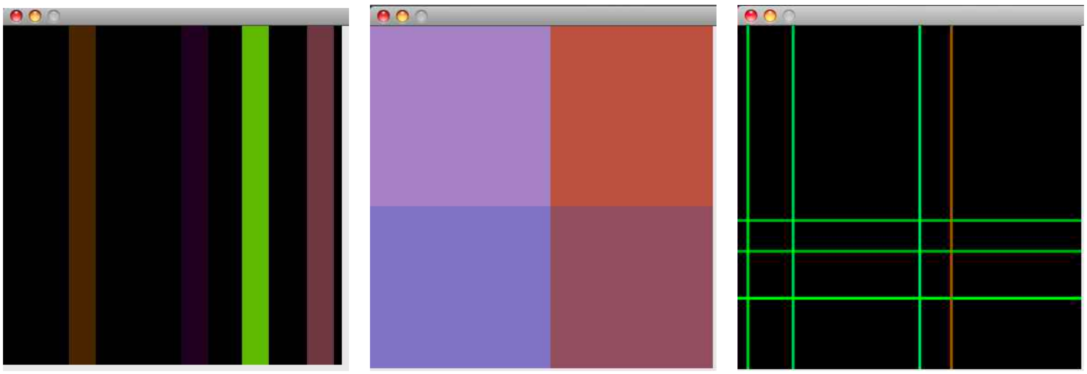
### 5.4.3 Video and Group Interviews

Each experiment session was video recorded to capture the discussions participants engaged in whist using the software. A group discussion was held at the end of each session. All dialogue from participants and interviewer were manually transcribed using the ELAN video annotation software [1]. This resulted in 11,300 words of dialogue. The discussions were structured around the prescribed themes of the *preferred interface*, *maintenance of awareness*, *the formation of roles* and *approaches to composing for the video loop*. The participants were also encouraged to talk about their experiences more generally, and were invited to ask their own questions about the research at the end of the group interview.

## 5.5 Experiment Task

Participant groups were set the task of working together to create music to compliment a short video loop. The task was put in place to give focus to the participants' collaboration by setting a common goal to work towards, encouraging discussion, and providing opportunities for consensus, disagreement and uncertainty during the course of the group collaboration. To give participants a fresh challenge with each interface condition, three different video loops were used, each with a similar aesthetic style (see Figure 5.6[2]). Each loop lasted around twenty seconds and was comprised of moving coloured geometric shapes. The videos were displayed in the top left corner of their graphical interface (see Figure 5.5) and were synchronised between participants, meaning all participants view exactly the same point in the loop at any given time. The videos were generated using an instance of Processing [3] which was embedded into the graphical interface. To eliminate ordering effects the videos were presented in a different order for each experimental session and ordered independently of the interface conditions.

[1] http://www.lat-mpi.eu/tools/elan/
[2] videos downloadable from http://www.eecs.qmul.ac.uk/~robinfencott/thesis_resources/
[3] http://processing.org/ Accessed 30 September 2012



(a) Video 1 (b) Video 2 (c) Video 3

Figure 5.6: Screenshots of videos presented to participants during the study

## 5.6 Pilot Study

A pilot study was carried out prior to participant recruitment. The pilot study was used to identify problems with the experimental design, test the data collection mechanisms, highlight software bugs and provide an opportunity to rehearse the training session and presentation of information to the participants. Pilot study participants were recruited from within Queen Mary University of London, and took part in an experimental session of exactly the form described above.

### 5.6.1 Software Bugs and Usability Issues

The pilot study identified a number of usability problems and software bugs which had not been apparent through informal testing during the development process. The following section describes these issues and the implemented solutions.

**Music Module Positioning** It was noted that music modules could be positioned outside of the on-screen workspace in such a way as to prevent them from being accessible or movable once the mouse was released. This encouraged participants to create more music modules than required, and as a consequence the resultant music was considered too chaotic. The solution to this bug was for music modules to automatically track their on-screen position and alter their position if placed outside bounds of the workspace.

**Software slowdown** It was noted that if too many music modules were created the software began to slow down and potentially crash. As a control for this, an upper limit of 25 was applied to the number of music modules that could be created at any given time.

**Chat Tool Bugs and Usability** Two bugs were identified in the way the chat tool parsed and displayed messages. The chat tool window was modified to correctly display messages using text wrapping automatic window scrolling.

**Musical expressivity in the software** The pilot study participants stated that the system did not allow them to be very creative or expressive. Improvements and modifications were discussed, and a number of changes were implemented to address this issue. Additional parameters were added to the music modules as follows: to introduce more control over the audio mix volume and stereo panning position sliders were added; additional tone controls were added to the other modules to provide more timbral variety; finally, the Pulse module design was revised through discussion with the pilot study participants to allow control over the pitch and bar position of the pulse. Multiple pulse module instances can be used to create melodies or rhythmic parts.

### 5.6.2 Experiment Conduct

The following alterations were made to improve the running of the experiment:

**Pre-test questionnaire** Typographic errors with the pre-test questionnaire were corrected, several questions were identified as ambiguous.

**Experiment 'Script'** A detailed script or plan was written to structure the introductory briefing and explanations given to participants. A list of key points was formulated, including ethical issues, the structure of the session, and what should be said to participants between experimental conditions.

**Music software training and free time** Explaining to the participants how to use the music software was difficult. After running the pilot, a more structured training session was developed to step participants through the interface components and the process of creating music. It was also noted that participants needed some unstructured time to explore the system before starting the experimental conditions.

## 5.7 Participants and Recruitment

Nine groups of three participants (27 participants in total) were recruited for the study. All participants were recruited via mailing lists related to art, electronic music, sonic arts, music technology and academic research concerning sound and music computing. The recruitment e-mail stated that a new music application was being developed at Queen Mary University of London, and that for testing purposes, the project required 'people with an interest in creating music, for instance composers, musicians, DJs, and students of Music, Music Technology or related fields'. The message stressed that people would work in a group with two other people, and that they did not need to be able to play an instrument or have formal musical training to take part. The message invited people to make contact via e-mail if interested. In line with ethical procedures for the recruitment of human subjects, no individuals were directly approached or asked to participate. Participants received £10 financial compensation for taking part in the experiment. For practical reasons the participants were assigned to groups on a first-come-first-serve basis, taking into account their availability to attend the study session.

17 participants were male (65%) and 9 were female (35%). The average (mean) age was 28 years (SD = 7.4), and the modal age was 21 years old. 11 participants (41%) were students, 4 participants (15%) were unemployed, 1 participant (4%) was self-employed, and 10 participants (37%) were in employment. One participant (4%) did not provide details of their employment status. 24 participants (89%) could play a musical instrument, 2 participants (7%) described their level of proficiency as 'beginner', 8 participants (30%) described their level of proficiency as 'intermediate, 9 participants (33%) described themselves as 'semi-professional', 4 participants (15%) described themselves as 'professional' and 4 participants (15%) gave no response. 24 participants (89%) reported having written musical compositions on their own. Of these, 7 participants (29%) had written between 2-5 pieces of music, 2 participants (8%) had written 5-10 pieces, and 15 participants (63%) had written more than 10 pieces. 19 participants (70%) reported having written musical compositions with other people. Of these, 1 participant (5%) had jointly written one piece of music, 11 participants (58%) had written jointly 2-5, 2 participants (11%) had jointly written 5-10 and 5 participants (26%) had jointly written more than 10. Figure 5.7 shows the participants reported musical preferences.

When describing their level of computer literacy, 2 participants (7%) selected 'beginner', 12 participants (44%) chose 'intermediate' and 13 participants (48%) described themselves as 'expert'. 1 participant (4%) used the internet once a week, 8 participants (30%) used the internet once a day, and 18 participants (52%) used the internet every hour. 16 participants (59%) had previously used multi-user computer software such as collaborative document editors or online games.

In summary, participants recruited for the study generally had a high level of musical experience. Almost all participants could play a musical instrument, and the majority had previously written music, either by themselves or with other people. The majority of participants were also experienced in the use of computers.

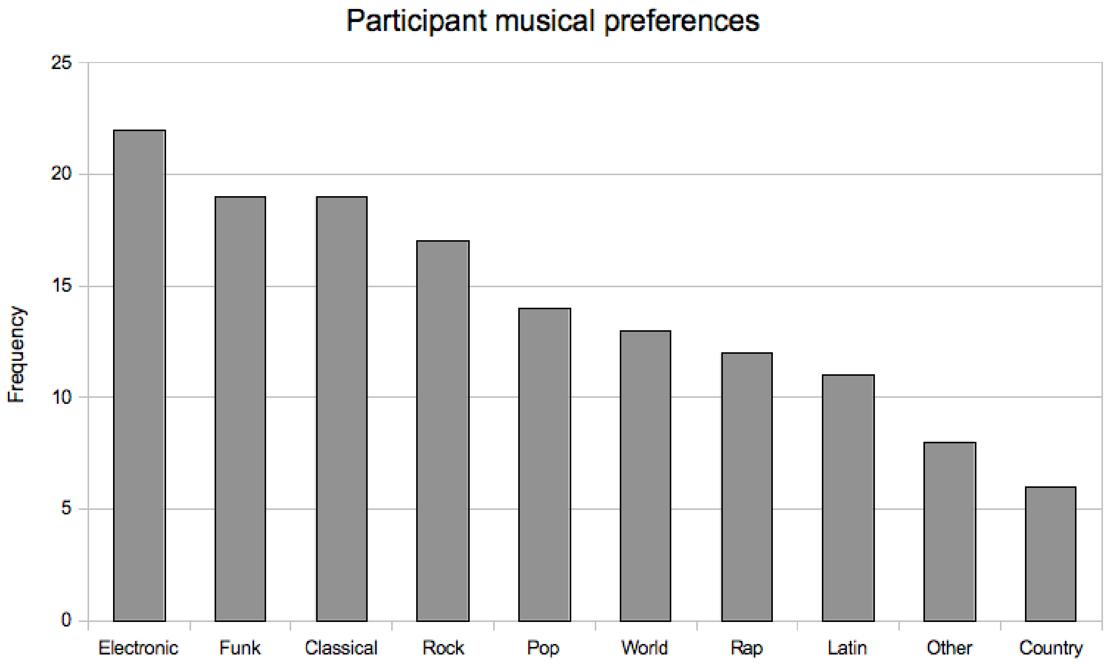



Figure 5.7: Participant musical preferences.

## 5.8 Procedure and Apparatus

Apple Mac computers were used to run the collaborative software environment for making music, each with a screen resolution of 1280x800. The computers were arranged around a table, and chairs placed at the table. Audio was presented through headphones connected to the computers.

Participants were asked to meet in a reception area close to the experiment laboratory. Once all participants were present they were led to the experiment laboratory, and asked to take a seat around the computer table. The researcher then sat at a chair near the table and presented participants with information sheets and consent forms. The researcher then proceeded to verbally describe the purpose of the study session, stressing the voluntary nature of the research participation. In line with research ethics guidelines, the data collection procedures were explained, along with reassurance that all data would be anonymised, and that images, video and audio from the research would not be published or made available online.

Participants were encouraged to ask questions throughout the introductory explanation. They were then given time to read the information sheet and consent form, and ask any further questions before being invited to sign the consent form. After signing consent forms the participants were given the pre-test paper based questionnaire to collect demographic information and details about their previous musical background, computer literacy and experience using collaborative computer software.

Participants were then given a training session with the software, which covered how to create, edit and delete music modules, how to use the chat tool, the use of the Private workspace, and the way modules could be transferred between these spaces. The location of the video loop was also made clear to the participants. During training the participants were given a short amount of free time to use the software (around five minutes), however time constraints prevented more time from being permitted. Throughout the training participants were invited to ask questions about the software. After training the groups were given fifteen minutes with each software condition. To control for learning effects the order of conditions and video loops was randomly assigned within each group.

After switching interface conditions, the researcher pointed out to the participants the difference between the current interface and the previous, and answered any questions that were asked. When stating these differences, the researcher focused on the objective differences, and did not present the interface features as intended for a particular purpose. Once all three conditions were presented the researcher handed out paper based post-test questionnaires as described in Section 5.4.1. After completing the final questionnaire a group discussion was conducted by the researcher. Towards the end of this discussion the participants were invited to ask questions of the researcher. At this stage the researcher explained more generally about the research being conducted, and the purpose of the study. Before leaving, the participants filled out a form to acknowledge receipt of £10 financial compensation. Finally the participants were thanked for their time and led to the exit.

## 5.9 Sample Size and Post-Hoc Effect Size Analysis

The Friedman test was selected as an appropriate test for the analysis of data from a three condition within subjects design. As the Friedman Test is used several times within this thesis, Appendix A presents a description and justification for the selection of this test.

The Friedman test is not widely used in HCI, CSCW or computer music research, and there are few guidelines for performing *a priori* sample size estimation (as was performed for Study Three of this thesis - see Chapter 7). Such calculations also require estimated mean and standard deviation based on prior studies, related literature or expert judgement (Prajapati et al., 2012). As this study was the first of its kind, such data was not available. In addition, recruiting large numbers of human participants for HCI studies is a challenging task Rogers et al. (2009). The sample size for Study One was therefore driven by the availability of participants. As noted about, 27 individuals were recruited and placed into nine groups of three participants.

To validate the significance of the findings given this sample size, significant results from the Friedman test were analysed using a post-hoc effect size analysis technique. This approach has been proposed for situations in which a sample size cannot be estimated *a priori*, or in cases where it is not possible to recruit a large enough sample in the first place (Prajapati et al., 2012).

The post-hoc test performs pairwise comparisons using the Wilcoxon signed-rank test to identify which pairs of conditions out of the three caused a significant effect. A Bonferroni correction is applied at this stage to account for the increased chance of detecting a significant result when multiple pairs are compared. The final part of the post hoc analysis is to calculate the statistical effect size for each pair, taking into account the sample size using the formula:

$$r = \frac{Z}{\sqrt{N}}$$

Where r is the effect size, Z is produced by the Wilcoxon Signed-Rank test, and N is the number of matched pairs in the analysis. In this instance N=27, as there are 27 sets of within-subjects data (Corder and Foreman, 2009) The effect size shows the degree of association between the groups (Corder and Foreman, 2009), and range from 0.0 to 1.0, with 0.10, 0.30 and 0.50 being widely accepted as small, medium and large sizes respectively (Corder and Foreman, 2009).

## 5.10 Results

### 5.10.1 Interaction Log Analysis

The Friedman Test was used to compare the number of creations in the Public Space between all conditions. The Friedman test is a non-directional, non-parametric statistical test used to identify differences across multiple test treatments within a repeated measure experimental design (Greene and D'Oliveira, 1999). As the Friedman test is used at several points throughout this thesis, Appendix A presents a description of this test, and provides justification for its use.

Significantly more creations occurred in the Public space in C0 than in C1 or C2 ($p<0.0001$, df=2 $csq_r=25.8$). A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 ($p<0.001$, r = 0.8614851) and between C1 and C3 ($p<0.001$, r = 0.7548208), indicating large effect sizes for both comparisons.

The Friedman Test was used to compare the total creations in all spaces. Significantly more creations occurred in total in C0 (p=0.0029, df=2, $csq_r=11.69$), with no statistical difference between C1 and C2. A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 ($p<0.001$, r = 0.6577167) and between C1 and C3 ($p<0.001$, r = 0.6763419), indicating large effect sizes for both comparisons.

For conditions where the both Public and Private workspaces were available (Conditions C1 and C2) the Wilcoxon Signed-Ranks Test was used to compare the number of module creations which took place in the Public Space compared with creations in either Public+Private condition (C1) or Public+Private+Awareness condition (C2). The Wilcoxon Signed-Ranks Test is a non-parametric test for comparing related samples on a single sample population (Greene and D'Oliveira, 1999). In this instance, the test was used to perform a pairwise comparison between the data for two of the experimental conditions. The the two data samples are *creations in Public* and *creations in Private*.

In both cases significantly more module creations took place in a participant's Private Space than in the Public space (for C1: p=0.0001, w=-331, z=3.97. For C2: p=0.0002, w=-307, z=-3.68). Figure 5.8 shows the mean module creations in each workspace across all conditions, demonstrating the strong tendency for module creations to be made in the Private workspaces when these workspaces were present in the interface.

The Friedman Test was used to compare the amount of module editing which took place in the Public Space between all conditions. Significantly less editing overall took place in condition C0 (where participants had only a Public Space) than in conditions where participants also had a Private Space (p=0.0344, df=2, $csq_r=6.75$). A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 ($p<0.001$, r = 0.5155887) and between C1 and C3 ($p<0.001$, r = 0.5363585), indicating large effect sizes.

There was no statistical difference in the amount of editing between conditions C1 and C2. See Figure 5.9.

The Friedman Test was used to compare the amount of module co-editing which took place

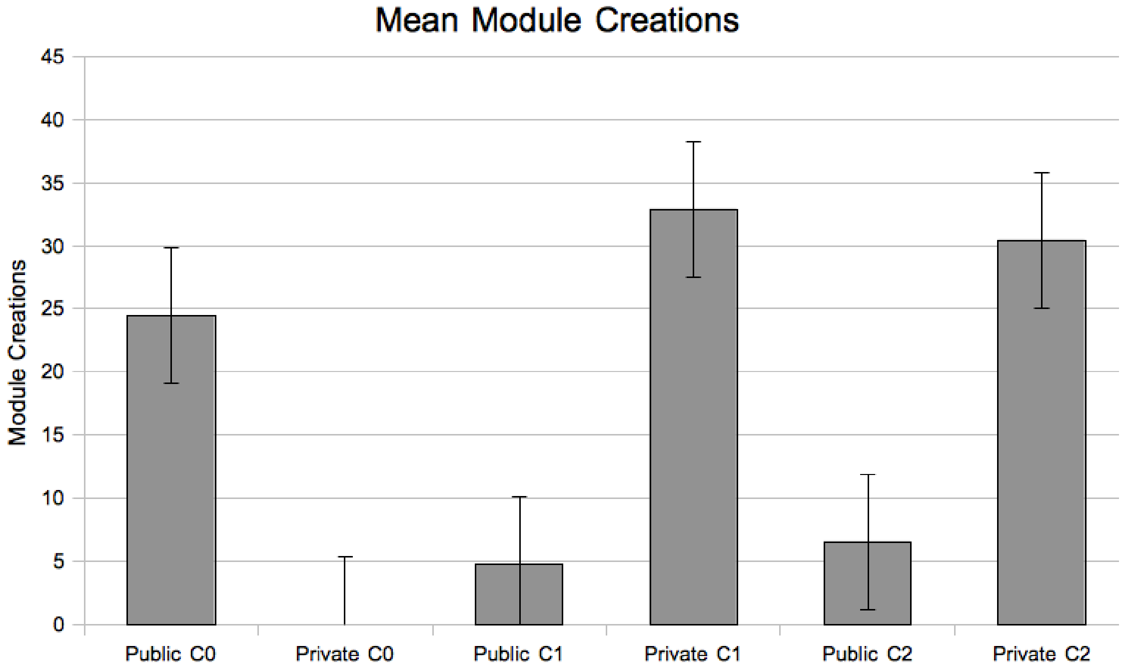



Figure 5.8: Mean number of module creations in Public and Private workspaces. Error bars show Standard Error

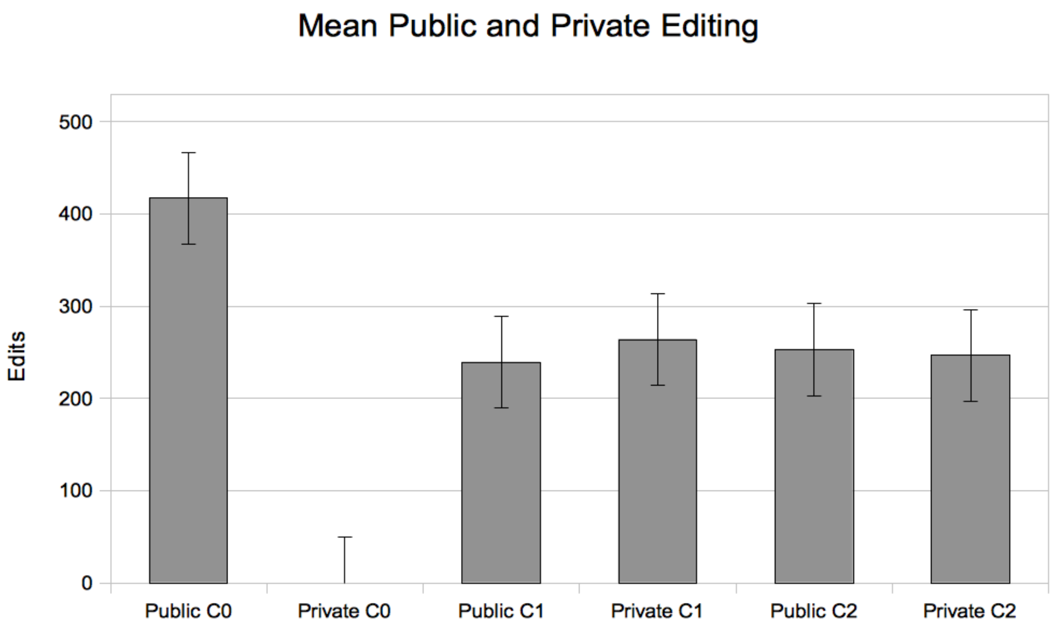



Figure 5.9: Mean Public and Private Editing. Error bars show Standard Error.

in the Public Space between all conditions. Co-editing was calculated per participant as a proportion of all the editing which took place. The Friedman test was then applied to test for significance on the proportion of co-editing per experimental condition. Significantly more co-editing took place in condition C0 (where participants only had a Public Space) than in conditions where participants also had a Private Space (p=0.0019, df=2, $csq_r$=12.57). A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 (p<0.001, r = 0.7259073) and between C1 and C3 (p<0.001, r = 0.6889184), indicating strong effect sizes.

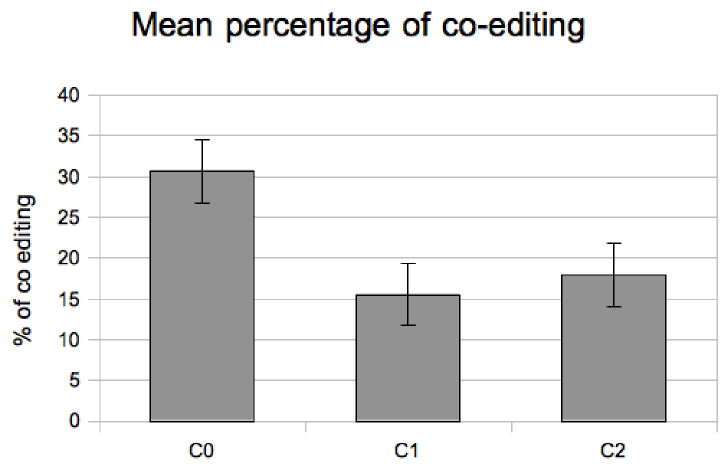



Figure 5.10: Mean amount of co-editing, calculated as a percentage of all editing. Error bars show Standard Error.

The Friedman Test was used to compare the amount of module deletions which took place in the Public Space between all conditions. Significantly more deletions took place in the Public space for condition C0 (where participants only had a Public Space) than in the Public Space for conditions where participants also had a Private Space (p=0.0293, df=2, $csq_r$=7.06). Figure 5.11 shows the mean deletions for each condition. A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 (p<0.001, r = 0.573964) and between C1 and C3 (p = 0.01, r = 0.4456084), in indicating medium and strong effect sizes respectively.

Module transfers between workspaces could only occur in conditions C1 and C2, as both of these interfaces feature two workspaces. The C0 condition featured only a Public space, so no transfers could occur. The Wilcoxon Signed-Ranks Test was used to compare the total number of transfers between conditions. There was no statistical differences between the overall number of

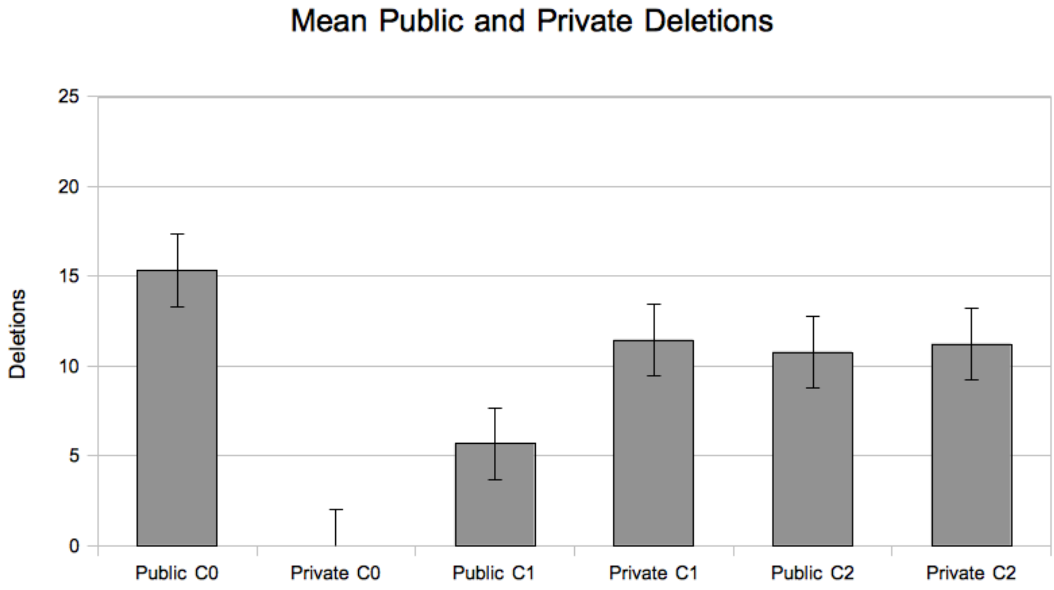



Figure 5.11: Mean Public and Private Deletions. Error bars show Standard Error.

| Feature | C0 | C1 | C2 |
|---|---|---|---|
| Public Module Creations | 220 | 43 | 59 |
| Private Module Creations | n/a | 296 | 274 |
| Public Module Deletions | 138 | 51 | 97 |
| Private Module Deletions | n/a | 103 | 101 |
| Public Module Edits | 3752 | 2152 | 2277 |
| Private Module Edits | n/a | 2375 | 2220 |
| Module Transfers to Public | n/a | 237 | 232 |
| Module Transfers to Private | n/a | 74 | 96 |

Table 5.1: Raw log file data

transfers, the number of transfers from *Private to Public*, or the number of transfers from *Public to Private*.

In Condition C2 participants spent on average 8.3% of their time viewing and listening to each others' Private Spaces. This equates to on average 75 seconds out of the total 15 minute session. Figure 5.12 shows the mean time (in seconds) participants spent viewing their own Private Space and the Private Space of other users.

### 5.10.2 Spatial use of workspaces

Visualisations were generated from the interaction logs to investigate how participants used the graphical workspace to organise and arrange their music modules within the shared workspace.

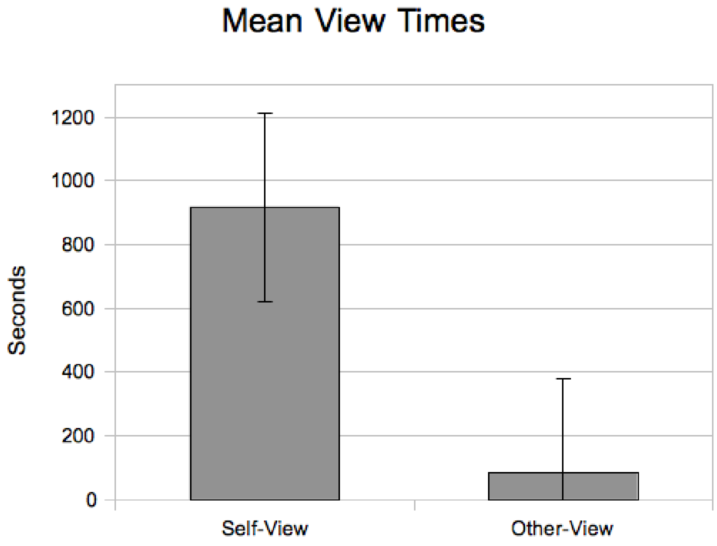



Figure 5.12: Mean time spent looking at own and others' Private Space (in seconds). Error bars show Standard Error.

Figure 5.13 shows an example visualisation. The full collection is included in Appendix B. The top box represents the Public space, and the three boxes below are the Clients Private spaces. From the visualisations it was not possible to identify any generalised patterns of behaviour. Chapter 6 presents a more detailed investigation of spatial arrangements within a variation of the interface used for this study.

### 5.10.3 Post-Test Questionnaire

| Statement | C0 | C1 | C2 | Total Responses | Chi test P value |
|---|---|---|---|---|---|
| The best music | 5 | 12 | 8 | 25 | 0.23 |
| I felt most involved with the group | 6 | 9 | 10 | 25 | 0.59 |
| I enjoyed myself the most | 5 | 13 | 8 | 26 | 0.07 |
| **I felt out of control** | 12 | 2 | 8 | 22 | **0.04** |
| I understood what was going on | 6 | 10 | 7 | 23 | 0.51 |
| **I worked mostly on my own** | 3 | 10 | 13 | 26 | **0.05** |
| We worked most effectively | 6 | 11 | 9 | 26 | 0.16 |
| Other people ignored my contributions | 10 | 6 | 4 | 20 | 0.22 |
| **The interface was most complex** | 7 | 3 | 14 | 24 | **0.01** |
| I knew what other people were doing | 8 | 2 | 11 | 21 | 0.06 |
| I felt satisfied with the result | 5 | 9 | 9 | 23 | 0.5 |
| We edited the music together | 4 | 11 | 8 | 23 | 0.2 |

Table 5.2: Post-Test questionnaire results for Study One. Statements with a significance level of $p<0.05$ are highlighted in bold.

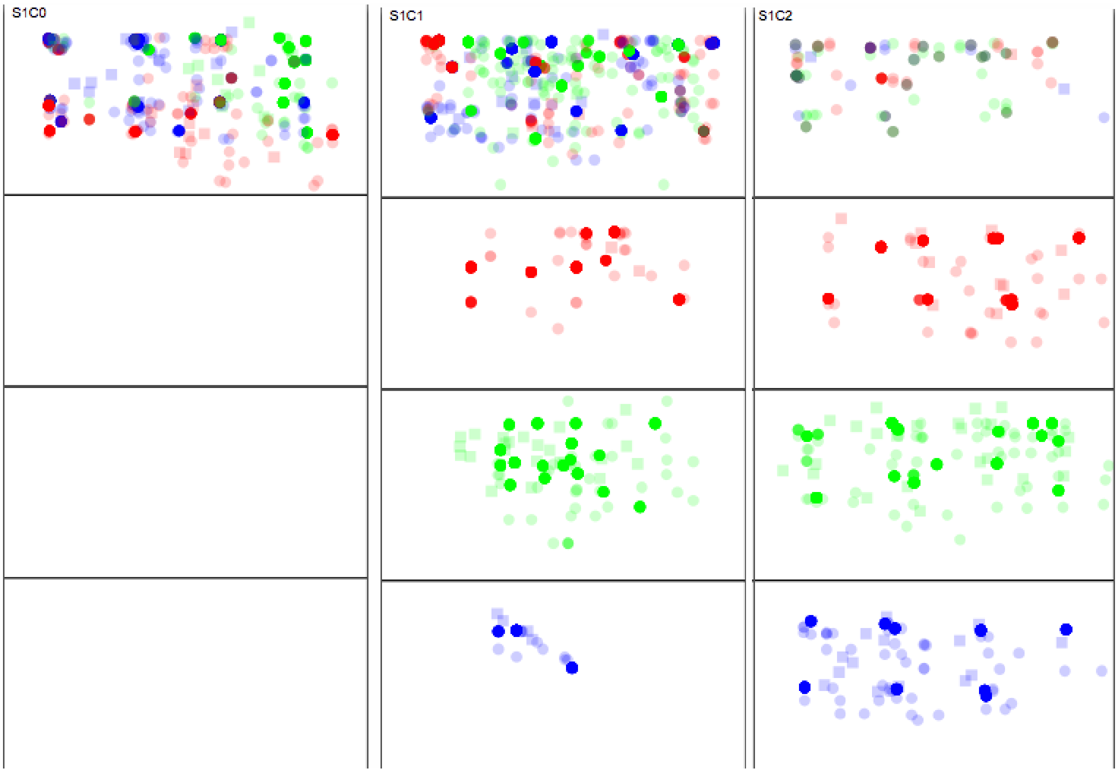



Figure 5.13: Visualisation of spatial usage for an experimental session. Experimental conditions are from left to right (C0, C1, C2). Circles represent edits, Squares represent module Creations, colours represent participants.

The chi-square test was used to determine the statistical significance of the post-test questionnaire responses. The chi-square test compares the observed and expected frequencies for multiple categories of data (Greene and D'Oliveira, 1999). In this instance, the observed frequencies are the responses from each participant to the statements in the post-test questionnaire, shown in columns C0, C1 and C2 of Table 5.2. If the experimental conditions had no influence on the experiences of the participants the expected frequencies for each questionnaire statement would be evenly distributed across all three choices. However, if one condition was frequently selected by participants for a given questionnaire statement the observed frequencies would differ from the expected frequencies, potentially indicating a significant effect within the data.

Using the chi-square test, the following statistically significant ($p<0.05$) effects were identified within the questionnaire data, summarised in Table 5.2.

- Participants felt 'out of control' the least in condition C1
- Participants stated working 'mostly on their own' the least in condition C0
- Participants stated that interface C2 was the 'most complex'

The following trends were also identified:

- Participants had least awareness of each other's activities in C1 (p=0.06).
- Participants enjoyed themselves least in C0 (p=0.07).

## 5.11 Discussion

This study set out to explore the role of privacy and awareness functions in a distributed software environment for CDMI, by addressing the research question *how do interface features designed to support privacy and awareness alter the way groups of co-located musicians engage in collaborative musical interaction?* This question was operationalised by the hypotheses stated in Section 5.2 and discussed below.

### 5.11.1 Hypothesis H1

H1 hypothesised that the inclusion of privacy would cause measurable differences in the way the software was used. This hypothesis was supported through a number of findings which indicate the inclusion of the privacy function caused participants to work more in isolation from one another. Comparing the location of module creations for conditions when both public and private workspaces were available shows that module creations happened almost exclusively in the Private Space when it was present within the interface. This finding suggests that participants used the Private Space to create musical contributions prior to sharing them with the other group members. Analysis of the post-test questionnaire also supports this finding, insofar as participants reported working mostly on their own in conditions which featured the privacy mechanism. This was true for the condition which also included the additional awareness mechanism, and the inclusion of the awareness mechanism did not appear to cause participants to work any less in their Private space.

Comparing all three conditions, significantly more co-editing occurred in condition C0 (Public Only). One explanation for this could be that in the Public+Private and Public+Private+Awareness conditions, users worked more on their own by creating and editing modules in their Private Space, therefore limiting the potential for co-editing as it was not possible to edit modules in other users Private Spaces. However, it may have also been the case that when participants were given the option to work in privacy, they concentrated on formulating finished contributions there, and used the Public space more as an area to share, and less as an area to edit the musical material. Several other results also support this notion. Firstly, significantly less editing took place overall when participants were presented with only the Public space, indicating that they were less exploratory when they were forced to work in public. Secondly, the interaction logs indicate that significantly more public deletions occurred in C0. This suggests that there was a higher turnover of contributions, with more contributions being rejected or discounted when participants could only work in the Public Space. One explanation for this is that when participants were forced to use the Public Space for experimentation, they were less inclined to spend time refining their contributions when they were in the Public space.

Further support for this claim can be found in the interview transcripts, where the Private Space was often described by participants as an area for experimentation and development, with participants often describing their Private Space as a place to 'prepare', 'sketch', 'test' and 'draft' contributions. These kinds of descriptions imply that the Private Space was used as an area in which musical ideas could be developed and refined before sharing, and that such a process was important part of the collaborative activity. Participants also described transferring or 'dropping' contributions into the Public Space once they were satisfied with the way they sounded. It is important to state that the participants were not lead to using the Private Space in this way during the training period at the start of the session. The following extracts from the interviews show these sentiments.

PARTICIPANT 2: You could prep. something, and say
PARTICIPANT 1: Here it is, and you knew no-one else had heard it before,
PARTICIPANT 2: Uh hu
PARTICIPANT 1: so you could prepare it and then drop it in, yeah.
PARTICIPANT 2: You could alter it a bit, so it's kind of a different mental space
PARTICIPANT 1: yeah, that's true. (Session 3 Interview)

PARTICIPANT 2: when I used Personal space, it was just like, drafting ideas, before I put it in there, just, yeah that, I didn't really look at that many people's personal spaces that much, it was just, oh yeah, see if they are doing the same thing as me, so I wouldn't have to, later on, but that was it really (Session 8 Interview)

In contrast to these descriptions, participants used terms such 'the main space' and 'the actual composition' to describe the Public Space Such descriptions suggest that the Public Space was regarded as the primary, shared focus of the collaboration, as illustrated by the following interview vignette:

INTERVIEWER: Which version did you prefer?
PARTICIPANT 1: the last one
PARTICIPANT 3: yeah, I might preferred the last one

PARTICIPANT 2: yeah, I agree
INTERVIEWER: why was that?
PARTICIPANT 3: I felt like I kind of worked on my own, and found bits that I liked, put it into the main space (Session 1 Interview)

The privacy mechanism did introduce some problems for participants. In particular, participants frequently noted being unsure which musical contributions were at group level and which were part of their Private space. This problem is in part caused by the separation within the interface between the graphical controls and the sonic output. For instance although the interface clearly separates the graphical representation of the musical contributions into public and private through their presence in one of the two workspaces, there is no auditory distinction between sounds which were in the Public Space and sounds in the Private Space. This issue was unforeseen in the design of the interface, but hints at an important design implication, that collaborative features such as Private Spaces also need to be designed to be usable from the perspective of a single user, and not only with concern for groups of people.

### 5.11.2 Hypothesis H2

The awareness mechanism introduced in condition C2 allowed users to view the contents of other users Private workspaces. Hypothesis H2 conjectured that increased awareness would promote more collaborative use of the software, as evidenced through a greater degree of co-editing in the Public space. This hypothesis was refuted by the results of the study. In fact, the awareness mechanism caused no measurable differences in the way the software was used by participants. In particular, the inclusion of the awareness mechanism did not cause participants to create more modules, perform more editing, or delete more modules. This contradicts previous research which suggests that increased awareness can cause reduced participation in collaborative activities (Olson et al., 1993, Pinsonneault et al., 1999).

Despite this, analysis of the interaction logs shows that the feature was used by the participants for on average over one minute during the fifteen minutes they were presented with it, and a significant proportion of participants identified interface C2 as the most complex, suggesting that the awareness mechanism was not ignored by the participants. Turning to the interview transcripts, participants often expressed a preference for condition C2 as it allowed them to get an overview of each other's work and therefore aided in reducing redundancy or duplication of work. This, coupled with the short amount of time the feature was used for, suggests that gaining a brief glimpse of their collaborator's activities provided enough information to support co-ordination. Based on this, it could be argued that providing additional awareness information through the Views mechanism changed the *quality* of the interaction, despite having little impact on the way participants used the software, as evidenced by the interaction log analysis.

Participants often stated that it was difficult to remember what each users Private Space sounded like when switching between them. This implies a design flaw in the awareness mechanism as the tabbed window prevents users from simultaneously listening to their own private contributions and the private contributions of others. This issue was investigated further in subsequent studies.

There is also evidence that the software did not provide a sufficient level of information about the activities of others. For instance in some cases participants would make verbal enquiries as to who was doing what, even when the Views awareness mechanism was present within the interface. This following excerpt, taken from the dialog during the interaction, suggests that using the software alone, the participants were not able to ascertain who was responsible for which actions or contributions. Study Three (Chapter 7) presents an awareness mechanism which enables participants to discover the authorship of musical contributions.

PARTICIPANT 1: *Glances at participant to right.* Who is playing with the pulse
PARTICIPANT 2: *Removes headphones.* Say what?
PARTICIPANT 3: Me
PARTICIPANT 1: You, Ok
PARTICIPANT 2: *laughs*
PARTICIPANT 3: alright
PARTICIPANT 1: it's, it's fine, I just wanted to see what it's doing. Keep playing.
(Session 1, Condition C2)

There were two other awareness problems, related to the disconnection between graphical interface and sonic output. The first was participants describing difficulties knowing which modules were producing which sounds, and secondly, in knowing which person was responsible for creating which modules. These issues are explored further in the subsequent studies (Chapters 6 and 7).

### 5.11.3 Awareness

*Referential Identity*

Participants rarely moved around the table or leaned over to view each other's screens, however pointing with fingers and talking was noted when participants assisted each other in the use of the software interface. While in conversation participants would occasionally make use of the shared software interface to gesture at specific music modules, for instance by continually moving a slider or music module window to indicate to others what they were referring to. The following excerpt demonstrates an instance of participants establishing reference to aspects of the interface through pointing and talking:

PARTICIPANT 2: *it's really difficult to single out what's doing what*
PARTICIPANT 1: *exactly, yeah*
PARTICIPANT 3: *yeah*
PARTICIPANT 1: *there's someone somewhere that's, there's one of them there that's making the*
PARTICIPANT 1: *how about this one down here? (repetedly clicks)*
PARTICIPANT 2: *erm*
PARTICIPANT 3: *no, I didn't*
PARTICIPANT 2: *take that texture*
PARTICIPANT 3: *no, I, I put the volume down, and it's*
PARTICIPANT 2: *is it*
PARTICIPANT 3: *very strange*
PARTICIPANT 3: *can't find it*

There are two points to draw from this extract. Firstly, Participant 2 expresses the problem of identifying the music module responsible for a certain sound within the musical mix, stating it is difficult to single out what is doing what. Participant 2 is attempting to gather awareness information about the source of a sound, and in doing so draws attention to the way graphical interfaces for music detach the means of visually defining a musical event from the auditory result; the two operate in different modalities.

The second point about this incident relates to the way Participant 1 draws the attention of Participant 2 to a specific item within the interface. During the incident participant 1 uses the spatially consistent layout of the workspace across all users screens and the knowledge that his on-screen actions are immediately reflected to all screens to make an indexical reference a music module by repeatedly modifying a parameter, and referring to the module as 'down here'. It is important to note that Participant 2 is not at this stage trying to ascertain who created the music module he is searching for, this highlights a situation which the Workspace Awareness framework (Gutwin and Greenberg, 2002) cannot fully describe, as it does not take into account awareness of autonomous elements within a shared workspace by non-human entities. It would therefore be feasible to suggest adding an additional authorship dimension to the WA framework to account for an awareness of which non-human element is responsible for which sounds within the unfolding music. This concept is explored further in Chapter 7.

### 5.11.4 Approaches

Through discussion, some groups of participants introduced more structured management into the collaboration, by for instance assuming specific roles, such as working on bass-lines, or creating rhythms. While in some instances a group member took the initiative and assigned roles to others, participants often noted in the group discussion that roles emerged spontaneously during the course of the interaction. Participants reflected on the assumption of roles based on what their fellow collaborators appeared to be doing, for instance noting that they would try to avoid duplicating other people's work by choosing to work on different sounds or instruments. During the interviews, participants often noted that this approach helped them remember who was responsible for which aspects of the music. In cases where participants had access to a Private space, a parallel approach was observed in which participants would designate time for each member to develop ideas in private, and then in turn share their ideas with the group for scrutiny, selection and further revision. This observation is similar to the finding of Morris et al. (2004), who described participants adopting a 'divide and conquer' approach to solving a collaborative task in instances where the interface supported individual work.

### 5.11.5 Feature Suggestions and Usability Issues

The study highlighted a number of usability issues encountered by participants while using the collaborative music software.

*Transfer of Modules*

It was frequently noted that transferring modules between the Private space and the Public space could have been simplified by means of drag-and-drop, rather than via a button on the module window.

*Musical Limitations*

Some of participants discussed a number of musical limitations within the software. One common limitation was the size of the musical loop, which participants often noted to be too short. Some suggested the need for multiple compositional sections to move between, such as verse and chorus, while others stated that a timeline feature would be a useful inclusion. Other participants noted a tempo control should be included. Another frequently raised issue was the limitations of the algorithmic music modules, and participants suggested that allowing control over note-level parameters would open up a lot of creative possibilities.

*Muting Private and Public Spaces*

As a feature suggestion, participants often stated it would be useful to solo the individual spaces, so as to avoid the problem of constantly hearing the sum of both spaces simultaneously. Such a feature would have enabled participants to distinguish between shared and private musical contributions. Study Three (Chapter 7) explores the notion of allowing participants to solo specific aspects of the shared audio.

### 5.11.6 Musical Expression

Most groups expressed frustration at the lack of expressive potential, and the limited scope for musical novelty with the software may have impacted on the participant's commitment to the task. The algorithmic nature of the music modules is too simplistic for a 'real' application, although the issues raised in this study demonstrate that it was sufficient as a research vehicle. The issue of musical complexity is addressed in subsequent studies through the inclusion of more advanced music making capabilities and sound manipulation options within the software.

### 5.11.7 Video Task

As described in 5.5, participants were set the activity of composing music for an abstract video loop. While some participants stated not paying a great amount of attention to this task, others stated in discussions that they often oriented around the video loop and found it a useful means of initiating conversation with their fellow collaborators. When asked, some participants discussed taking musical inspiration from the video loop. The 'video task' is used in the following two studies.

## 5.12 Summary

This concludes the first study of this thesis. Nine groups of three musicians used three software interfaces to make music together using an interface shared across a network of computers. Each interface variant provided users with different degrees of privacy and awareness. Changes in the interface design caused significant differences to the way the participants used the software and the way the groups worked together. Specifically, when given the opportunity, participants made extensive use of the ability to work individually and control access to, and release of their musical contributions. When made available, the ability to work privately was exploited by participants from all groups to formulate ideas in isolation before making them available for others in the group to hear and edit. The awareness mechanism changed the subjective quality of the interaction for a number of participants, but from the viewpoint of the interaction logs, it did not alter the way in which the software was used.

Analysis of the results pointed towards a number of areas for further exploration, specifically the means by which audio can be presented so as to indicate its status as shared or private, and the role of awareness mechanisms in aiding users to identify the ownership of musical contributions. These issues are addressed by the studies presented in following two chapters.

# Chapter 6

# Study Two: Audio Delivery

---

*The second study of this thesis investigated the effects of different forms of audio delivery for Private and Public audio. Results suggest that presenting Personal audio through speakers discourages participants from using their Personal audio channel. Analysis of the way participants arranged the shared workspace also points to different styles of interaction. Portions of this work are presented in Fencott and Bryan-Kinns (2012).*

The previous chapter (Chapter 5) presented a study investigating the way groups of musicians used Public and Private workspaces in a collaborative graphical environment for creating music. The study showed that given the opportunity, musicians used their Private workspace to formulate musical contributions in isolation from the group before sharing their ideas in the Public workspace. Incorporating a Private workspace into the interface also affected factors such as the overall number of musical contributions and amount of editing, as well as the proportion of musical contributions which were editing by multiple group members.

Despite the advantages of including a Private workspace within the collaborative interface, participants in Study One encountered the problem of not knowing which sounds were currently audible to all group members via the public workspace, and which sounds were only audible to themselves via their Private workspace. This ambiguity may be seen as a product of the auditory nature of music (see Section 3.2), and a consequence of the modality separation between the visual control mechanisms of the graphical user interface and the resulting auditory output.

In Study One audio was presented using headphones (see Section 5.8), therefore allowing each group member to receive a personalised audio mix containing sounds which were not necessarily present in the audio directed to other group members' headphones. This in turn allowed group member to formulate musical contributions in auditory isolation from their collaborators. Using headphones may have also influenced the group interaction in other ways, such as disrupting or impeding verbal conversation, and causing participants to become more focused on the computer interface than the people around them. Acoustic instruments feature an intrinsic physical connection between the means of control and the sound source (Jordà, 2005a), whereas in computer based instruments in which audio is generated electronically the sound can be presented in a variety of ways. Speakers and headphones are ubiquitous devices for this purpose, and are therefore important technological artefacts for the production of computer based music.

## 6.1 Audio Presentation

In Study One (Chapter 5), audio was presented to the participants exclusively through headphones. There are many other ways audio could have been presented to members of the group. One option, which was suggested by participants from several groups in Study One, would be to present public audio via shared speakers, while providing each group member with individual headphones for their Private audio. This configuration is also similar to the arrangement often employed by DJs, where headphones are used to pre-listen to a record before it is made audible to the audience through speakers (Pfadenhauer, 2009). As discussed in Chapter 2, there is limited research on the effects of audio delivery within HCI and CSCW.

To explore this issue, the study presented in this chapter examines how sound can be presented to a group of musicians so as to afford an understanding of the location and status of musical elements within a collaboratively created musical work, and provide information to individuals about the activities of the group members. The study compares the effects different devices for audio delivery (speakers, headphones) have on the process of collaborative music making using a shared software interface distributed across multiple computers for use in a real-time co-present musical interaction. Three combinations of speakers and headphones, referred to as the audio presentation modes, are used to investigate these effects. It is important to stress that the focus of this chapter lies in a comparison of the affordances presented by personal and public means of audio delivery, rather than simply a shallow comparison of speakers and headphone devices.

The research question this study addresses is *how does the mode of audio presentation affect co-located collaborative digital musical interaction using an interface shared across multiple computers?* More specifically, how does the way audio is presented through combinations of speakers and headphones affect:

- the way musicians gather or maintain awareness information during musical interaction
- the way groups arrange and manage the shared interface
- how musical material is contributed, edited or co-edited

Within the field of collaborative music making there are no detailed comparisons of headphones, speakers and combinations thereof for co-located group musical interaction. This study therefore informs the design of collaborative computer music interfaces by pointing to the implications of choices made about the way audio is physically presented to users, and the presentation of shared and private musical material. This study also contributes to an understanding of how groups of musicians collaborate in real-time, considering the provision of awareness and authorship information, and support for the user defined spatial arrangement of resources in distributed interfaces. More generally, as audio and sonic interaction design has not been a central feature for CSCW research the findings of this study contribute to an understanding of how different devices for presenting audio can influence collaborative group work in non-musical activities.

## 6.2 Collaborative Music Software

The Study One software (described in section 5.1) imposed a number of constraints on the musical outcomes participants could realise. The algorithmic nature of the music modules predetermined the types of melodies and beats which could be created, and these constraints, combined with the limited sound pallet fixed tempo may have caused participants to create music which they may not otherwise have wished to create. The algorithmic nature of the Study One music modules also caused the music to sound relatively harmonious and pleasant, regardless of what combinations or parameters were used by participants. This is problematic, as if all musical contributions sound relatively homogeneous there may be less opportunity for participants to discuss or debate the quality of each others contributions, and less incentive to perform editing.

The software musical interface developed for this study is a multi-user patching environment designed to promote collaborative activity between the participants. The software allows a number of software instruments to be routed through audio processing effects before being connected to an audio output which is associated with either headphones or speakers. The connection metaphor used in the software is similar to other commercial music environments such as Propellerhead's Reason software (Duignan et al., 2004). Although the design is visually similar to patching environments such as Max/MSP and Pure Data, the experiment software does not offer the low-level programming capabilities of such systems.

The new software design attempts to address some of the creative limitations imposed in Study One by replacing the algorithmic modules of the first study software with sequencer style instrument modules. These modules provide the users with control over the musical contents of a 16 step looping bar, and therefore offer more fine grain control over the musical material being created. They also feature a much wider variety of timbral controls, while the effects modules can be used to achieve more advanced and unpredictable sonic manipulations. The instruments available are the Drum Sequencer and Step Synth. The effects are Distortion, Delay, Reverb, and Filter. Compared to the Study One software, including more sounds and effects gives participants more opportunities to customise the style and character of their contributions, while providing note-level editing options (rather than higher level algorithmic controls) allows participants to create their own melodies and drum patterns, and in addition opens up the opportunity for participants to create discordant, out of time, and unpleasant sounding contributions. This in turn may provide more opportunity for discussion, disagreement and decision-making within the collaboration. With reference to game design theory, Overholt (2009) supports this concept by stating that 'good instruments should be able to make bad sounds', so as to present a challenge and provide an element of reward to their users.

Furthermore, the effects routing options provide more opportunity for users to interconnect and inter-associate their musical contributions in ways which were not possible using the Study One software.

### 6.2.1 Shared Workspace

While the software in Study One enforced a separation between Public and Private areas by means of a split-screen design, the Study Two software uses a single shared workspace which is consistently distributed across all computers (see Figure 6.3). This move to a single shared workspace, rather than a split-screen design is significant as it provides more opportunities to study how participants develop or enforce territorial behaviour and use spatial partitioning of the interface to support their collaboration.

### 6.2.2 Shared and Personal Audio

In order to allow participants to work in privacy the software introduces Public and Personal audio outputs. Outputs are represented as grey rectangles within the workspace. Instruments and effects can be 'patched' to these outputs. All users share a Public output which is found in the centre of the workspace. Figure 6.1 shows modules patched to the public output.

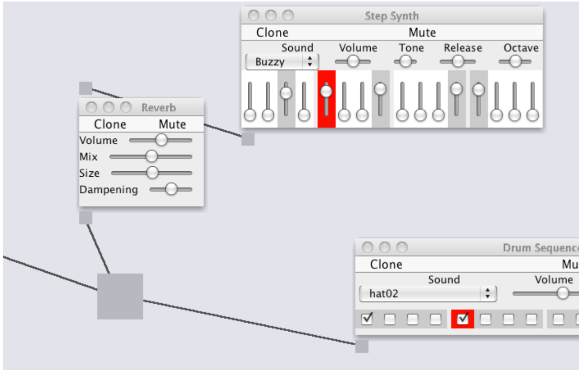



Figure 6.1: Public Audio Output

All users are also provided with a Personal output which is located in the bottom left of the screen. Figure 6.2 shows the Personal audio outputs. Modules connected to a users' Personal output appear as greyed out rectangles on other users' screens.

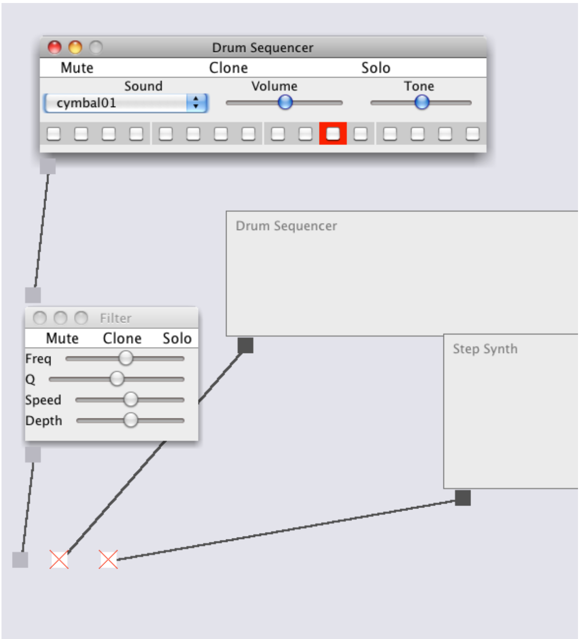



Figure 6.2: Personal Audio Outputs shown in bottom left of figure

Audio from modules patched to the Public Output is routed to all participants. Depending on the experimental condition this may mean routing audio to all participants headphones, or to all participants speakers, participants can route audio to their headphones or individual speaker by connecting modules to their own Personal Output. Participants are able to mute their Personal audio, and are also able to listen to each-other's Personal audio by un-muting the other users' Personal Outputs. The output a module is connected to also affects its visual representation and editing access for other users. Modules patched to the Public output are freely editable by all users, while modules patched to a Personal output are only editable by the owner of that output, and appear as grey boxes on all other participants screens.

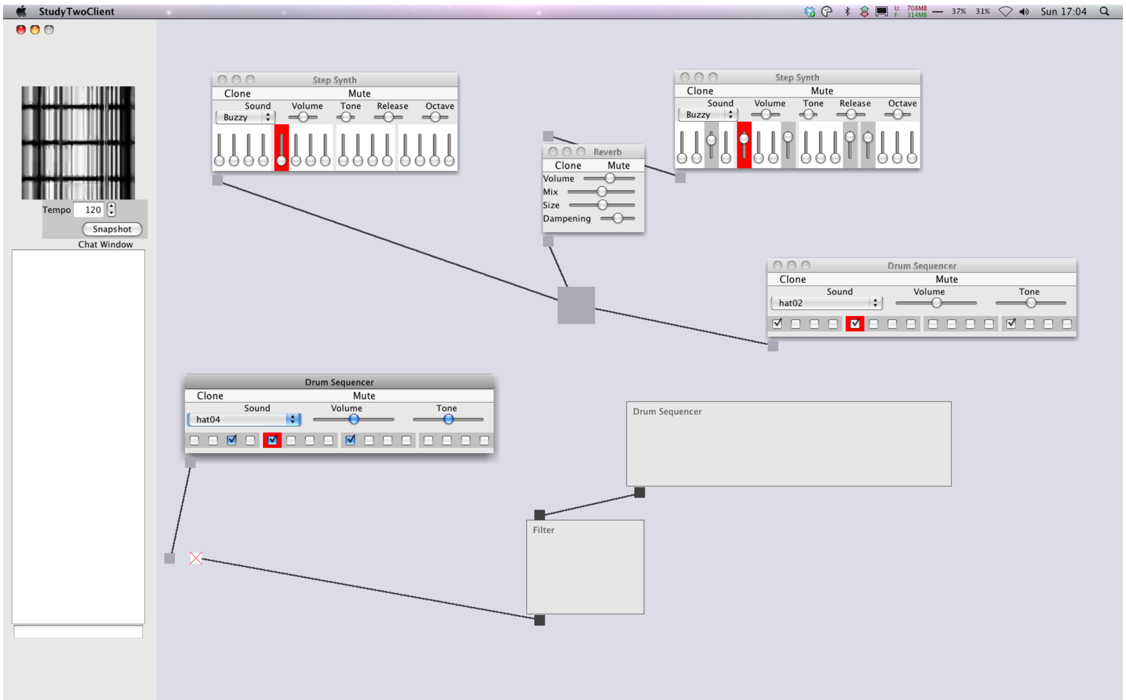



Figure 6.3: Interface for Study Two.

### 6.2.3 Global Controls

As well as a requirement for more control over musical features such as melody and rhythm, in Study One participants frequently mentioned the need for global controls, such as the tempo or musical key. The software used in Study One offered no such controls, and therefore a tempo control was introduced into the revised software environment for this study. A global tempo control was also proposed as an area of future investigation by Laney et al. (2010). The tempo operates globally for all users, and as it's effect have implications for every user, the points at which it is discussed and manipulated are a potentially interesting feature to study.

## 6.3 Hypotheses

The research question from Section 6.1 were used to derive the following hypotheses:

- H1 Presenting audio exclusively through headphones will encourage more individual work. Indicators of individual work are: more use of Personal output; less co-editing of contributions; less verbal or embodied acknowledgements.
- H2 An interface with the most auditory distinction between Public and Private audio channels will be preferred by participants. This hypothesis is based on the assumption that privacy is a key factor in the user's preference (Fencott and Bryan-Kinns, 2010).
- H3 An interface which presents Personal and public audio entirely through speakers will cause participants to work more collectively, as evidenced by: more audio played in the Public channel; less use of privacy features; observation of spatially indexical cues to audio sources being used as a collaborative resource

## 6.4 Experiment Tasks

### 6.4.1 Video Task

Participants were set the challenge of composing music for a short video animation. Due to a chance in the way the software's graphical interface was implemented, the Processing library[1] used to generate the video animations in the Study One software (see Section 5.5 became unstable in the new software. For this reason, the videos used in this study were different to those used in Study One. Rather than using the animated geometric shapes from Study One, the video animations for this study were taken from output of the interactive VJ software developed by Bergstrom and Lotto (2009). These animations were played in the software using the Java Media Framework[2], which proved to be more reliable. The new animations were abstract compositions of line and colour, similar in style to the videos presented in the first study, although featuring more visual complexity. As with the first study, the video task was in place to spark discussion and provide a talking point for the participants.[3]

[1] http://processing.org/ Accessed 30 September 2012

[2] http://www.oracle.com/technetwork/java/javase/tech/index-jsp-140239.html Accessed 30 September 2012

[3] videos downloadable from http://www.eecs.qmul.ac.uk/~robinfencott/thesis_resources/

### 6.4.2 Snapshot Task

In addition to the video task, the participants were asked to make recordings or 'snapshots' of their music composition during each experimental condition. Snapshots are 30 second snippets of music, which are recorded using a button on each of the participant's GUIs. A snapshot begins when all participants vote to start, by each pressing their Snapshot button. The motivation for inviting participants to create snapshots of their music is to encourage discussion amongst the participants, for instance about when to take the snapshot, and which musical material should be included in it.

## 6.5 Independent variables

The study employed a three condition within-subjects design in which all participants were exposed to each of the conditions. The musical software (discussed in Section 6.2) remained constant throughout all conditions, and the independent variable was the way audio is presented. This was manipulated by means of three configurations of speakers and headphones, presented as three different experimental conditions. Ordering was randomised between groups to counteract learning effects. The experimental conditions are described below, and outlined in table 6.1.

| **Condition** | **Public Audio** | **Personal Audio** |
|---|---|---|
| C1 | Shared Speakers | Personal Speaker |
| C2 | Shared Speakers | Headphones |
| C3 | Headphones | Headphones |

Table 6.1: Table showing audio delivery configurations for experimental conditions

*Condition C1: Speakers only*

In this condition each participant has his/her own speaker, which is used to present both Personal audio and Public audio. Using Personal speakers would be similar to conventional instrumental playing, where each person's instrument comes from a distinct spatial location. Giving participants individual speakers creates a situation in which participants each have a personalised version of the music which can be overheard by others in the workspace. Research from Auditory Scene Analysis (Bregman, 1990), suggests that participants will be able to localise the sources of different sounds from multiple loudspeakers. The spatial quality of audio played through speakers may provide a form of feedthrough awareness, allowing participants to identify more easily who is doing what in the collaboration. Using speakers instead of headphones may also encourage more verbal communication.

*Condition C2: Headphones and Public Speakers*

In this condition each participant has their own speaker, which is used to present the Public audio. Personal audio for each participant is played through headphones. Playing public audio through shared speakers and providing non-shared audio through individual headphones is not uncommon in conventional musical contexts, for example a common feature of DJ practice is to use headphones to cue new records in private before they are mixed with the front of house speakers for the audience to hear (Pfadenhauer, 2009). In this instance by wearing headphones a participant may indicate to others (as a form of consequential communication) that they are concentrating on private contributions.

*Condition C3: Headphones Only*

In the third condition participants hear both Public and Personal audio through headphones. Musical contributions routed to the Public Channel go to all headphones, while contributions routed to the Personal channel are only routed to the owner's headphones.

When listening to audio through headphones other co-located individuals are not able to listen in. Headphones may also inhibit natural conversation, and may make the wearer less conscious of other activities around them. In Study One it was noted that participants would sometimes remove their headphones to indicate availability for conversation or to initiate group discussions (see 5.11.3).

## 6.6 Dependant variables

### 6.6.1 Questionnaire data

Pre and Post-test questionnaires were used. The pre-test questionnaire collected demographic information about the participants, while the post-test was designed to collect information about the participants experiences with the experimental conditions. The post-test questionnaire required participants to order the experimental conditions in terms of how they applied to a list of statements. This questionnaire was based on the Mutual Engagement Questionnaire (Bryan-Kinns and Hamilton, 2009). A computer based questionnaire system was used to administer the questionnaires and automatically log the results. The pre-test questionnaire questions are included in Appendix C. The post-test questionnaire questions were as follows:

- The best music
- I felt most involved with the group
- I enjoyed myself the most
- I felt out of control
- I understood what was going on
- I mostly worked on my own
- I lost track of time
- Other people ignored my contributions
- We worked most effectively
- The interface was most complex
- I had the most privacy
- I knew what other people were doing
- We edited the music together
- I made my best contributions
- I was influenced by the other people
- The condition I preferred the most

### 6.6.2 Interaction Log Analysis

The software logged all interaction, including mouse clicks, key presses, and details of which software functions and operations were used, as well as how much time was spent using each function. These measures are similar to those taken in Study One (see Section 5.4.2). Various software tools were developed using Python and Java to extract quantitative data from the logs. Table 7.2 summarises the features studied, and how they were identified in the log files.

### 6.6.3 Group Discussions

Semi-structured discussions were held with the participants at the end of the session to gather data about their perceptions and experiences. Discussions were video and audio recorded. A set of questions were generated to focus on general preferences, perceived differences between the speakers and headphones, use of the Personal and Public channels, awareness of each-other's activities, roles and working strategies and spatial use of the shared on-screen workspace. Participants were also invited to raise any other issues, talk about the experience in their own words, and ask their own questions. Analysis of interview transcripts used Grounded Theory methods, as

| Feature | Description |
|---|---|
| Creations | Creating a module |
| Deletions | Deleting a module |
| Cloning | Cloning a module using the clone button |
| Snapshots | Taking a group snapshot of the music |
| Text chat | Number of text chat events |
| Editing (raw) | Modifying a module control (slider, button, etc.,) |
| Patching to Public | Connecting a module to the Public output |
| Patching to Personal | Connecting a module to a Personal output |
| Movement | Movement of modules |
| Number of Snapshots | |
| Tempo changes | |

Table 6.2: Summary of log file features studied

adopted by many others for the study of digital musical interactions, CSCW and HCI (Dow et al., 2008, Laney et al., 2010, Luther and Bruckman, 2008, Michael Gurevich and Marquez-Borbon, 2010, Muller and Kogan, 2010, Thom-Santelli et al., 2009).

## 6.7 Participants and Recruitment

Thirty individuals were recruited and organised into ten groups of three people. These groups were scheduled to attend experiment sessions at Queen Mary University of London. Participants received financial compensation of £10 for taking part. All participants were recruited via mailing lists related to art, electronic music, sonic arts, music technology and academic research concerning sound and music computing. The recruitment e-mail stated that a new music application was being developed at Queen Mary University of London, and that for testing purposes, the project required 'people with an interest in creating music, for instance composers, musicians, DJs, and students of Music, Music Technology or related fields'. The message stressed that people would work in a group with two other people, and that they did not need to be able to play an instrument or have formal musical training to take part. The message invited people to make contact via e-mail if interested. In line with ethical procedures for the recruitment of human subjects, no individuals were directly approached or asked to participate.

Over a period of several months approximately eighty people replied to the recruitment e-mails, either expressing an interest in taking part, or asking for additional information. Each respondent was contacted via e-mail, with answers to any questions they had about the research. Respondents were then invited via e-mail to take part in an experiment session. Electronic copies of the information sheet and consent form were sent to participants prior to the experiment session, and e-mail reminders were sent close to the date of the session. Participants were informed they did not need to bring instruments or equipment with them on the day of the session.

Demographic information for each participant was acquired via a computer administered pre-test questionnaire at the start of the experiment session. Participants were able to skip questions if they so wished. 25 of the 30 participants completed the questionnaire in full, while five participants omitted details about age and/or occupation. 19 participants (68%) were male (based on 29 responses). The mean age of participants (based on 25 responses) was 33 years. 20 participants (69%) could play a musical instrument. 6 participants ( 22%) classified themselves as beginner level in their musical proficiency, 11 participants (40%) classified themselves as of 'intermediate' level, 5 participants (18%) classified themselves as 'semi-professional' and 5 participants (18%) classified themselves as 'professional' level. 25 participants (86%) had previously composed songs on their own. 4 participants (16%) had composed 2-5 pieces, and 21 participants (84%) had composed more than 10 pieces. 22 participants (75%) had previously composed songs with others. 2 participants (9%) had composed 1 piece with others. 4 participants (18%) had composed 2-5 pieces with others, 3 participants (14%) had composed 5-10 pieces with others, and 13 participants (59%) had composed more than 10 pieces with others.

One participant (3%) described their level of computer literacy as 'beginner'. 17 participants (58%) identified their level of computer literacy as 'intermediate', and 11 participants (37%) identified as 'expert'. 28 participants (97%) had previously used an Instant Messenger application to communicate online, and one participant had not used an Instant Messenger (3%). 8 participants (28%) had not used collaborative software before. 7 participants (25%) had played online multi-player computer games, 4 participants (14%) had used collaborative document editors, 2 participants (7%) had used collaborative writing software, 14 participants (50%) had previously used some form of collaborative music making software, one participant (4%) had used another form of collaborative software (SVN)[4].

Information about musical preference was collected by asking participants to select from a multiple choice list of sixteen musical genres. The most frequently selected genre of music was 'Jazz' (22 participants) closely followed by 'Electronic' (21 participants). The least popular genres were 'Metal' (7 participants), 'Country' (8 participants) and 'Rap' (9 participants). The mean number of genres selected was 7.37, indicating that most participants appreciated a

[4]Percentages do not add up to 100 for multiple choice questions

range of different musical styles. The Spearman rank-order correlation coefficient indicated no correlation between age and musical tolerance (n=24, rs=-0.0309, t=-0.15, df=2, $p_1$=0.441066 $p_2$=0.882131). In summary, whilst not all participants were instrumentalists, there was a high level of musicality and musical experience within the participant sample, as well as a high level of computer literacy.

It might be expected that people interested in testing new computer music software would appreciate electronic music, yet it is also interesting that jazz was such a popular genre within the participants for this study. One possibilities is that the term 'jazz' is too broad and generic to reflect any specific style, while another explanation is that people willing to work with other musicians in an experimental musical context may also be inclined to enjoy and/or participate in music such as jazz, which frequently incorporates aspects of improvisation. Although data on musical preferences was collected by the pre-test questionnaire, this information was not factored into the analysis.

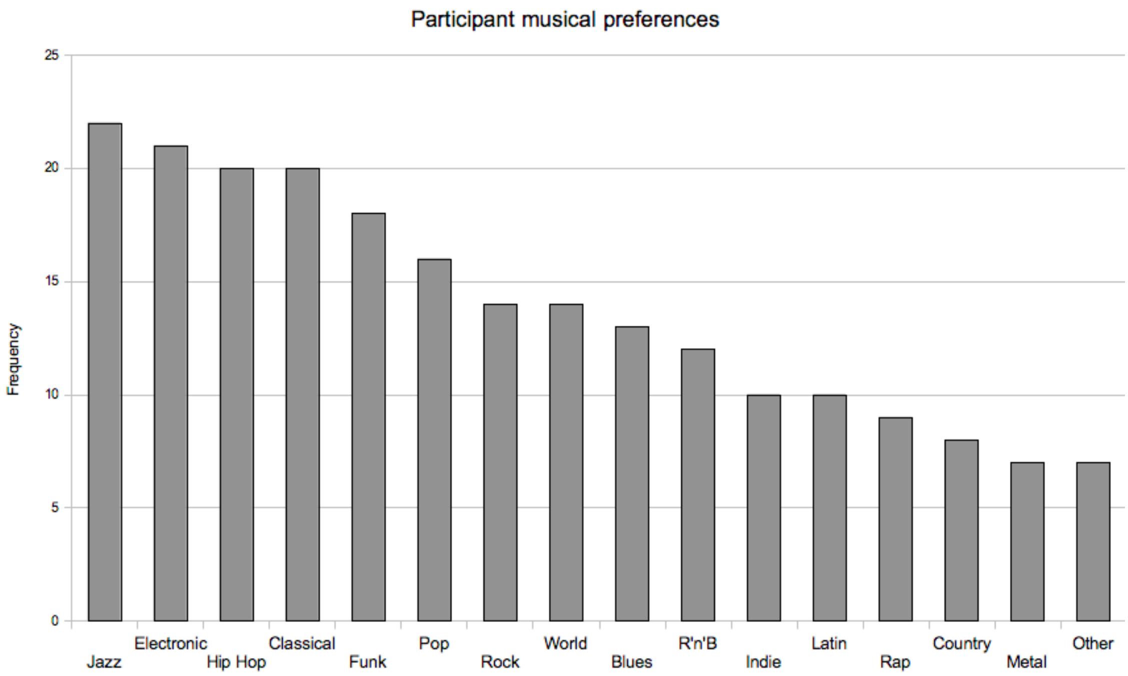



Figure 6.4: Participant Musical Genre Preferences

## 6.8 Pilot Studies

Two pilot studies were conducted prior to commencing with the full study. The first pilot identified a number of usability and stability issues within the software, while analysis of the captured pilot data highlighted problems with the data collection mechanisms. These issues were addressed before the second pilot.

To aid in analysis, the data logging methods were modified to record the state of the entire system every time a change was made. This simplified the analysis process at the expense of increasing the size of the log files.

## 6.9 Procedure and Apparatus

Three studio quality near-field Yamaha MSP5 monitors were used for conditions requiring speakers. Sony MDR-7509HD headphones were selected for conditions C2 and C3. These headphones feature freely rotatable ear-pieces and therefore allow the wearer to comfortably hold them to their ear in the style of a DJ, as well as wear them in the conventional way. Mac-Mini computers with 21 inch widescreen displays were used. The computer displays were lowered as far as possible to allow participants to see over them. Each computer had an identical two button mouse and Apple keyboard. The workstations were set up on a large round table (see Figure 6.5) with the Mac Mini units in the centre. A Yamaha MSP5 monitor was positioned to the right of each computer screen, and a pair of Sony MDR-7509HD was headphones placed on top of the speaker at the start of each session. A chair was positioned at each of the computer workstations.

Figure 6.5: Second study set-up.

Each session was structured as follows:

- Introduction, Consent form signing, pre-test questionnaire - 10 minutes
- Training - 10 minutes
- Experimental Conditions - 45 minutes
- Post-test Questionnaire - 5 minutes

- Discussion and Debriefing - 15 minutes

Participants were asked to meet in a reception area close to the experiment laboratory. Once all participants were present they were led to the experiment laboratory, and asked to take a seat around the computer table. The researcher then sat at a chair near the table and presented participants with information sheets and consent forms. The researcher verbally described the purpose of the study session, stressing the voluntary nature of the research participation. In line with ethics guidelines, the data collection procedures were explained, along with reassurance that all data would be anonymised, and that images, video and audio from the research would not be published or made available online.

Participants were encouraged to ask questions throughout the introductory explanation. They were then given time to read the information sheet and consent form, and ask any further questions before being invited to sign the consent form. After signing consent forms the participants were presented with a computer administered pre-test questionnaire to collect demographic information and details about their previous musical background, computer literacy and experience using collaborative computer software.

Participants were then given a training session with the software, which covered how to create, edit and delete music modules, how to use the chat tool, the use of the Personal and Private workspaces, and the way modules could be transferred between these spaces. The location of the video loop within the interface was made clear to the participants. During training the participants were given a short amount of free time to use the software. Throughout the training participants were invited to ask questions about the software. After training the groups were given fifteen minutes with each software condition. To control for learning effects the order of conditions and video loops was randomly assigned within each group.

After exactly fifteen minutes the Server application automatically closed the interface and presented participants with a blank screen. At this point the researcher loaded the next experimental condition via a software interface running in the control room. The researcher then explained to the participants how the new conditions was different from the previous, and answered any questions the participants had, before letting the participants begin.

After the three conditions had been presented the participants were presented with a computer administered post-test questionnaire, as described in Section 6.6.1. After completing the final questionnaire a group discussion was conducted by the researcher. Towards the end of this discussion the participants were invited to ask the researcher questions about the study. At this stage the researcher explained more generally about the research being conducted and the purpose of the study. After the group discussion the participants filled out a form to acknowledge receipt of the £10 financial compensation for their time. Finally the participants were thanked for their time and led by the researcher to the exit.

## 6.10 Sample Size

As this study was based on the same design as Study One, a similar sampling approach was taken, although a larger sample of 30 participants was recruited. The same post hoc analysis was performed on the Friedman Test results to determine their effect size. This process is described in Section 5.9 and Appendix A.

## 6.11 Results

This chapter presents the results obtained from analysis of the experiment sessions. A discussion of the findings proceeds in Section 6.12.

### 6.11.1 Post-Test Questionnaires

All participants were invited to complete a multiple choice post-test questionnaire administered via an onscreen interface. This questionnaire was in place to investigate the participants' interpretations of the experimental conditions, their feelings towards music created and their thoughts about how well the group worked together (see Figure 6.4). Of the 30 participants, 28 completed the questionnaire in full, one participant provided responses to almost all statements and one participant did not provide any responses.

The post-test questionnaire asked participants to rank a series of statements in order of their applicability to the experimental conditions presented to them. As experimental conditions were randomly ordered for each session the participants were asked to use the order in which the conditions were presented to them. The response orders were subsequently placed in a uniform order for analysis. The Friedman test was used to identify questionnaire statements which elicited a statistically significant trend. A significant number of participants identified condition C2 as the one in which they 'lost track of time' ($p=0.0406$, $df=2$, $csq_r=6.41$). A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between

C1 and C2 ($p<0.001$, r=0.4553862), with a medium effect size.

A significant number of participants (p=0.0211, df=2, $csq_r$=7.72) also rated condition C2 as the one in which they 'had the most privacy'. No other statements elicited a statistically significant result or indicated a strong trend. A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 (p = 0.001, r = 0.5528049), and between conditions C1 and C2 (p = 0.01, r = 0.4100125), with a strong and medium effect size respectively.

The post-test questionnaire also included a comments field. Comments left by the participants in the post-test questionnaire are given in Appendix C, and were not analysed.

| **Statement** | **C1** | **C2** | **C3** | **Total** | $csq_r$ | df | p |
|---|---|---|---|---|---|---|---|
| The best music | 2 | 1.9 | 2.2 | 29 | 1.16 | 2 | 0.5599 |
| I felt most involved with the group | 1.9 | 1.9 | 2.2 | 29 | 1.31 | 2 | 0.5194 |
| I enjoyed myself the most | 2.1 | 1.9 | 2.1 | 29 | 0.83 | 2 | 0.6603 |
| I felt out of control | 2.2 | 1.8 | 2 | 29 | 2.14 | 2 | 0.343 |
| I understood what was going on | 1.9 | 2 | 2.1 | 29 | 0.48 | 2 | 0.7866 |
| I worked mostly on my own | 2 | 1.9 | 2.1 | 29 | 0.9 | 2 | 0.6376 |
| **I lost track of time** | 2.2 | 1.6 | 2.1 | 29 | 6.41 | 2 | **0.0406** |
| Other people ignored my contributions | 1.9 | 2 | 2.1 | 29 | 1.1 | 2 | 0.5769 |
| We worked most effectively | 2.1 | 2 | 2 | 29 | 0.21 | 2 | 0.9003 |
| The interface was most complex | 1.9 | 2.1 | 2 | 29 | 0.48 | 2 | 0.7866 |
| **I had the most privacy** | 2.4 | 1.7 | 1.9 | 29 | 7.72 | 2 | **0.0211** |
| I knew what other people were doing | 2 | 2 | 2 | 28 | 0.07 | 2 | 0.9656 |
| We edited the music together | 2 | 2.1 | 1.9 | 28 | 1.14 | 2 | 0.5655 |
| I made my best contributions | 2 | 2 | 2 | 28 | 0.07 | 2 | 0.9656 |
| I was influenced by the other people | 2 | 1.9 | 2 | 28 | 0.21 | 2 | 0.9003 |
| The condition I preferred the most | 2.1 | 1.9 | 2 | 28 | 0.64 | 2 | 0.7261 |

Table 6.3: Summary of Post-Test questionnaire results for Study Two, showing mean rank for each condition, ordered by experimental condition. Questions with a significance level of $p<0.05$ are highlighted in bold.

*Ordering Effects in Post-Test Questionnaire*

The post-test questionnaire responses were analysed in their presentation order to show the effects of time and ordering on participants experiences. A significant proportion of participants noted mostly working on their own in the first condition they were presented with (p=0.0063, df=2, $csq_r$=10.14). In the first presented condition a significant number of participants stated that their contributions were ignored the most in condition C3 (p=0.0052, df=2, $csq_r$=10.5). A significant number of participants reported that they they least understood what was going on (p=0.0046, df=2, $csq_r$=10.76) and a significant number stated that they worked most effectively in condition C1 (p=0.0167, df=2, $csq_r$=8.19). A significant number of participants noted that they made their best contributions in the second and third conditions (p=0.0107, df=2, $csq_r$=9.07), while response to the statement ‘the best music’ shows a strong trend towards later conditions (p=0.0553, df=2, $csq_r$=5.79), as do responses to the statement ‘We edited the music together’ (p=0.0596, df=2, $csq_r$=5.64).

| **Statement** | **C1** | **C2** | **C3** | **Total** | $csq_r$ | df | p |
|---|---|---|---|---|---|---|---|
| The best music | 2.3 | 1.9 | 1.7 | 29 | 5.79 | 2 | 0.0553 |
| I felt most involved with the group | 2.1 | 2 | 1.9 | 29 | 0.93 | 2 | 0.6281 |
| I enjoyed myself the most | 2.2 | 2 | 1.8 | 29 | 1.78 | 2 | 0.4232 |
| I felt out of control | 1.8 | 2 | 2.2 | 29 | 2.97 | 2 | 0.2265 |
| **I understood what was going on** | 2.5 | 1.9 | 1.7 | 29 | 10.79 | 2 | **0.0046** |
| **I worked mostly on my own** | 1.6 | 1.9 | 2.4 | 29 | 10.14 | 2 | **0.0063** |
| I lost track of time | 2.1 | 1.8 | 2.1 | 29 | 1.86 | 2 | 0.3946 |
| **Other people ignored my contributions** | 1.6 | 1.9 | 2.5 | 29 | 10.5 | 2 | **0.0052** |
| **We worked most effectively** | 2.4 | 1.8 | 1.7 | 29 | 8.19 | 2 | **0.0167** |
| The interface was most complex | 2.2 | 1.8 | 2 | 29 | 1.72 | 2 | 0.4232 |
| I had the most privacy | 1.9 | 1.8 | 2.3 | 29 | 4.34 | 2 | 0.1142 |
| I knew what other people were doing | 1.9 | 2 | 2.1 | 28 | 0.29 | 2 | 0.865 |
| We edited the music together | 2.4 | 1.9 | 1.8 | 28 | 5.64 | 2 | 0.0596 |
| **I made my best contributions** | 2.5 | 1.8 | 1.8 | 28 | 9.07 | 2 | **0.0107** |
| I was influenced by the other people | 2.3 | 2 | 1.8 | 28 | 3.07 | 2 | 0.2155 |
| The condition I preferred the most | 2.3 | 1.9 | 1.8 | 28 | 3.5 | 2 | 0.1738 |

Table 6.4: Summary of Post-Test questionnaire results for Study Two, showing mean rank for each condition, ordered by presentation order. Significance level of $p<0.05$ highlighted in bold.

### 6.11.2 Interaction Log Analysis

The Friedman Test was used to compare interaction log features between the experimental conditions. Appendix A provides justification for selecting this statistical test. Table 6.5 summarises these results. Interaction log analysis using the Friedman Test showed that where participants used Speakers Only (Condition 1) they made significantly less use of the Personal channel to listen to musical contributions (p=0.01216, df=2, $csq_r$=8.82). See Figure 6.6. A post-hoc test using Wilcoxon sign-rank tests with Bonferroni correction showed the significant differences between C1 and C2 ($p<0.001$, r=0.629554) and between C1 and C3 (p = 0.006, r = 0.480526), indicating strong and medium effect sizes respectively.

There was no significant effect on the amount of times participants patched modules to the Public channel (p=0.1496, df=2, $csq_r$=3.8). See Figure 6.7.

Interaction log analysis using the Friedman Test showed that the experimental conditions

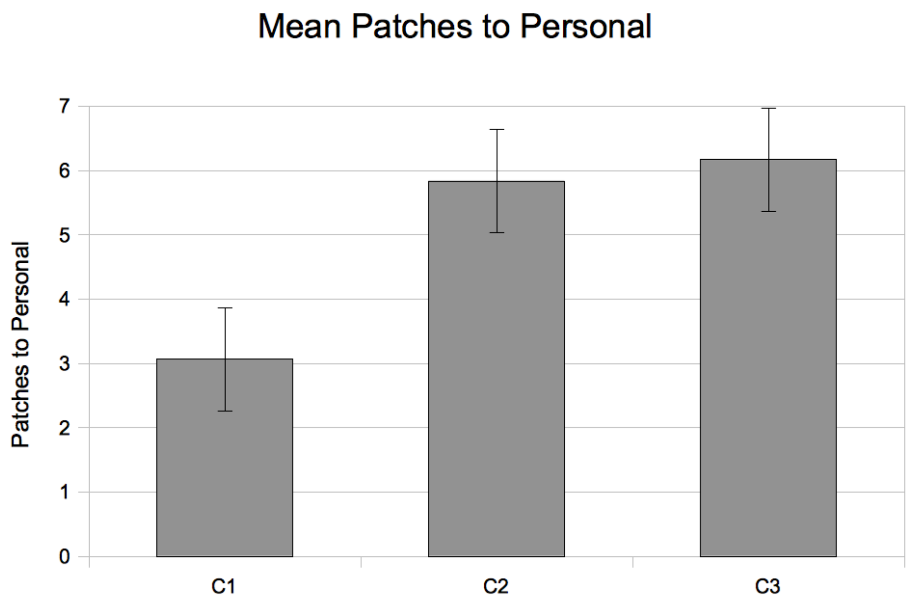



Figure 6.6: Mean number of times modules were patched to Personal outputs. Error bars show Standard Error

caused no significant difference in the amount of modules individuals contributed (p=0.7225, df=2, $csq_r$=0.65) or deleted (p=0.5169, df=2, $csq_r$=1.32). See Figures 6.8 and 6.9.

The Friedman test showed no significant differences between the amount of editing which individuals performed on their own modules (p=0.8395, df=2, $csq_r$=0.35), or the amount of co-editing which took place (p=0.3413, df=2, $csq_r$=2.15). See Figures 6.10 and 6.11.

Participants spent a similar amount of effort organising the workspace in each condition, as evidenced by the Friedman test, which indicated no significant difference in the amount of window position movements between conditions (p=0.4677, df=2, $csq_r$=1.52). See Figure 6.12.

| **Feature** | C1 | C2 | C3 | **df** | **$csq_r$** | **p** |
|---|---|---|---|---|---|---|
| Creations | 1.9 | 2 | 2.1 | 2 | 0.65 | 0.7225 |
| Deletions | 2.1 | 1.8 | 2.1 | 2 | 1.32 | 0.5169 |
| Individual editing | 1.9 | 2 | 2.1 | 2 | 0.35 | 0.8395 |
| Co-editing | 2.2 | 1.9 | 1.9 | 2 | 2.15 | 0.3413 |
| Movements | 2 | 1.8 | 2.2 | 2 | 1.52 | 0.4677 |
| Patch to Public | 2 | 2.3 | 1.8 | 2 | 3.8 | 0.1496 |
| **Patch to Personal** | 1.6 | 2.3 | 2.2 | 2 | 7.35 | **0.0253** |

Table 6.5: Interaction log results summarised to rank averages. Significance of $p<0.05$ in bold.

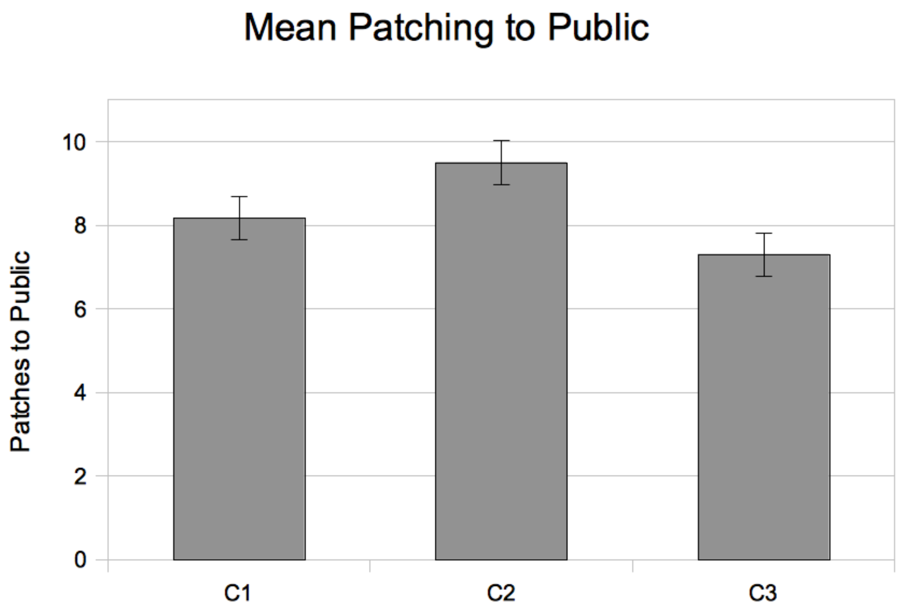



Figure 6.7: Mean number of times modules were patched to the Public output. Error bars show Standard Error

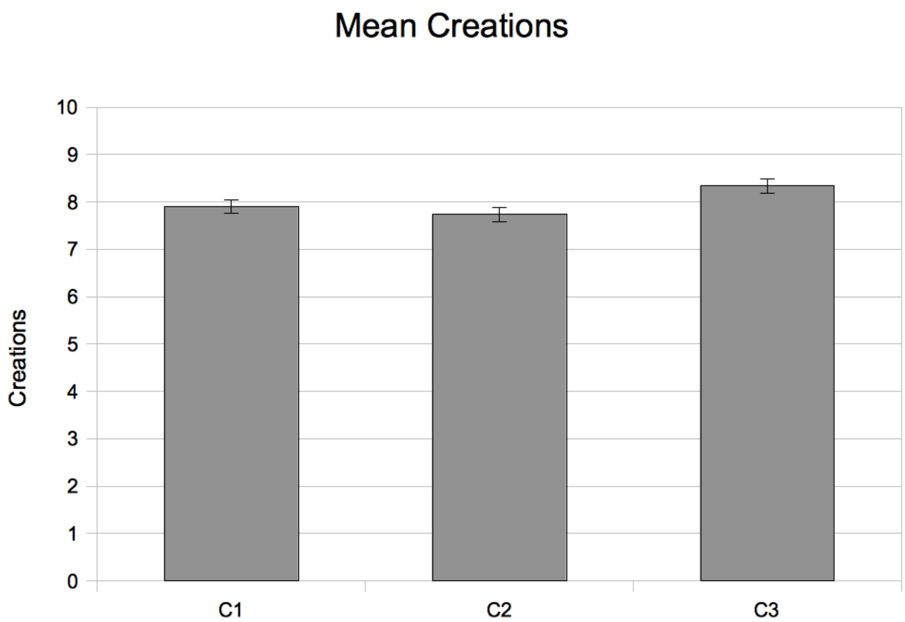



Figure 6.8: Mean number of creations. Error bars show Standard Error

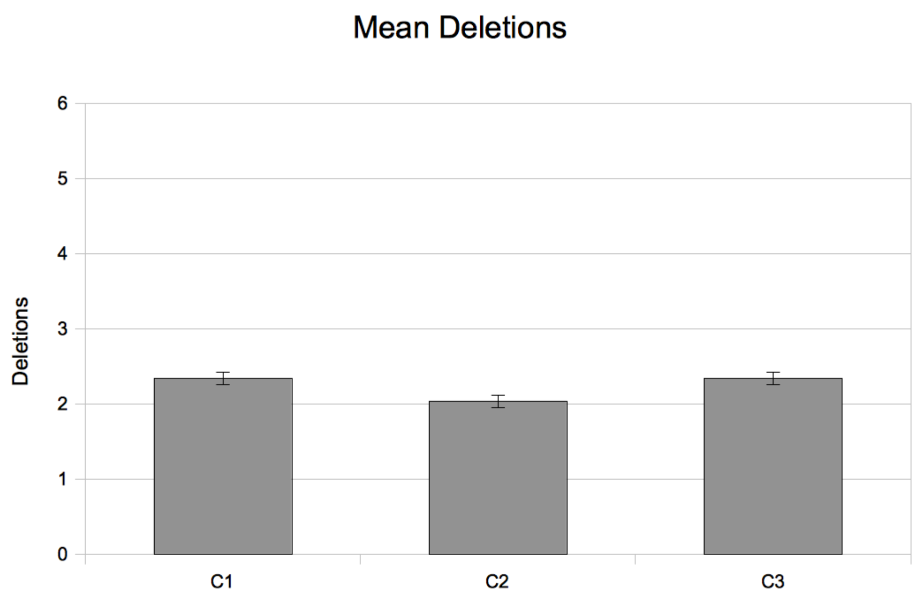



Figure 6.9: Mean number of deletions. Error bars show Standard Error

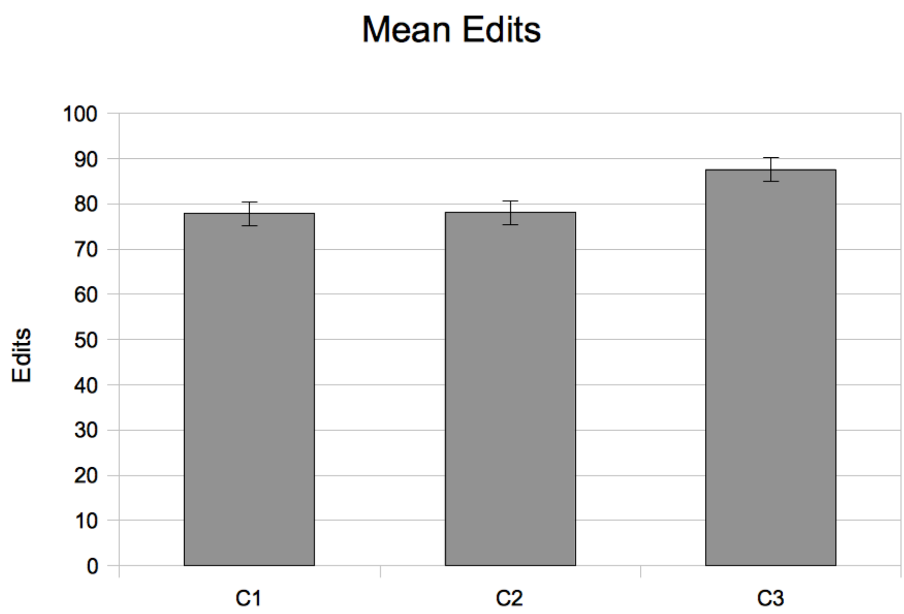



Figure 6.10: Mean number of edits. Error bars show Standard Error

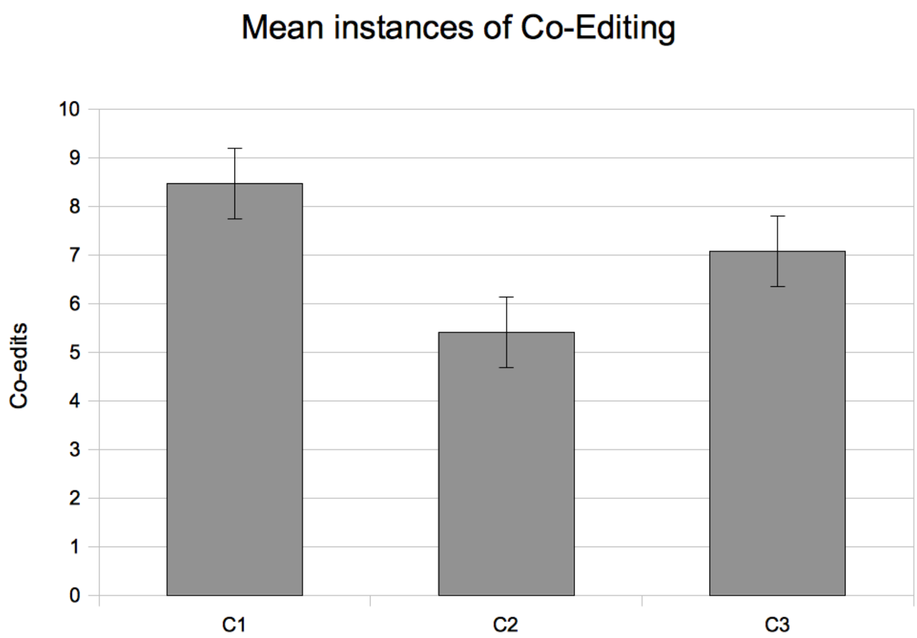



Figure 6.11: Mean number of co-edits. Error bars show Standard Error

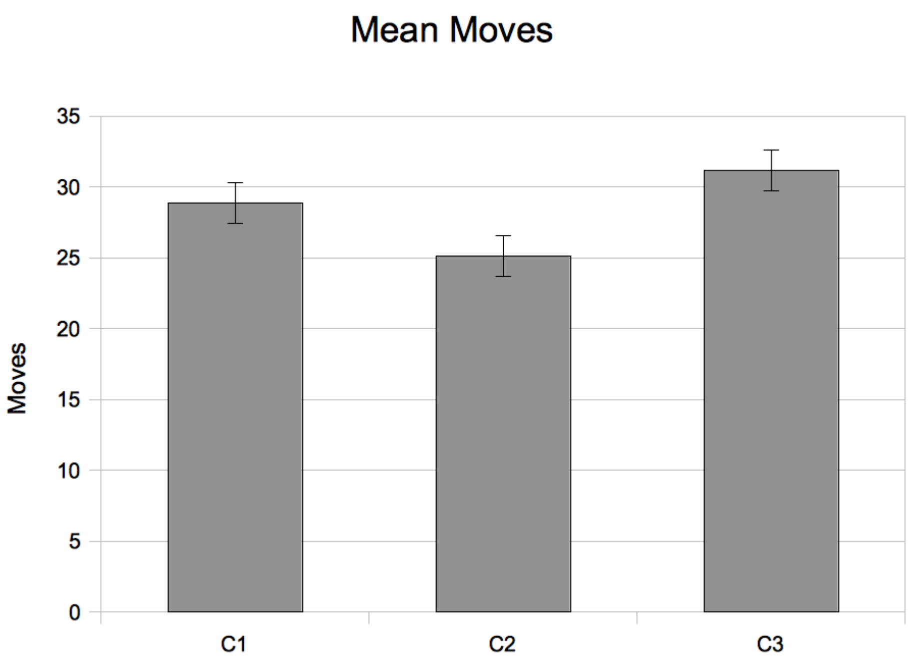



Figure 6.12: Mean number of module position moves. Error bars show Standard Error

*Snapshots*

Groups were asked to record 30 second ‘snapshots’ of their compositions using a button within the interface. In order to initiate a snapshot all participants were required to press the button. Five groups (50%) did not take any snapshots, two groups (20%) took one snapshot, one group (10%) took two snapshots, and one group took six (10%). Table 6.6 shows snapshot totals organised by experimental condition. The least snapshots were taken in condition C3 (headphones only), and the most snapshots were taken in condition C2 (headphones+speakers). Based on the limited use of the snapshot feature, no further analysis was performed.

| Session/Group | C1 | C2 | C3 | Total |
|---|---|---|---|---|
| 1 | 0 | 1 | 0 | 1 |
| 2 | 0 | 1 | 0 | 1 |
| 3 | 1 | 0 | 0 | 1 |
| 4 | 0 | 1 | 1 | 2 |
| 5 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 |
| 7 | 3 | 3 | 0 | 6 |
| 8 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 |

Table 6.6: Number of snapshots taken

*Tempo Control*

The Friedman test showed that there was no significant difference in the use of the tempo control between conditions (p=0.5916, df=2, $csq_r$=1.05).

### 6.11.3 Spatial Organisation

Visualisations of how participants spatially arranged the music modules within the workspace were created using interaction log data. The visualisations were created with the Processing language[5]. These visualisations ploted the co-ordinate position of music modules for the duration of each experiment condition. An example is presented in Figure 6.13, and the full collection is included in Appendix C. In the visualisations, each participant is associated with a colour to signify authorship, and each circle represents the position of a music module during the course of the interaction. The same module will appear twice in the visualisation if its position was

[5] www.processing.org

changed. As participants rarely edited each others' music modules (see Table 6.5) the initial creator of the module was used as the indicator of authorship. The visualisations are arranged with columns to represent experimental conditions from left to right C1, C2 and C3, and experiment groups are arranged in rows.

The visualisations were visually categorised into 'grouped' and 'intermingled' sets based on the degree to which the areas of coloured dots appeared to be grouped together. Grouped sets refer to where the participants were more territorial in the layout of the interface, by grouping there contributions together, while intermingled groups were less territorial, and interspersed their contributions across the interface. The categorisations were independently performed by a third party rater (see Table 6.7), and inter-rater agreement was assessed using the Cohen's Kappa test. This test provides a measure of the level of consensus there exists within the ratings of independent judges. The test produces a score ranged from 0 to 1, where 0 indicates total disagreement, and 1 indicates complete agreement. When applied to the ratings of 'groups' and 'intermingled' categorisations, Cohen's Kappa produced a rating of 0.6667 agreement. The inter-rater categorisations are presented in Table 6.7.

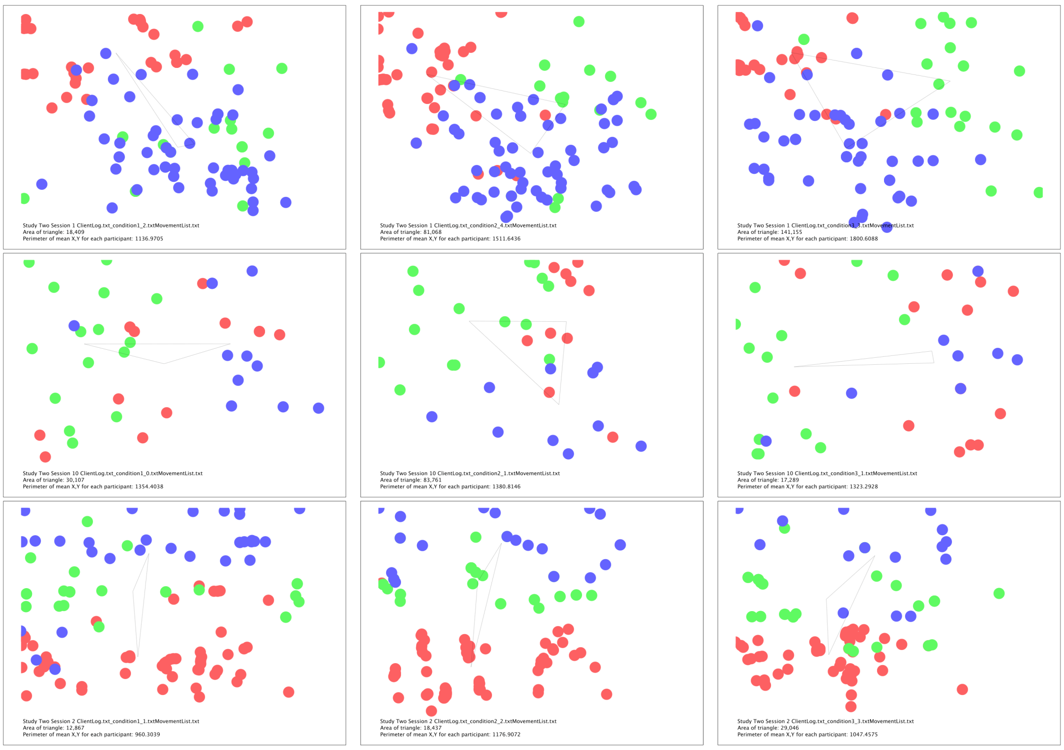


Figure 6.13: Example workspace visualisation

| Session | Condition | Rater 1 | Rater 2 |
|---|---|---|---|
| 1 | Condition 1 | Intermingled | Intermingled |
| 1 | Condition 2 | Intermingled | Intermingled |
| 1 | Condition 3 | Grouped | Intermingled |
| 2 | Condition 1 | Grouped | Grouped |
| 2 | Condition 2 | Grouped | Grouped |
| 2 | Condition 3 | Intermingled | Grouped |
| 3 | Condition 1 | Grouped | Grouped |
| 3 | Condition 2 | Grouped | Grouped |
| 3 | Condition 3 | Grouped | Grouped |
| 4 | Condition 1 | Intermingled | Intermingled |
| 4 | Condition 2 | Intermingled | Intermingled |
| 4 | Condition 3 | Intermingled | Intermingled |
| 5 | Condition 1 | Intermingled | Intermingled |
| 5 | Condition 2 | Grouped | Intermingled |
| 5 | Condition 3 | Intermingled | Intermingled |
| 6 | Condition 1 | Grouped | Intermingled |
| 6 | Condition 2 | Intermingled | Intermingled |
| 6 | Condition 3 | Grouped | Intermingled |
| 7 | Condition 1 | Grouped | Grouped |
| 7 | Condition 2 | Grouped | Grouped |
| 7 | Condition 3 | Grouped | Grouped |
| 8 | Condition 1 | Grouped | Grouped |
| 8 | Condition 2 | Grouped | Grouped |
| 8 | Condition 3 | Grouped | Grouped |
| 9 | Condition 1 | Grouped | Grouped |
| 9 | Condition 2 | Grouped | Grouped |
| 9 | Condition 3 | Grouped | Grouped |
| 10 | Condition 1 | Intermingled | Intermingled |
| 10 | Condition 2 | Intermingled | Intermingled |
| 10 | Condition 3 | Intermingled | Intermingled |

Table 6.7: Table of inter rater classifications for spatial visualisations

The visualisations and categorisations provide evidence that five out of ten groups partitioned the two-dimensional space of the shared workspace during their interaction. The visualisations also show that groups used a similar spatial arrangement in every condition, and the spatial arrangements appear not to be influenced by the audio presentation modes.

It is important to emphasise that the spatial position of modules within the virtual workspace does not influence the musical outcome of the software, meaning that the way windows were arranged by participants served no musical function, yet may have been relevant from an interactional standpoint as a support for awareness, coordination, or organisation of some form. Furthermore, as noted in Section 6.11.2 the experimental manipulations did not influence the amount of module movements participants made.

### 6.11.4 Spatial and non-spatial analysis

The Mann-Whitney U test was used to analyse the interaction logs for grouped and intermingled data sets independently, revealing several differences in the patterns of interaction between the participants who grouped their modules territorially and the participants who intermingled their modules. Groups who intermingled their modules performed significantly more coediting than groups who were spatially grouped ($U_a$ = 1354.5, z=-2.76, $p_1$=0.0029, $p_2$=0.0058). Groups with more pronounced spatial partitioning also created more modules ($U_a$ = 696, z=2.55, $p_1$=0.0054, $p_2$=0.0108) and made more use of the public channel ($U_a$ = 697.5, z=2.54, $p_1$=0.0054, $p_2$=0.0111).

### 6.11.5 Video Observation and Group Interviews

All interaction during the sessions was video recorded, and all groups were subjected to a video recorded group discussion at the end of the experiment session. This resulted in approximately ten hours of video footage. Videos of each experimental session were transferred to computer and divided into four separate files; the three audio presentation conditions plus the interview. The ELAN software[6] was used to transcribe speech from the videos. This resulted in 11,644 words of interview dialog, and 3,696 words of dialogue during from during the interaction itself.

The videos were analysed using coding techniques adopted from Constructivist Grounded Theory (Charmaz, 2006). Although the experimental approach adopted in this study precludes performing a full grounded theory analysis, the use of Grounded Theory Methods (GTM) is a recognised approach to analysing transcripts of interaction in HCI and CSCW research (Muller and Kogan, 2010). Grounded Theory methods have also been used to study digital musical interactions (Kiefer, 2010, Laney et al., 2010, Michael Gurevich and Marquez-Borbon, 2010, Morrison et al., 2007, Xambó et al., 2011).

To perform analysis using a Grounded Theory method, the transcripts were read through several times, and then openly coded. This involved a line-by-line reading of the text, and the assignment of codes to represent meaning. Following (Charmaz, 2006), codes were assigned using a gerund form[7] to retain an emphasis on the processes inherent in the activity and interac-

[6]http://www.lat-mpi.eu/tools/elan/
[7]words ending -ing

tion. Words and phrases used by the participants were also employed as codes, so as not to put distance between the analysis and the data. After the initial coding process, the codes were compared line-by-line with the transcripts to assess their fit and suitability. Appendix C.1presents a list of initial codes generated at this stage.

Similar codes were then grouped into higher level categories (Axial Codes) to form more broad descriptions. These categories were refined through further comparison with the transcripts and extended through notes and written descriptions to clarify their meaning. At this stage, excerpts were taken form the text to exemplify the categories. Whist not adhering to a full Grounded Theory methodology, this analysis provided a means of exploring and making sense of the large corpus of interview text collected. Interview extracts are presented throughout the discussion section to elaborate on specific features of the analysis. The following axial codes were identified from the interviews:

*Generating new ideas*

Participants described the experience in terms of generating ideas and brainstorming, and noted the benefits of getting three people's ideas at once.

*enjoying the experience*

Participants in the interviews talked about their enjoyment.

*Working in individual areas of the screen*

Participants work in individual areas of the screen to help them keep track of which contributions are theirs and which belong to other people. Sometimes the groups discuss where to work, but often this process of claiming screen space happens without discussion.

*Assuming a role*

Sometimes participants assume specific roles within the collaboration, for instance 'doing the drums', 'doing the textures'.

*Becoming familiar with the group*

Becoming familiar with the group as time moves forward appears to be a natural aspect of the group interaction, and is often mentioned by the participants.

*Close listening with headphones*

Participants often describe the close listening experience of using headphones, and the affordance of hearing more detail in the sound that way.

*Using personal channel before sharing*

Participants often describe the process of creating musical ideas in their personal channel before sharing them with the group. This appears to be an integral aspect of the collaborative activity, and a strong emergent code from the interviews.

*Identifying ones' self*

Identifying the sounds one was making

*Identifying others*

Attempting to identity others in the mix

*Losing track of time*

Some people note losing track of time, as they become more immersed in the activity. This code clearly echoes the idea of Flow and optimal experience.

## 6.12 Discussion

This study took as a starting point the question *how does the mode of audio presentation affect co-located collaborative digital musical interaction using an interface shared across multiple computers?* Particular concerns were how different forms of audio presentation influenced styles of working, discussion over the music and collaboration, and management of the shared interface. The results suggest a number of ways in which manipulating the way audio is delivered in a collaborative music-making activity changed the way participants used the software, and influenced the perceived quality of the interaction. This section begins by using the quantitative interaction log measures, questionnaires, and analysis of the video data to assess the hypotheses presented in Section 6.3. The discussion then expands on a number of more general issues surrounding the interaction which took place, drawing heavily on the Grounded Theory Methods (GTM) analysis of the experiment sessions and group interviews.

### 6.12.1 Hypothesis H1

Hypothesis H1 (Section 6.3) stated that *Participants will work more individually when audio is presented exclusively through headphones*. Key indicators of individual work were increased use of the Personal audio output, a reduction in module co-editing and less discussion between participants. In addition, some of the post-test questionnaire questions were in place to investigate the participants' reported impression of the interaction related to this hypothesis.

Interaction log analysis for editing and co-editing showed no significant effects caused by the experimental manipulation of audio devices. The same level of co-editing in the conditions suggests that the mode of audio delivery did not encourage participants to edit each others contributions more or less. There was also no effect on the amount of creations, deletions, or movement of modules within the interface. This suggests that manipulating the audio delivery devices did not alter the turnover of contributions or music modules.

Participants used the Personal audio channel more when it was routed through headphones (Conditions C2 and C3), although this did not appear to be further influenced by routing Public audio through speakers or headphones. According to the post-test questionnaire, participants reported having least privacy in the speakers only condition ($p=0.0211$, $df=2$, $csq_r=7.72$), although there was no significant reported distinction between the headphones only condition (C2) and the headphones+speakers condition (C3). This suggests that the largest difference in interaction was brought about by presenting the Personal channel through speakers, and not through presenting all audio exclusively through headphones.

During the interview, participants identified a number of distinctions between working in headphones and working in speakers. There was no strong indicator that presenting audio exclusively through headphones caused groups to work more individually, although participants often noted that the headphones encouraged more concentrated listening than with speakers, causing them to 'hear close things', or focus on 'texture' and 'tiny details'. The property of headphones to facilitate a more intimate, close and immersive listening experience was also identified by Kallinen and Ravaja (2007). This tendency for headphones to promote more focused listening had contradictory effects on reported experiences of involvement in the group. On the one hand some participants reported being less involved with the group as they become more focused on the details of the sounds they were creating, at the expense of engaging with the publicly shared music. On the other hand, for different participants, the more concentrated listening resulted in their becoming more attentive to the changes made by others, and consequently they noted feeling as though they were working more as a group when using headphones. These two effects are illustrated by transcript extracts from interviews with two separate groups of participants:

> PARTICIPANT 2: yeah, and I think you're much more aware of, of other people's changes, in the last one, having the headphones on [...] and I think, my feeling was that that encourages, encouraged us to change more in each others'. [...] that was my feeling anyway, that there was a bit more collaboration with each others' sounds
>
> (Session 6 Interview)

PARTICIPANT 3: well, erm, I see, the headphones definitely, shut me off from, erm, I mean I was able to concentrate solely on what I was doing, but I wasn't as involved as a group (Session 5 Interview)

Participant 2 from Session 6 described a situation in which headphones caused them to engage more with the group, as the close listening they promoted focused their attention on changes other people were making to the sounds. Conversely, Participant 3 clearly stated that the headphones allowed them to concentrate on their own ideas, but at the expense of less group involvement. The contradictory statements from participants do provide reassurance that the interview questions were not leading participants towards particular answers (Furniss et al., 2011), although the range of results also suggests that the individuals had quite widely different experiences of the interaction. It does appear that participants noted feeling more involved with the group when they were most aware of and attentive to the changes being made by others in the group, and this occurred primarily when public and Personal audio was presented via the same device (speakers *or* headphones), rather than when public and Personal audio were split into separate devices.

### 6.12.2 Hypothesis H2

Hypothesis H2 posited that *an interface which provides the greatest auditory separation between Personal and Public audio would be preferred by participants*. The post-test questionnaire data does not indicate that any of the conditions affected preference based on results for the statements 'The condition I preferred the most' or 'I enjoyed myself the most'. Analysis of post-test responses indicated that a significant proportion of participants rated condition 2 (Speakers + Headphones) as the one they most lost track of time in ($p<0.05$). This questionnaire statement is rooted in the theory of Flow, which suggests losing track of time is an indicator of enjoyment or engagement (Csikszentmihalyi, 1990), although this in and of itself is not enough to support the hypothesis.

During interviews participants expressed mixed responses to the headphones+speakers condition. Some participants expressed difficulty switching between headphones and speakers, noting in particular that it caused a disruption in their ability to focus on the shared aspects of the music. For instance one participant described using headphones initially to experiment with the software, before switching back to speakers to work with the group.

PARTICIPANT 1: yeah yeah, because I, I was just trying to get, see how it all worked, and, which I preferred, and then I found it was much better to just be on the same wavelength as everybody else (Session 10 Interview)

Similarly, one participant from another group noted that the headphones made them concentrate more on what they were doing individually, apparently at the expense of formulating a musical contribution which were coherent with the sounds playing through the Public audio channel on speakers. This contradicts the proposal by Morris et al. (2004) to use individual audio channels presented via in-ear headphones to support parallel working in a task based scenario.

> PARTICIPANT 2: I was more concentrating on what I was doing, and when I tried to erm, add it to the, you know, public thing, it was, it just didn't sound right.
> (Session 5 Interview)

In summary, the results do not support hypothesis H2, that splitting Public and Personal audio channels between devices would be most preferred. As with H1, it appears that splitting audio between headphones and speakers was actually problematic for a lot of the participants, although this is not reflected in the interaction logs analysis, which shows interaction in C2 and C3 were very similar.

### 6.12.3 Hypothesis H3

Hypothesis H3 posited that *an interface which presents Personal and Public audio entirely through speakers will cause participants to work more collectively*. Interaction log analysis shows that the audio presentation configurations did not influence individuals' tendency to edit modules created by other group members, and altering the form of audio presentation did not cause individuals to create significantly more or less music modes, or influence the amount of editing they performed. However, when Public and Personal audio was presented entirely through speakers (Condition 1), interaction log analysis using the Friedman Test showed that participants made significantly less use of the Personal channel to listen to musical contributions than they did in either of the other conditions ($p=0.0253$, $df=2$, $csq_r=7.35$). This result partially supports hypothesis H3, insofar as participants were less likely to listen to contributions via their Personal channel when Personal audio was routed to their speaker. This finding may however be more related to the depreciated benefit of using the Personal channel when it was routed through speakers.

Furthermore, response to the questionnaire self-reported statement 'I had the most privacy' shows participants reported experiencing least privacy in C1 ($p=0.0211$, $df=2$, $csq_r=7.72$). However, analysis of the post-test questionnaire results for the statements 'We edited the music together', 'We worked most effectively' and 'I worked mostly on my own' suggested that there was no trend for participants to report working more collectively in any of the audio presentation configurations.

One of the strongest descriptive codes in the interview analysis related to participants descriptions of control over the sharing of musical contributions, and the formulation of contributions away from the group. This is similar to the observed use of the Private space in Study one.

PARTICIPANT 1: yeah, you didn't want to introduce something terrible
(Session 5 Interview)

By reducing the privacy afforded by the Personal channel, the ability to formulate contributions in isolation before sharing would clearly have been negated, therefore reducing the usefulness of the Personal audio channel. However, forcing contributions to be more Public did not impact on the level of co-editing in the same way as reported in Study One.

When analysed for ordering effects, the post-test questionnaire results showed that a significant number of participants reported working most effectively in the final condition they were presented with, regardless of which audio presentation configuration this equated to. A significant number of participants reported working mostly on their own in the first condition they were presented with, and similarly that people also ignored their contributions more in the first presented condition. The final condition presented to the participants was most frequently selected for the statement 'I understood what was going on'.

Findings related to the order of conditions are to be expected, as it makes sense that over the the group members would have become more familiar with the software, and with each other. This familiarity would explain their reported feelings of working more effectively together and understanding more about what was happening in the later conditions. However, as the experiment controlled for learning effects by altering the condition presentation order for each group (see Section 6.5) this finding does not negate the other findings of the study, which have been identified independently of the condition ordering.

Some participants reported an unanticipated use of the Personal channel, as a mechanism for discovering which modules were producing which sounds. During group interviews participants talked in more detail about the importance of knowing what was making a specific sound than knowing who was responsible for a given contribution. For instance one participant noted of the Personal audio channel:

PARTICIPANT 1: yeah, I err, I used, yeah, I used it for two things. For, for monitoring

> before I submitted something, but also if I got, confused between, which sounds were, were what, then I would, would listen to it (Session 6 Interview)

Furthermore, analysis of the interviews suggested that the idea of becoming aware of *what* was creating sounds arose much more frequently than the category of *who* was creating a sound, and participants frequently noted not being aware of who was responsible for particular contributions.

In summary, presenting Personal Audio through speakers reduced the participants' reported level of privacy, and also reduced the tendency for the Personal audio channel to be used. This can be explained through the reduced level of privacy afforded by presenting Personal audio through speakers. This finding relates strongly to the findings of Study One (Section 5.11) by showing that the option for privacy gets exploited when made available to participants. However, as the experimental manipulations did not influence the way participants used the software (module creations, deletions, editing), it is difficult to find strong support for the hypothesis that presenting audio entirely through headphones caused the group to work more collectively. Instead, analysis of the self-report questionnaire data also implied that the time the group spent working together and/or using the music software was a stronger predictor of how collectively they reported working together.

### 6.12.4 Spatial Use of Workspace

Analysis of the way participants spatially organised the interface showed strong cases of territorial behaviour in half the groups, although there was also no evidence that the form of audio presentation influenced the way individuals or groups spatially arranged the modules within the shared on-screen workspace. The evidence for participants employing strategies for spatial organisation of the shared interface supports the work of Tse et al. (2004), who suggest people naturally adopt spatial partitioning strategies in shared Single Display Groupware interfaces. Tse et al. (2004) argue that users primarily partition their shared workspaces to avoid interference with each another's work, and that some collaborative tasks lend themselves to an inherent spatial partitioning strategy. Tse et al. (2004) also observe that seating arrangement around a shared screen can influence participants' on-screen working area. Similarly in interaction over conventional tabletops it has been shown that group members generally modify aspects of the group product which are within their proximity and reach (Scott et al., 2004). Unlike in the case of people seated or standing around a shared screen or table, the circular nature of the equipment and seating arrangement (see Figure 6.5), combined with the consistent spatial position of modules within the shared and distributed workspace means that the physical position of participants around the table could not have contributed to the way participants organised spatial layout of elements on the screen, while the primarily auditory activity presents no inherent cues or suggestions towards specific spatial arrangements. It is possible that the position of the public patching block in the middle of the screen may have directed participants towards a natural spatial partitioning strategy around the patching block, however the interface was identical for all conditions and the plurality of layout approaches (corners, horizontal and vertical stripes, non-uniform) therefore suggests that this was not a major contributing factor.

Participants may have spatially partitioned the interface to reduce interference with one anothers work. This statement is supported by interaction log evidence; using the Mann-Whitney test to compare instances of co-editing between the sets of partitioned and intermingled groups reveals that participants in groups with weaker spatial organisation strategies conducted significantly more co-editing (i.e., editing music modules created by others) than participants in groups with more strict spatial organisation ($U_a$=1354.5, z=-.2.76, $p_1$=0.0029, $p_2$=0.0058). However, conversation during the sessions indicate that the role of spatial partitioning stretched beyond the minimisation of interference, and was used to signify and help manage awareness of authorship of musical contributions. This follows the work of Thom-Santelli et al. (2009), who argue that during collaboration, territoriality serves the communicative function of indicating ownership over a particular object or space. It is also important to note that whilst the shared workspace supported the territorial arrangement of contributions, such strategies were not encouraged through the interface design itself. This finding is similar to the observation in (Thom-Santelli et al., 2009) that existing software features can be re-appropriated as markers of territory in a collaborative environment.

When asked in interview, the vast majority of participants seemed aware that they themselves were working in a particular area of the screen, and could remember which portions of the screen they were mostly working in. However, the participants did not always know where other group members were working. This implies that the spatial organisation of the interface did not contribute strongly to awareness of specific individuals, although it may have contributed to a more general awareness of ownership. Studying the video recordings, an absence of discussion over where in the shared workspaces people worked indicates that in many cases this partitioning was the result of tacit agreement, rather than the result of verbal negotiation.

While in many cases the groups did not discuss spatial layout, the following conversation extract from an experiment session demonstrates how some of the participants identified difficulties maintaining awareness of each other's actions. The extract then goes on to show the participants' negotiated solution to these difficulties by means of an explicitly agreed strategy to spatially organise the shared interface.

PARTICIPANT 3: The beat's pretty cool, hey?
PARTICIPANT 1: yeah, it's alright, man
PARTICIPANT 1: it's just hard to keep up with so much, what's going on. You don't know who's, who's is doing what, you know?
PARTICIPANT 2: *(laughs)*
PARTICIPANT 3: ok, ok well how about this, how about this.
PARTICIPANT 1: uh uh
PARTICIPANT 3: Why doesn't everybody, like, lets say, you go to one side with yours, I go to one side with mine, and you go to one side with yours. Like, it's to move it to one side so we know what everyone is doing.
PARTICIPANT 1: yeah, yeah
PARTICIPANT 3: that make sense?
PARTICIPANT 2: yeah
PARTICIPANT 3: so if I, can I go down, I can go down on the lower
PARTICIPANT 3: oh sorry
PARTICIPANT 3: I put myself down there, so
PARTICIPANT 2: you start moving your things
PARTICIPANT 3: yeah
PARTICIPANT 2: and then you
PARTICIPANT 3: yeah
PARTICIPANT 2: yeah
PARTICIPANT 1: yeah
[...]
PARTICIPANT 3: that's me
PARTICIPANT 1: where are you, at the bottom?
PARTICIPANT 3: yeah, that's me at the bottom, yeah
[...]
PARTICIPANT 2: you're going on the left, yeah?
PARTICIPANT 1: yeah, well I'm, I'm on the middle
PARTICIPANT 2: the middle?
PARTICIPANT 1: yeah
PARTICIPANT 2: so I'll go to the top, yeah?
PARTICIPANT 3: yeah, if you, if you climb, yeah
PARTICIPANT 3: so If I'm down low, and maybe someone is on the left hand side, and someone is on the right hand side, then everyone has enough space, you know what I mean? so
PARTICIPANT 1: or we can do top, middle and bottom, see, can do it either way really. I'm in the middle now
PARTICIPANT 3: OK
PARTICIPANT 1: OK

PARTICIPANT 3: that's you? In, that's you in the middle, yeah?
PARTICIPANT 1: yeah, just the three boxes
PARTICIPANT 3: OK

(Study Two Condition C2)

This extract demonstrates the way in which participants negotiated an informal mechanism or agreement to scaffold authorship awareness through partitioning of the workspace. The participants used the spatial division of the shared interface to create or claim workspace areas for themselves within the shared interface, even though such workspaces were not provided explicitly via the interface. This extract also highlights a common vocabulary used by participants to discuss the screen layout. During interviews, participants often described themselves in terms of working 'in the bottom left', 'top right', and so on, although some participants appeared not to have any sense of territory within the interface and talked about working 'all over the place', or 'putting stuff anywhere there was space'.

Another important feature of the participant's spatial arrangement of the workspace is that independent of any particular configuration (corners, strips, etc), the groups maintained a very similar arrangement of modules between conditions. This indicates that territorialisation was not influenced or affected by the experimental manipulations, but remained relatively constant. Gutwin and Greenberg (2002) used the example of people's physical bodily location in a workspace environment to highlight the notion that certain aspects of workspace awareness do remain static during collaboration, and therefore do not necessarily need explicit support in a software interface.

There were points at which the participants' schemes for spatial organisation broke down. For instance one participant noted during the interview

PARTICIPANT 1: you watch around for space, and start just putting stuff wherever

(Session 9 Interview)

Participants also occasionally used the shared and consistent spatial layout of the interface as a resource to discuss aspects of the musical arrangement, as demonstrated in the following excerpt. Here Participant 1 uses the spatially consistent workspace as a resource to draw Participant 2's attention to a particular music module.

PARTICIPANT 1: *....pointing with both fingers all over his screen* I can see what everyone else has got on their parameters
PARTICIPANT 2: yeah, that's right

PARTICIPANT 1: *.....pointing at right hand side of screen*because I've put, I've got, the second step sequencer, down on the right hand side, I've put the notes where your kicks are, *.....pointing to left of screen*
PARTICIPANT 1: so it's sort of
PARTICIPANT 2: ah, right, paralleled
PARTICIPANT 1: yeah

(Session 7 Condition C2)

In this incident, Participant 1 begins by gesturing with both hands towards the workspace on his screen, and then begins to point at an on-screen music module while talking about how he has configured it. Due to the physical arrangement of screens on the circular table, Participant 2 is unable to ground Participant 2's deictic reference to the information on his screen (Gutwin and Greenberg, 2002). Participant 1 therefore verbally refers to the music module's spatial position within the shared on-screen workspace, by stating 'the second step sequencer, down on the right hand side' of the workspace. In this way Participant 1 and Participant 2 use their inter-subjective knowledge of the spatially consistent layout of the workspace to discuss an aspect of the shared interface.

Participant 2 then notes that Participant 1's sequencer is 'paralleled', to acknowledge the observation that A's step sequencer is playing notes at the same time as the kick drums from Participant 2's drum module. This extract also demonstrates that participant Participant 1 was using visual access to Participant 2's music modules as a resource for creating musically coherent contributions. Finally, this scenario demonstrates the difficulty of identifying collaborative co-creation. Although both participants have created music modules on their own, and appear not to have edited each others modules, Participant 1 has used the visually shared workspace to create a contribution which is musically cohesive with Participant 2's existing contribution.

### 6.12.5 Roles

Participants discussed and adopted a variety of roles within the interaction. This observation is based on video evidence from during the group interaction and discussion with the participants at the end of the sessions. During interviews, the majority of participants talked about what could be termed *instrumental roles* associated with the generation of musical material, for instance 'doing the bass', 'doing the drums' and 'doing the melodies'. Some participants described their activities using more abstract terminology, for instance 'filling gaps in the sound', trying to 'fill out the other spectrum' or 'experimenting with weird stuff'. Finally, some participants noted moving from the role of generating musical material to that of mixing or taking on a more *compositional* role, as exemplified by the following extract:

PARTICIPANT 1: I was the one who was sort of, taking things in and out, mixing, mixing role, towards the end. And I saw because there was some elements that really worked I thought, and then others I was just into kind of putting in and out, and just thinking more of sort of a mix, which was nice to get to that point

(Session 6 Interview)

In the majority of cases groups did not explicitly assign or discuss roles, however, one group explicitly discussed and assigned roles at the start of their session, before starting to create music. They then decided to rotate these roles for the three conditions, to ensure each member of the group had a chance to perform each role. The following extract from during their interaction shows how these roles were chosen.

PARTICIPANT 3: I think the best way to do it, is, if, one person does the drums, one person does, the, the step sequencer, and somebody else does something else, because, if all of us are working on the drums, I think it's going to get too random
PARTICIPANT 2: yes
[...]
PARTICIPANT 2: what about drums, bass and lead, or something, just as three roles
PARTICIPANT 1: yep, yeah, could try that, OK, makes sense
PARTICIPANT 3: and, then, then we can rotate, because there's three to do isn't there
PARTICIPANT 2: yeah, so we can try different
PARTICIPANT 1: Right, which do you want?
PARTICIPANT 2: I don't mind at all
PARTICIPANT 3: You want to start with err, drums?
PARTICIPANT 1: OK, yeah, sure
PARTICIPANT 3: Lead?
PARTICIPANT 2: OK
PARTICIPANT 3: and I'll do, what did we say? Bass?
PARTICIPANT 1: Bass
PARTICIPANT 2: Bass (Session 7 Condition C1)

### 6.12.6 Preferences and Learning Effects

The post-test questionnaire responses were analysed in their presentation order to show the effects of time and ordering on participants experiences. This identified a number of ordering effects, particularly in relation to the quality of the music made and the degree to which the group members responded to each other's contributions. These ordering effects suggest that participants' enjoyment and reported preferences were influenced by the time spent using the software, and could have been related to their familiarity with the other group members. This finding is supported by studies of collaborative document writing, which indicate that as groups become more mature they are more inclined to suggest and make changes to each others' contributions (Posner and Baecker, 1992). The following statements from participants support this:

PARTICIPANT 1: just felt I was more involved in a way, I just felt, I just felt like I was more involved (Session 2 Interview)

PARTICIPANT 2: cos the ball start to roll, you know, a little more, more erm, into, into the things and so on (Session 2 Interview)

PARTICIPANT 3: just sounded, sounded better. 'Cos the other one's a little bit too fast to what I normally do, if anything, that's the only reason why (Session 2 Interview)

In another session, the same participant reported a preference for their final conditions, due to familiarity with the software, while Participant 2 gave the second condition, as he/she felt the group was working more effectively together.

PARTICIPANT 1: yeah, I think that I'd gotten more used to sort of, using the, the software better (Session 3 Interview)

PARTICIPANT 2: err, I preferred two, I just thought it had a groove to it, I don't know, it seemed like everyone was working together (Session 3 Interview)

### 6.12.7 Feedback from participants

Participants commented on the positive experience of creating music with other people using the software.

PARTICIPANT 3: I mean, it was fun you know, and it, and, that's the main thing, it's fun to be, to be working in real time with two other people (Session 10 Interview)

Other participants noted the potential of using it live, for instance when asked how they would describe the experience, one participant noted:

PARTICIPANT 3: I would say, we need to get one of these, and start playing it live (Session 9 Interview)

## 6.13 Design Implications

Based on the findings above, this section outlines a number of design implications for systems to support collaborative musical interaction. So as to be appropriate in a variety of contexts, these design implications are presented in a generalised manor. New research directions are highlighted as this discussion progresses.

### 6.13.1 Implications for Multiple Devices for Audio Delivery

In single user performance contexts such as DJing, the separation of audio into different devices has been identified as a central aspect of the practice. However in a real-time, co-located collaborative context, where multiple people are listening to a variety of sound sources simultaneously, splitting audio across multiple devices appears to be problematic from a design and usability perspective.

Although Mixed Focus interaction (discussed in Section 3.1.5) was supported through the results of Study One (Chapter 5), the findings of this chapter suggest that performing this separation by means of different audio presentation devices is detrimental or problematic for a number of reasons. Firstly, separating audio into different devices affected the participants ability to coordinate and manage their collaboration. In particular, during interview, some participants noted feeling less involved and less aware of their collaborators when the Personal and Public audio channels were split between two devices.

Secondly, problems arose due to logistic issues such as switching between headphones and speakers, balancing the level between speakers and headphones, and the disruption of headphones on both conversation and monitoring of audio played through speakers. These issues could be counteracted by using less acoustically isolated headphones, providing each user with individual control over the level of their audio outputs, and using wireless headphones to make switching between headphones and speakers less awkward.

Finally in regard to split audio delivery, a clear drawback to the findings of this study is that although the participants were musically inclined, they had limited experience of collaborative software, and specifically had limited experience of using the software developed for this study. Had the participants become accustomed to the split audio delivery design of Condition 2 they may have developed ways to deal with the problems they encountered. Consequentially, a strong implication emerging from this discussion is that for *first time users* of a collaborative music making environment, audio should be presented via a single device where possible (either headphones *or* speakers) as it appears to encourage stronger feelings of group involvement, and a greater sense of awareness between users.

### 6.13.2 Implications for Single Device Audio Delivery

Interaction log analysis of the data suggests that using speakers for shared and individual audio presentation discouraged people from using Personal audio channels. Previous research suggests that incorporating the ability for users to work in auditory isolation from the group had a positive impact on collaboration, by allowing people to formulate more complete contributions before sharing them. Designers must therefore balance the choice to discourage individual work against the benefits of allowing users to control how and when their ideas are shared with the rest of the group.

If the system design is intended to encourage users to work publicly, rather than supporting more mixed focus activities, then using speakers might be preferable. A system which is designed to promote mixed focus individual work might benefit from headphone presentation, as this allows users to concentrate on their own contributions and take advantage of the detailed sound provided by headphones. This design implication is supported by the log file analysis, which indicates significantly less use of the Personal audio channel when all audio was routed through speakers. As a compromise, a collaborative system could also incorporate a switching mechanism which allows individuals to select between using headphones or speakers.

### 6.13.3 User Configurability and Layout Features

The ten groups used and configured the shared workspace in a variety of ways. While these configurations were in some cases emergent, rather than the result of discussion or verbal negotiation, they appeared to be important ways for individuals to keep track of their own contributions, and on a larger scale important for groups to maintain a spatially organised representation or ownership. Given the importance of module layout during collaboration (and the precedent this has in previous studies of collaboration), a redesigned interface could incorporate additional layout and organisation features to support additional scaffolding for collaboration, awareness and joint attention. These features would not have a direct influence on the sonic output of the software, but could aid groups of people in structuring a collection of interface elements so as to support and scaffold their collaboration. Additional features could include user configurable dividers, partitions, annotations and colour coded areas. The ability to group or bundle associated modules together (e.g., a collection of drum sequencers forming a rhythm) could be another useful feature.

However, it is important to remember that groups who employed more spatially territorial layouts were less inclined to edit the contributions made by others. Designers should therefore keep in mind that supporting spatial layout through interface features might impact on the amount of joint editing which takes place.

In collaborative environments which do not feature a workspace metaphor (e.g., a timeline based sequence application), separate visual representations could be incorporated, as well as other mechanisms to highlight, group or visually organise musical contributions. A new research direction might be to investigate the extent to which these organisational features need to be consistently duplicated across all users, and how groups or individuals might exploit the affordances of these features to support personalised and group level organisation of the interface.

## 6.14 Summary

This study has started to unpack the complexity surrounding the way Public and Personal audio can be presented during CDMI. A key finding was that presenting Public and Personal audio through different devices did not appear to be a successful design decision, as it caused confusion and made it difficult for participants to keep track of both channels. Presenting Public and Personal audio through speakers caused less use of the Personal channel, but may have affected the ability for participants to confidently formulate musical contributions by themselves. For some participants, headphones was a more focused listening experience, in some cases leading to increased awareness of others. This plurality of responses suggests that designers should pay particular attention to the way audio is presented in a collaborative environment, as it can have strong and varied effects on users.

The analysis identified spatial organisation strategies in the way the shared interface was used by the participants. These strategies appeared to be independent of the experimental interventions studied, and were in some cases emergent, rather than the result of negotiation between group members. The identification of these patterns led to further analysis of the data, revealing that more spatially territorial groups performed less co-editing.

The findings have suggested a number of design implications for future systems. One implication is that using audio delivery as a means of providing awareness information about the activity of others is problematic. The following chapter picks up on the design implications related to the identification of contributions and authors, and performs a detailed analysis of the effects different interface provisions for interrogation of the audio mix have on the way a collaborative music making environment. This approach puts participants in control of what awareness information they receive by letting them create their own mix of the musical contributions currently shared within the interface.

# Chapter 7

# Study Three: Source and Authorship Awareness

---

*The third study of this thesis investigated the effects of two different awareness mechanisms, one for presenting information about the Source of sounds, and one for presenting information about Authorship of contributions. Results suggested that presenting Authorship Awareness discouraged participants from collectively editing, and encouraged individual work. Source Awareness had a limited impact on the way participants used the software.*

The previous study investigated the effects of audio delivery in a collaborative music environment. Analysis showed that when Public and Personal audio were combined and presented through speakers only (rather than headphones or a combination of speakers and headphones), participants were less inclined to use the privacy features of the software to listen to contributions individually. In addition, it was discovered that the groups of participants adopted spatial approaches to arranging the shared workspace, and that this contributed to their ability to manage and maintain awareness of the authorship of contributions. Groups of participants for the most part used the same spatial arrangements in all their conditions, and groups who adopted more strictly separated layouts between group members' contributions were statistically less likely to co-edit one anothers' musical contributions. One possible implication of this finding is that an increased awareness of *who* created a particular musical contribution had an influence on the tendency for collaborators to edit items which they did not initially create.

Finally, participants reported using their Personal audio channel within the interface as a means of identifying the source of particular sounds within the mix. This behaviour could be regarded as representing a deficiency in the awareness provision of the software, as it indicated participants were re-appropriating other features within the software to compensate for a lack of awareness information pertaining to the source of particular sounds. Furthermore, by routing audio to a personal channel to determine its source, a user would also be disrupting the shared music by removing this sound from the collection of modules routed to other users via the Public audio channel. This could potentially cause interference with the activities of others within the group, particularly if they were currently listening to, or editing a shared musical contribution which was then removed from the Public audio channel by another user. The study presented in this chapter investigates the impact of awareness mechanisms to support both awareness of who contributed, and what the source of sounds are within the interface.

## 7.1 Authorship (Who) Awareness and Source (What) Awareness

The findings of the previous study indicate that users of a shared environment for collaborative music-making could benefit from an interface which separates the presentation of awareness information about which interface component represents a sound, and who was responsible for creating it. In Study Two, authorship awareness appears to have been maintained through spatial organisation of the interface, while some participants described re-appropriating the Personal audio channel as a means of discovering which modules were creating which sounds.

This leads to the delineation of two distinct forms of awareness within collaborative digital musical interaction. The first is an awareness of *who* is responsible for the authorship of particular sounds or contributions within the shared music. Secondly; within the shared interface an awareness of *what* is responsible for producing a particular sound, or in other words, an awareness of the *sound source*. While based in the theory of Workspace Awareness (Greenberg, 1996), these concepts extend the theory to apply to auditory media such as music.

Related concepts have been partially addressed in previous research. Within the context of Acousmatic music, Smalley (1997) used the term ‘Source Bonding’ to describe the attribution of a sound to a known or unknown source, however Smalley’s theory does not touch on the issue of authorship, and is focused primarily on individuals listening to pre-recorded material rather than groups of people co-creating music in real time. Within research on laptop ensembles Merritt et al. (2010) identified the problems of ‘who makes what sound?’ and ‘how is the sound being altered?'. While their work is highly relevant, it considers musicians using individual or personal means of sound generation, rather than the case of interaction via a shared interface which allows all users to see and modify each others' contributions. As a consequence of this distinction, the study presented in this chapter addresses the subtly different questions of 'who is responsible for which sound?' and 'which interface element is responsible for which sound?'.

The study presented in this chapter investigates the practical application of these concepts through posing the research questions *how do mechanisms for authorship awareness and source awareness measurably alter the way people engage in the process of collaborative digital musical interaction?* These questions were addressed by observing the effects of providing different auditory and visual indicators of authorship (who) and sound source (what) awareness in a collaborative music environment. In particular the study investigates how allowing users to individually isolate audio from different musical contributions and from different people within the on-going musical collaboration affects the process of musical collaboration and the way the collaborative software environment was used to make, share and edit musical contributions. Specific features of interest are quantitative differences in the way individuals use the software as a result of the additional awareness mechanisms, differences in the way groups co-ordinate their activities and the influence of these awareness mechanisms on self-reported measures such as preference.

## 7.2 Hypotheses

The following hypotheses operationalise the research questions discussed above:

H1 an environment which incorporates additional *Source* awareness will elicit the following effects:

- Increase instances of co-edited as participants will be able to accurately identify the source of specific aspects of the music
- a positive impact on preferences
- Increase participants' reported levels of awareness of one another

H2 environments which incorporate additional *Authorship* awareness will elicit the following effects:

- Decrease instances of co-editing
- feature less discussion over authorship management (e.g., spatial partitioning, discussion over who is doing what)

– a positive impact on preferences

## 7.3 Experiment Task

The participants were set the same task as in the previous studies; creating music based on a short video loop. As different participants were to be recruited, the same videos were used as in Study Two (see Section 6.4.1). The 'Snapshot' task was not used in this study.

## 7.4 Collaborative Music Software

The software used for this study was extended from the collaborative music environment developed previously in Study Two (see Section 6.2). The music-making functionality was left unchanged, although two additional awareness mechanisms were implemented as a means of allowing the user to select different subsets of sounds to listen to within the mix of music modules and effects. These awareness mechanisms were named Module Soloing and Client Soloing, after the 'soloing' function found in mixing desks and digital audio workstations.

### 7.4.1 Module Soloing

The Module Soloing function enables a user to solo the audio from a particular module or collection of modules within the software interface. Modules are soloed using a 'Solo' button found on the top of the module, as shown in Figure 7.4. Once a module has been soloed the sound from all other non-soloed modules is no longer audible. Soloed modules are also visually highlighted with a yellow background against the grey software workspace (see Figure 7.1). Any number of modules can be soloed simultaneously using the Module Solo function, resulting in the user hearing a combination of all soloed modules. If no modules are soloed then all are made audible.

### 7.4.2 Client Soloing

The Client Soloing function allows a user to highlight all the Music Modules created by a specific person within the collaboration. This action is performed by ticking against the users' name in a box to the left of the screen (see Figure 7.3). Once a Client has been soloed, only the modules created by this Client are audible, and modules created by other Clients within the collaboration are no longer audible. Any number of Clients can be soled at the same time. It is possible for a user to solo themselves in order to listen to only the modules they have created. Modules soloed using the Client Soloing feature are also colour highlighted within the workspace, as shown in

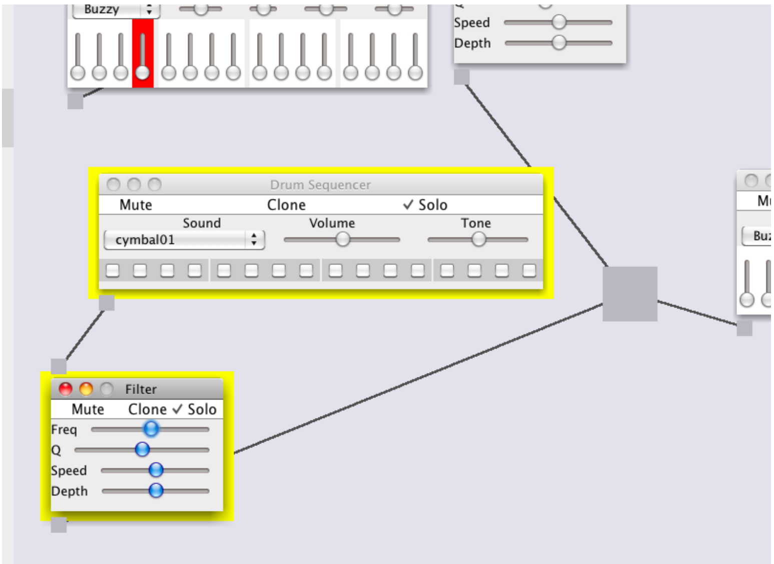



Figure 7.1: Screenshot of Module Soloing, showing the Solo buttons at top right of each Module Window and colour highlighting for soloed modules

Figure 7.2. Each user is given a different colour. These colours were consistent on the interfaces for all participants. The Client Soloing checklist itself was not colour coded as it was possible to determine which colour was associated with which user by checking and un-checking.

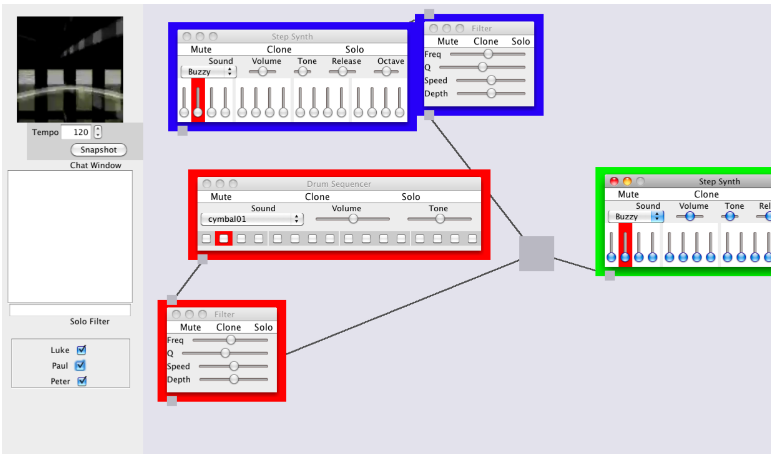



Figure 7.2: Screenshot of colour coding for Client Soloing

Soloing an audio effect (for example a Reverb or Delay module) causes all modules connected to the input of that effect to also be soloed. Soloing an effect module at the base of a long chain of modules causes all modules above it in the chain to also be soloed. Similarly, soloing an instrument module (Step Synth or Drum Sequencer) connected through an effect causes the processed sound to be heard. It is possible to use the Client and Module soloing functions at the same time, causing all soloed modules to be made audible while non-soloed modules are rendered inaudible. Soloing a module using both Client and Module soloing simultaneously has no additional effect on which modules are audible.

### 7.4.3 Individual Audio Mixes

Both forms of soloing are applied to individual users, rather than operating globally for all users. This means that with a group, each user may hear different combinations of music modules, depending on which solo functions they are currently using. The software provides no indication of which modules other users are currently soloing at any given moment.

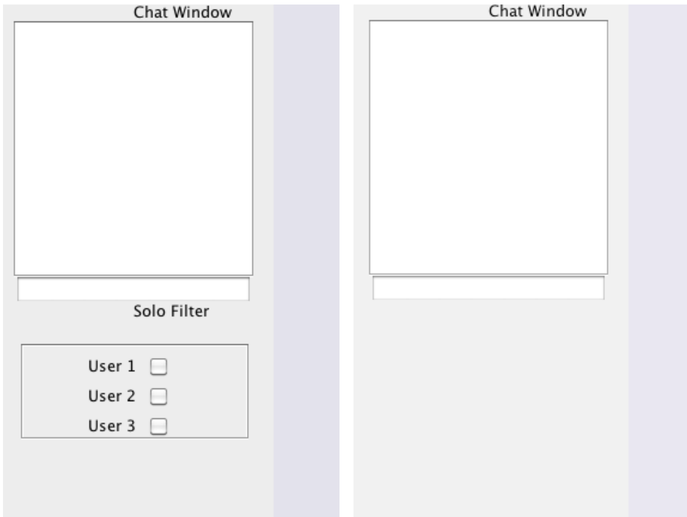
### 7.4.4 Implementation of Soloing Functions

Implementation of the soloing functions required a number of core modules within the code base to be modified. The Server was modified to store for each connected Client a representation of which other Clients were soloed by the particular user. The Server was also required to store information about which modules were currently soloed by each user. This stored information is then used to route or block audio generation messages to each connected Client, while ensuring that soloing and un-soloing does not affect all users receiving an identical representation of the audio mix when no modules are soloed.

The Server-side data logging mechanisms were modified to record the state of the Module and Client soloing functions. The Server control panel was modified to allow the Client Soloing and Module Soloing functions to be disabled via checkboxes. Figure D.2 (in appendix) shows the modified Server window. See Figure 7.3 for a comparison of the interfaces with and without Client Soloing, and Figure 7.4 for a comparison of Music Modules with and without Module Soloing.

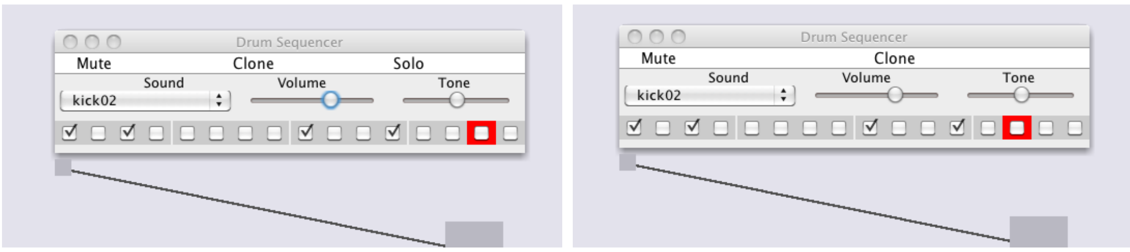
### 7.4.5 Other Application Changes

Privacy was not a central feature of this study, and consequently the Private outputs (described in Section 6.2.2) were disabled from the interface to reduce the complexity and simplify the training process. Future research should address the effect of incorporating privacy into the environment (see Section 9.2).




(a) with Client Soloing (b) without Client Soloing

Figure 7.3: Interface with and without Client Soloing enabled




(a) with Module Soloing (b) without Module Soloing

Figure 7.4: Module with and without Module Soloing function enabled. Note absence of Solo button in (b)

## 7.5 Independent Variables

For reasons outlined below, the study uses a four condition experimental design adopted from a previous study of collaborative music making by Bryan-Kinns and Hamilton (2009). In this design each participant group is subjected to two conditions, and the experiment features both within-subjects and between-subjects conditions. All participant groups were exposed to a within subjects condition of Authorship Awareness which was operationalised through the inclusion of a ‘Client Soloing’ function that enabled users to visually highlight and aurally solo contributions made by specific users within the group. A between subjects comparison of Source awareness was also performed by presenting half the subject groups with an additional ‘Module Soloing’ interface feature that allowed specific Music Modules to be visually highlighted and aurally soloed regardless of who created them.

This experimental design allowed for pair-wise testing for the within subjects effect of incorporating Authorship awareness through the inclusion of Client Soloing, and allowed for pair-wise testing for the between subjects effect of Source awareness through the inclusion of Module Soloing. Table 7.5 outlines the four conditions. Authorship awareness was selected for the within-subjects comparison as this would show how adding or removing awareness information about how is responsible for which contributions would affect interaction within the same group of people. Source awareness was selected for the between subjects comparison as this form of awareness is by definition more anonymous than Authorship awareness.

The four condition mixed design differed from the previous two three condition designs employed in Chapters 5 and 6, and was selected in response to several factors. Firstly, so as to address the learning effects observed and reported by participants in the previous study, it was decided that participants should be provided with a longer period of time for training and familiarisation with the software. However, ethical and financial constraints prevented the study sessions from being extended in time beyond one and a half hours. By presenting subjects with only two conditions, more time could be given over to training and providing participants with unstructured free time to familiarise themselves with the software before being presented with the experiment conditions. In association with this, participants also had more time to ask the researcher questions about the software. Secondly, in addition to a longer training period, using two conditions within an experiment session (instead of three) enabled the participant group to be assigned twenty minutes with each study condition, instead of the previous fifteen minutes per condition, as was the case in studies one and two. Finally, the mixed design enabled a more complex comparison of conditions to be performed across multiple participants within the limited time frame of the study sessions. More detail about the procedure and apparatus is given in Section 7.9.

### 7.5.1 Study Three Condition Naming Convention

The remainder of this chapter refers to the conditions using an alpha-numeric abbreviation with the M and C letters representing Module soloing and Client soloing, suffixed with 0 or 1 to represent the presence (1) of absence (0) of this interface element within the condition. For instance M0C1 represents a condition which does not contain Module soloing but which does contain Client Soloing. This naming convention is also used in Table 7.5.

| | Within Subjects (Columns) | |
|---|---|---|
| | Groups 1-8 | Groups 9-16 |
| Between Subjects (Rows) | **Condition M1C0**<br>specific Module soloing (M1)<br>no soloing by Client (C0) | **Condition M0C0**<br>no specific Module soloing (M0)<br>no soloing by Client (C0) |
| | **Condition M1C1**<br>specific Module soloing (M1)<br>soloing by Client (C1) | **Condition M0C1**<br>no specific Module soloing (M0)<br>soloing by Client (C1) |

Table 7.1: Summary of experiment conditions for Study Three. Columns represent the within subjects groups, and table rows show the between subjects comparisons. Therefore a left-right comparison of the groups shows the Module Soloing condition which was compared using independent groups of participants, whilst a top-bottom comparison shows the Client Soloing conditions which were compared using the same subjects.

## 7.6 Dependant variables

This section describes the measures taken during and after the participants' interactions with the software.

### 7.6.1 Post-Test Questionnaire Data

A post-test questionnaire asked participants to select which of the software conditions was most applicable to a list of statements. This questionnaire was adapted from the Mutual Engagement Questionnaire (Bryan-Kinns and Hamilton, 2009). A computer based questionnaire system was used to administer the questionnaires and automatically log the results. The post-test questionnaire statements were:

**The best music**

**I felt most involved with the group**

**I enjoyed myself the most**

**I felt out of control**

**I understood what was going on**

**I mostly worked on my own**

**I lost track of time**

**Other people ignored my contributions**

**We worked most effectively**

**The interface was most complex**

**I had the most privacy**

**I knew what other people were doing**

**We edited the music together**

**I made my best contributions**

**I was influenced by the other people**

**The condition I preferred the most**

### 7.6.2 Interaction Log Analysis

The software logs all interaction, including mouse clicks, key presses, and details of which software functions and operations were used, as well as how much time was spent using each function. Software tools were developed using Python and Java to extract quantitative data from the logs. Table 7.2 summarises the features studied, and how they were identified in the log files.

| Feature | Description |
|---|---|
| Creations | Creating a module |
| Deletions | Deleting a module |
| Cloning | Cloning a module using the clone button |
| Snapshots | Taking a group snapshot of the music |
| Text chat | Number of text chat events |
| Editing (raw) | Modifying a module control (slider, button, etc.,) |
| Patching to Public | Connecting a module to the Public output |
| Movement | Movement of modules |
| Tempo changes | Alterations to the global tempo control |

Table 7.2: Summary of log file features studied

### 7.6.3 Group Discussions

Semi-structured discussions were held with the participants at the end of the session to gather data about their perceptions and experiences. Discussions were video and audio recorded. A set of questions were generated to focus on general preferences, perceived differences between the conditions, the use of the awareness mechanisms, roles and working strategies and spatial use of the shared on-screen workspace. Participants were also invited to raise any other issues, talk about the experience in their own words and ask their own questions. Analysis of interview transcripts used the Thematic Analysis methodology described by Braun and Clarke (2006).

## 7.7 Sample Size

The required sample size for the study was calculated using the G*Power application[1]. This program can be used to calculate a sample size based on the the alpha value and desired statistical power to be achieved. G*Power is able to calculate a sample size a priori using mean and standard deviations collected from, for instance, a pilot study (Prajapati et al., 2012). As the two previous studies had been based on a different design, mean and standard deviations were instead used from the study by (Bryan-Kinns and Hamilton, 2009), which featured an identical experimental design, and a similar, although less complex collaborative music interface.

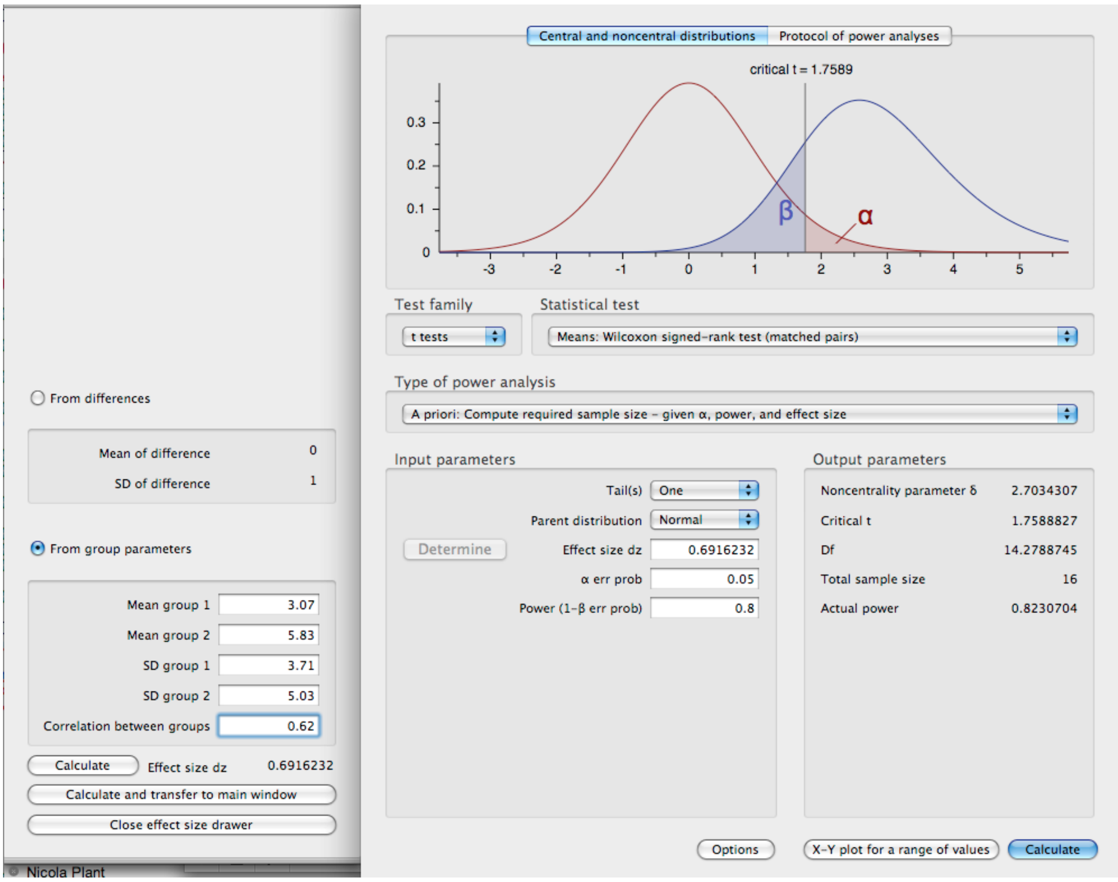



Figure 7.5: Screenshot of the G*Power readout to calculate sample size.

G*Power calculated a sample size of 16 as sufficient to achieve significant P values at a statistical power of 0.8, which is defined in Prajapati et al. (2012) and Faul et al. (2007) as a ‘large’ effect size (see Figure 7.5). As the experiment design required comparisons across two sets of groups (see Section 7.5), a sample size of 24 participants was chosen for each set, comprising of 8 groups of 3 participants, making the sample for each group larger than suggested

[1]http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/ Last Accessed 15th October 2012

by G*Power. In total, this required the recruitment of 48 participants, a sample size larger than used in the study by Bryan-Kinns and Hamilton (2009), which featured 39 participants in an identical experimental design.

## 7.8 Participants

Initially forty eight participants were recruited using online mailing lists to take part in the study. Each participant received £10 for taking part. These people were organised into seventeen groups of three participants. The first eight groups were presented with an interface featuring Module Soloing, and were exposed to a within subjects condition of Client Soloing or No Client Soling (conditions M1C1 and M1C0). This is shown in the first within subjects column in table 7.5. The second eight groups were presented with an interface featuring no Module Soloing, and were exposed to the same within subjects condition of Client Soloing or no Client Soloing (Conditions M0C1 and M0C0). This is shown in the second within subjects column in Table 7.5.

One of the first eight groups was excluded from the data set due to one of the recruited participants choosing not to take part in the experiment. This session was repeated with three new participants to make up the numbers.

A pre test questionnaire gathered demographic information about the participants. Appendix D presents a detailed account of the participant demographics. In summary, the participants had a high level of musical competence and knowledge. The majority of participants (88%) could play a musical instrument, almost all participants had previously written pieces of music on their own or with other people. On average the participants were also experienced in the use of computers for a variety of activities.

## 7.9 Procedure and Apparatus

Three identical Apple Mac Mini computers with keyboards, mice and 21 inch widescreen displays were arranged on a round table in a laboratory room (shown in Figure 7.6). Headphones were used exclusively to present audio, as this allows each user to be delivered a personalised audio mix depending upon how they choose to solo specific music modules within the interface. Identical sets of Audio-Technica ATH M30 headphones were connected to each computer for the purpose of audio delivery.

Despite differences in the lengths of training period and interaction, and a change in the num-

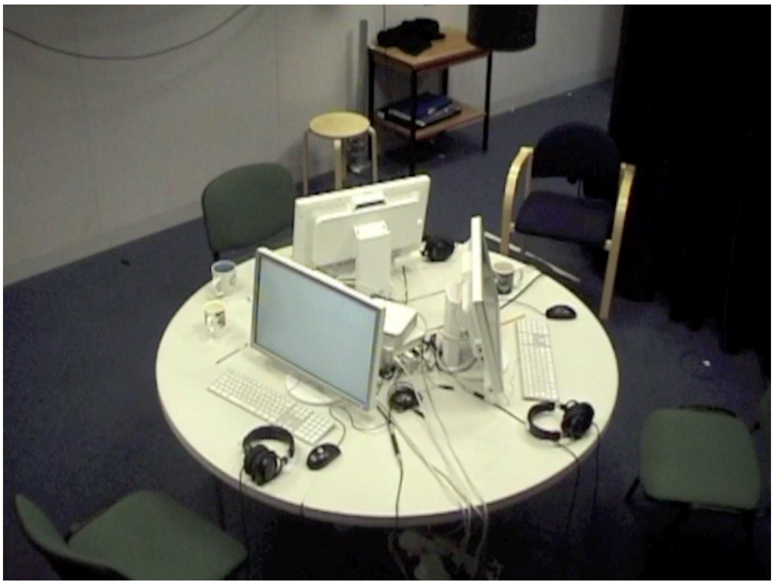


Figure 7.6: Hardware setup

ber of experimental conditions, the procedure and practical running of the experiment sessions was largely identical to Study Two (see Section 6.9). The session began with the researcher outlining what would happen during the session, and inviting participants to ask any questions before reading an information sheet and signing a consent form. A computerised pre-test questionnaire was then presented to collect demographic information (presented in Section 7.8). A training period followed, which began with a hands on demonstration of how to use each feature of the software. Participants were then given ten minutes of unstructured time to explore the software. Participants were invited to ask questions about the software throughout the training period. Each group of participants was then subjected to two experimental conditions, each lasting twenty minutes. The computerised post-test questionnaire was then presented, and the session closed with a group discussion. Each session was structured as follows:

- Introduction, Consent form signing, pre-test questionnaire - Approximately 20 minutes
- Training - 20 minutes
- Experimental Conditions - 40 minutes
- Post-test Questionnaire - Approximately 15 minutes
- Discussion and Debriefing - Approximately 15 minutes

## 7.10 Testing and Pilot Studies

The software was tested thoroughly for software errors and bugs before being used in pilot studies. Pilot study participants were recruited from within Queen Mary University of London. Pilot sessions were run to test the software environment, the data collection mechanisms, questionnaires and general running of the experiment sessions. The pilot studies highlighted a number of usability issues with the software interface, specifically related to the position of the soloing functions and the layout of the interface. Initially the module soloing buttons were located in a column to the left hand side of the screen, although the pilot study indicated that it would be more useful if it was located on the module window itself. After this modification was implemented a further pilot was run with this design to test for bugs and other unforeseen issues. In addition to testing the software interface and associated hardware, the pilot studies also enabled testing of the automatic data logging and analysis tools.

## 7.11 Quantitative Results

This section presents the results obtained from analysis of the experiment session data. As noted above, the experimental design and use of statistical tests is identical to the study presented in Bryan-Kinns and Hamilton (2009). A discussion of the findings follows in Section 7.13.

### 7.11.1 Effects of Client Soloing on Post-Test Questionnaires

The experimental design used a within-subjects comparison to observe the effects of Client Soloing, meaning all participants were exposed to a condition with Client Soloing and a condition without Client Soloing. The Wilcoxon Signed-Rank test was used to identify the effects of Client Soloing on the post-test questionnaire responses. The Wilcoxon Signed-Rank test is a non-parametric statistical test used for the comparison of two related samples (Greene and D'Oliveira, 1999), and is therefor an appropriate test in this instance.

The Wilcoxon Signed-Rank test was applied to the post-test questionnaire responses for conditions M1C0 and M1C1 (groups with Module Soloing), and for conditions M0C0 and M0C1 (Groups without Module soloing). Table 7.3 presents the within subjects comparison of post-test responses for both sets of groups. As half the groups were presented with Module Soloing in both their conditions, and half the groups were not presented with Module Soloing in either of their conditions, the test was applied independently to both sets of data.

For groups who were presented with Module Soloing in both conditions (groups M1C0 and M1C1), there was a trend (w= 125, n s/r=24, Z=1.78, p=0.0516) for participants to express preference for interface M1C1 over M1C0. A significant proportion of participants stated that in interface M1C1 they 'knew what the other people were doing' (W=99, n s/r= 18, Z=-1.64, p=0.0436), and that they worked mostly 'on their own' in interface M1C1 (W=121, n s/r= 21, Z=2.09, p=0.0183).

For groups who were not presented with Module Soloing in either condition, a significant proportion indicated feeling 'out of control' in condition M0C0 (W=-99, n s/r=21, Z=-1.71, p=0.0436). A significant proportion indicated that they 'knew what other people were doing' in condition M0C1 (W=156, n s/r=23, Z=2.36, p=0.0091), and a significant number of participants indicated preference for condition M0C1 (W=125, n s/r=24, Z=1.78, p=0.0375).

| Statement | C1 | C0 | W | n s/r | Z | p |
|---|---|---|---|---|---|---|
| The best music | | | | | | |
| with Module Soloing | 12 | 11 | 12 | 23 | 0.17 | 0.4325 |
| without Module Soloing | 14 | 10 | 50 | 24 | 0.71 | 0.2389 |
| I felt most involved with the group | | | | | | |
| with Module Soloing | 10 | 12 | -23 | -23 | -0.37 | 0.3557 |
| without Module Soloing | 12 | 12 | 0 | 24 | 0 | 0.5000 |
| I enjoyed myself the most | | | | | | |
| with Module Soloing | 12 | 10 | 23 | 22 | 0.37 | 0.3557 |
| without Module Soloing | 14 | 9 | 60 | 23 | 0.9 | 0.1841 |
| I felt out of control | | | | | | |
| with Module Soloing | 8 | 12 | -42 | 20 | -0.77 | 0.2207 |
| without Module Soloing | 6 | 15 | -99 | 21 | -1.71 | **0.0436** |
| I understood what was going on | | | | | | |
| with Module Soloing | 12 | 10 | 23 | 22 | 0.37 | 0.3557 |
| without Module Soloing | 15 | 9 | 75 | 24 | 1.06 | 0.1446 |
| I mostly worked on my own | | | | | | |
| with Module Soloing | 16 | 5 | 121 | 21 | 2.09 | **0.0183** |
| without Module Soloing | 14 | 9 | 60 | 23 | 0.9 | 0.1841 |
| I explored more | | | | | | |
| with Module Soloing | 12 | 12 | 0 | 24 | 0 | 0.5000 |
| without Module Soloing | 12 | 12 | 0 | 24 | 0 | 0.5000 |
| I lost track of time | | | | | | |
| with Module Soloing | 7 | 12 | -50 | 19 | -1 | 0.1587 |
| without Module Soloing | 13 | 11 | 25 | 24 | 0.35 | 0.3632 |

Continued on next page

Table 7.3 – continued from previous page

| Statement | C1 | C0 | W | n s/r | Z | p |
|---|---|---|---|---|---|---|
| Other people ignored my contributions | | | | | | |
| with Module Soloing | 6 | 7 | -7 | 13 | -0.23 | 0.4090 |
| without Module Soloing | 10 | 6 | 34 | 16 | 0.87 | 0.1922 |
| I was more focused | | | | | | |
| with Module Soloing | 14 | 8 | 69 | 22 | 1.11 | 0.1335 |
| without Module Soloing | 14 | 10 | 50 | 24 | 0.71 | 0.2389 |
| The interface was most complex | | | | | | |
| with Module Soloing | 10 | 11 | -11 | 21 | -0.18 | 0.4286 |
| without Module Soloing | 14 | 8 | 69 | 22 | 1.11 | 0.1335 |
| I had the most privacy | | | | | | |
| with Module Soloing | 10 | 8 | 19 | 18 | 0.4 | 0.3446 |
| without Module Soloing | 8 | 13 | -55 | 21 | -0.95 | 0.1711 |
| We worked most effectively | | | | | | |
| with Module Soloing | 9 | 12 | -33 | 21 | -0.56 | 0.2877 |
| without Module Soloing | 16 | 8 | 100 | 24 | 1.42 | 0.0778 |
| I knew what other people were doing | | | | | | |
| with Module Soloing | 15 | 6 | 99 | 21 | 1.71 | **0.0436** |
| without Module Soloing | 18 | 5 | 156 | 23 | 2.36 | **0.0091** |
| We edited the music together | | | | | | |
| with Module Soloing | 5 | 13 | -76 | 18 | -1.64 | 0.0505 |
| without Module Soloing | 15 | 8 | 84 | 23 | 1.27 | 0.1020 |
| I made my best contributions | | | | | | |
| with Module Soloing | 12 | 12 | 0 | 24 | 0 | 0.5000 |
| without Module Soloing | 11 | 10 | 11 | 21 | 0.18 | 0.4286 |
| I was influenced by the other people | | | | | | |
| with Module Soloing | 5 | 14 | 1 -90 | 19 | -1.8 | **0.0359** |
| without Module Soloing | 12 | 12 | 0 | 24 | 0 | 0.5000 |
| The condition I preferred the most | | | | | | |
| with Module Soloing | 16 | 7 | 108 | 23 | 1.63 | 0.0516 |
| without Module Soloing | 17 | 7 | 125 | 24 | 1.78 | **0.0375** |

Table 7.3: Wilcoxon Signed-Rank analysis of post-test questionnaire responses, comparing Client Soloing and no Client Soloing, with and without Module Soloing

### 7.11.2 Effects of Client Soloing on Software Interaction

The effects of Client Soloing on the interaction logs was analysed through a within-subjects design, where all participants were exposed to a condition with Client Soloing and a condition without Client Soloing. The Wilcoxon Signed-Rank test was used to perform this within-subjects comparison on the interaction log features discussed in Section 7.6, and identify the effects of Client Soloing on the way participants interacted with the software. The Wilcoxon Signed-Rank test is a non-parametric statistical test used for the comparison of two related samples (Greene and D'Oliveira, 1999), and is therefor an appropriate test in this instance.

As half the groups were presented with Module Soloing in both their conditions, and half the groups were not presented with Module Soloing in either of their conditions, the test was applied independently to both sets of data. Table 7.4 presents the results for the effects of Client Soloing when the interface contained Module Soloing and also for conditions where the interface did not include Module Soloing.

For groups where the interface contained Module Soloing in both conditions, analysis showed a strong trend for participants to make more co-editing operations when the interface did not include the Client Soloing feature (See Figure 7.7), although this was slightly above the level of statistical significance (w=-87, n/sr= 20, z=-1.61, p=0.0537). There were no significant effect on Module Creations, Deletions, Editing, Module Cloning, Connection to the Public Audio Channel, or Movement of modules.

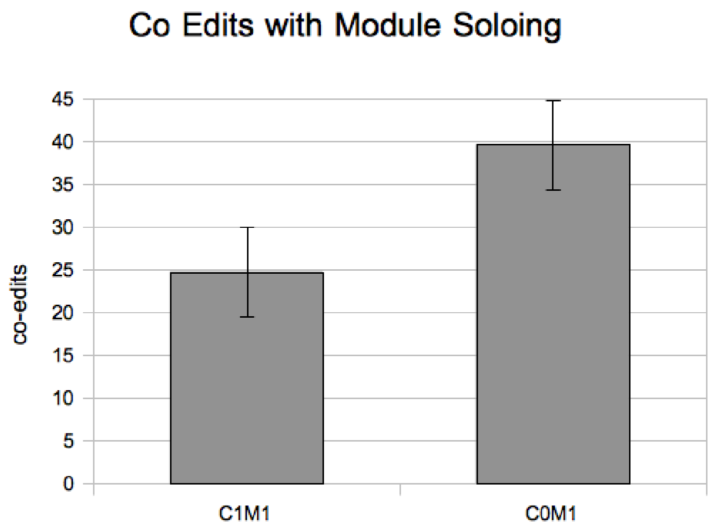



Figure 7.7: Mean co-edits compared between conditions C1M1 and C0M1. Error bars show Standard Error.

For groups where the interface did not include Module Soloing, participants made significantly less Deletions (w=–183, n/sr= 23, z=-2.09, p=0.02) when given the option of Client Soloing. (See Figure 7.8). There were no other significant effects.

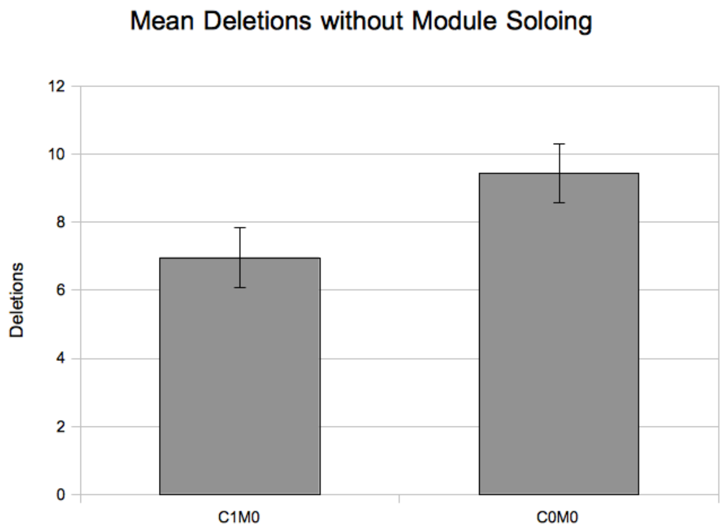



Figure 7.8: Mean Deletions compared between conditions C1M0 and C0M0. Error bars show Standard Error.

| Feature | C1 | C0 | W | n s/r | z | p |
|---|---|---|---|---|---|---|
| Creations | | | | | | |
| with Module Soloing | 219 | 210 | 20 | 22 | 0.32 | 0.3745 |
| without Module Soloing | 206 | 224 | -49 | 23 | -0.74 | 0.23 |
| Deletions | | | | | | |
| with Module Soloing | 83 | 110 | -30 | 19 | -0.59 | 0.2776 |
| without Module Soloing | 87 | 118 | -138 | 23 | -2.09 | **0.02** |
| OwnEdits | | | | | | |
| with Module Soloing | 3009 | 2716 | 41 | 24 | 0.58 | 0.2810 |
| without Module Soloing | 2972 | 2754 | 101 | 24 | 1.44 | 0.07 |
| coEdits | | | | | | |
| with Module Soloing | 310 | 496 | -87 | 20 | -1.61 | **0.0537** |
| without Module Soloing | 356 | 458 | -10 | 18 | -0.21 | 0.42 |
| Patch to Public | | | | | | |
| with Module Soloing | 311 | 312 | 3 | 22 | 0.04 | 0.4840 |
| without Module Soloing | 299 | 329 | -70 | 23 | -1.06 | 0.14 |
| Moves | | | | | | |
| with Module Soloing | 1109 | 1127 | 25 | 24 | 0.35 | 0.3632 |
| without Module Soloing | 1141 | 1009 | 86 | 23 | 1.3 | 0.1 |

Table 7.4: Effects of Client Soloing on interaction log features, comparing Client Soloing and no Client Soloing, using the Wilcoxon Signed-Rank test, for groups with and without Module Soloing

### 7.11.3 Effects of Module Soloing on Post-Test Questionnaire

The experimental design permitted an analysis of Module Soloing by means of a between-subjects comparison. In this approach, half the groups of participants were presented with an interface containing Module Soloing for both their conditions, and half the groups of participants were not presented with Module Soloing in either of their conditions. Both sets of participants (those with Module Soloing and those without) were also presented with two within-subjects conditions, an interface with Client Soloing and an interface without Client Soloing.

To perform the between-subjects comparison, The Mann-Whitney U test was used to analyse the post-test questionnaire responses across the M0C0 and M1C0 samples of data to test for the influence of Module Soloing where Client Soloing was not present within the interface. The Mann-Whitney U test was then applied across the M1C1 and M0C1 samples to test for the influence of Module Soloing where Client Soloing was also present within the interface.

The Mann-Whitney U test is a non-parametric statistical test for assessing whether one of two independent samples tends to have larger values than the other. The Mann-Whitney U can be regarded as the unrelated design equivalent to the Wilcoxon Signed-Rank (Greene and D'Oliveira, 1999), and is therefore an appropriate test in this instance.

Table 7.5 presents the results of this analysis. This analysis showed that when Client Soloing was present within the interface, independent sets of participants reported working most effectively (Ua=372, Z=-1.72, p=0.04), editing the music together (Ua=408, Z=-2.46, p=0.01), and being influenced by the other group members in the condition which did not include Module Soloing (Ua=372, Z-1.72, p=0.04).

When the interface did not feature Client Soloing, the presence of Module Soloing in the interface caused no significant effects on reported preferences. This suggests an interaction between Client Soloing and Module Soloing, where the presence of Client Soloing within the interface resulted in a stronger distinction between the reported effects of Module Soloing.

| Statement | M1 | M0 | Ua | Z | p |
|---|---|---|---|---|---|
| The best music | | | | | |
| with Client Soloing | 23.5 | 26 | 312 | -0.48 | 0.32 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| I felt most involved with the group | | | | | |
| with Client Soloing | 23.5 | 26 | 312 | -0.48 | 0.32 |
| without Client Soloing | 24.5 | 25 | 288 | 0.01 | 0.5 |

Continued on next page

Table 7.5 – continued from previous page

| Statement | M1 | M0 | Ua | Z | p |
|---|---|---|---|---|---|
| I enjoyed myself the most | | | | | |
| with Client Soloing | 23.5 | 26 | 312 | -0.48 | 0.32 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| I felt out of control | | | | | |
| with Client Soloing | 25.5 | 24 | 264 | 0.48 | 0.32 |
| without Client Soloing | 23 | 26 | 324 | -0.73 | 0.23 |
| I understood what was going on | | | | | |
| with Client Soloing | 23 | 26 | 324 | -0.73 | 0.23 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| I mostly worked on my own | | | | | |
| with Client Soloing | 25.5 | 24 | 264 | 0.48 | 0.32 |
| without Client Soloing | 22.5 | 27 | 336 | -0.98 | 0.16 |
| I explored more | | | | | |
| with Client Soloing | 24.5 | 25 | 288 | 0.01 | 0.5 |
| without Client Soloing | 24.5 | 25 | 288 | 0.01 | 0.5 |
| I lost track of time | | | | | |
| with Client Soloing | 21.5 | 28 | 360 | -1.47 | 0.07 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| Other people ignored my contributions | | | | | |
| with Client Soloing | 22.5 | 27 | 336 | -0.98 | 0.16 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| I was more focused | | | | | |
| with Client Soloing | 24.5 | 25 | 288 | 0.01 | 0.5 |
| without Client Soloing | 23.5 | 26 | 312 | -0.48 | 0.32 |
| The interface was most complex | | | | | |
| with Client Soloing | 22.5 | 27 | 336 | 0.98 | 0.16 |
| without Client Soloing | 26 | 23 | 252 | 0.73 | 0.23 |
| I had the most privacy | | | | | |
| with Client Soloing | 25.5 | 24 | 264 | 0.48 | 0.35 |
| without Client Soloing | 22 | 27 | 348 | -1.23 | 0.11 |
| We worked most effectively | | | | | |
| with Client Soloing | 21 | 28 | 372 | -1.72 | **0.04** |
| without Client Soloing | 26.5 | 23 | 240 | 0.98 | 0.16 |
| I knew what other people were doing | | | | | |
| with Client Soloing | 23 | 26 | 324 | -0.73 | 0.23 |
| without Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| We edited the music together | | | | | |
| with Client Soloing | 19.5 | 30 | 408 | -2.46 | **0.01** |
| without Client Soloing | 27 | 22 | 228 | 1.23 | 0.11 |

Continued on next page

Table 7.5 – continued from previous page

| Statement | M1 | M0 | Ua | Z | p |
|---|---|---|---|---|---|
| I made my best contributions | | | | | |
| with Client Soloing | 25 | 24 | 276 | 0.24 | 0.41 |
| without Client Soloing | 25.5 | 24 | 264 | 0.48 | 0.32 |
| I was influenced by the other people | | | | | |
| with Client Soloing | 21 | 28 | 372 | -1.72 | **0.04** |
| without Client Soloing | 25.5 | 24 | 264 | 0.48 | 0.32 |
| The condition I preferred the most | | | | | |
| with Client Soloing | 24 | 25 | 300 | -0.24 | 0.41 |
| without Client Soloing | 24.5 | 25 | 288 | 0.01 | 0.5 |

Table 7.5: Effects of Module Soloing with and without Client Soloing, using a between subjects Mann-Whitney U Test on post-test questionnaire responses

### 7.11.4 Effects of Module Soloing on Software Interaction

The experimental design called for an analysis of the effects of Module Soloing by means of a between-subjects comparison. As with the post-test questionnaire data described above, half the groups of participants were presented with an interface containing Module Soloing for both their conditions, and half the groups of participants were not presented with Module Soloing in either of their conditions. Both sets of participants (those with Module Soloing and those without) were also presented with two within-subjects conditions, an interface with Client Soloing and an interface without Client Soloing.

To perform the between-subjects comparison, The Mann-Whitney U test was used to analyse the interaction log data across the M0C0 and M1C0 samples to test for the influence of Module Soloing where Client Soloing was not present within the interface. The Mann-Whitney U test was then applied across the M1C1 and M0C1 samples to test for the influence of Module Soloing where Client Soloing was also present within the interface.

Module Soloing caused no statistically significant effects on the way the software was used for either set of groups, and on inspection, the figures were almost identical (See Figure 7.6), indicating that Module Soloing had a negligible effect on all aspects of the interaction logs or the way participants used the software.

| Feature | M1 | M0 | Ua | Z | p |
|---|---|---|---|---|---|
| Creations | | | | | |
| With Client Soloing | 25.9 | 23.1 | 254 | 0.69 | 0.2451 |
| Without Client Soloing | 23.8 | 25.2 | 305.5 | -0.35 | 0.3632 |
| Deletions | | | | | |
| With Client Soloing | 22.3 | 26.7 | 341.5 | -1.09 | 0.1379 |
| Without Client Soloing | 21.9 | 27.1 | 349.5 | -1.26 | 0.1038 |
| OwnEdits | | | | | |
| With Client Soloing | 24.5 | 24.5 | 288.5 | 0 | 0.5000 |
| Without Client Soloing | 24.5 | 24.5 | 288 | 0.01 | 0.4960 |
| coEdits | | | | | |
| With Client Soloing | 26 | 23 | 253 | 0.71 | 0.2389 |
| Without Client Soloing | 26.5 | 22.5 | 240.5 | 0.97 | 0.1660 |
| Patch to Public | | | | | |
| With Client Soloing | 25 | 24 | 275.5 | 0.25 | 0.4013 |
| Without Client Soloing | 24.3 | 24.8 | 294 | -0.11 | 0.4562 |
| Moves | | | | | |
| With Client Soloing | 25 | 24 | 275.5 | 0.25 | 0.4013 |
| Without Client Soloing | 25 | 24 | 277 | 0.22 | 0.4129 |

Table 7.6: Effects of Module Soloing on interaction log features by performing a Between Subjects comparison of conditions with Module Soloing and without Module Soloing, using the Mann-Whitney U Test. Table shows two sets of groups, those who were presented with Client Soloing and those who were not.

## 7.12 Thematic Analysis

Thematic Analysis (Braun and Clarke, 2006) was used to analyse topics of verbal exchange made by participants during their interaction with the software, and to analyse the interviews conducted at the end of each study session. This analysis was performed with the intention of gaining insight into how the participants discussed and conceptualised the software interface, their approaches to collaboration, the issues they faced while working together, and the conversations they had about the music they were creating. Thematic analysis provides an established and structured method for making sense of large bodies of textural data, and was therefore selected as an appropriate technique for this purpose. This section describes in detail the methodology employed in conducting the Thematic Analysis. This description is an essential criteria for reporting qualitative research as it provides a transparent account of steps taken, and enables other researchers to arrive at their own assessment of the work. Following a description of methodology, this section provides a lengthy discussion of the findings.

### 7.12.1 Thematic Analysis Methodology

Thematic Analysis is a qualitative analysis method which seeks to identify themes or patterns within a body of textual data (Braun and Clarke, 2006). This is performed through an iterative process of manually coding the data, grouping codes to identify common themes, and re-evaluating these themes against the data to assess their strength and applicability. While coding and comparison could potentially be a never-ending process (Braun and Clarke, 2006, Charmaz, 2006), there eventually comes a point at which the comparison between codes and data ceases to result in radical modifications to the previously identified themes. At this point the comparative process could be said to have produced a working set of themes which sufficiently 'fit' with the data.

Qualitative research commonly uses coding and comparative techniques, although in many cases the process and method employed are not sufficiently documented to provide external parties with enough information to effectively evaluate or assess the outcomes of the analysis (Braun and Clarke, 2006). For instance many research publications talk about a process of 'coding for themes' without documenting the coding process or talking in sufficient detail about where or how the themes were derived. The Thematic Analysis process presented by Braun and Clarke (2006) addresses this problem by providing a structured methodology for coding, grouping codes, and generating a report of the findings. Braun and Clarke (2006) stressed the role of researcher as taking an active position in the interpretation of data, as opposed to being presented as a passive channel through which themes and codes within the data are said to 'emerge', as though hidden or latent within the data itself. Braun and Clarke (2006) encouraged researchers to apply their knowledge, intuition and expertise in analysis the data, rather than denying it or attempting to suspend their informed judgements, as advocated in some branches of Grounded Theory (Glaser and Strauss, 1967)

Thematic Analysis is also proposed as a suitable alternative to 'lite' versions of more in-depth qualitative methods such as Grounded Theory and Discourse Analysis (Braun and Clarke, 2006). For instance within previous HCI and CSCW research Muller and Kogan (2010) identified examples of 'Grounded Theory Methods' which have been employed as a lightweight alternative to conducting fully fledged Grounded Theory research. In these Grounded Theory Methods (GTM) the emphasis is usually placed on comparative coding, categorisation and comparative analysis, for the purposes of analysing previously collected interview or textual data. However without adhering to the other requirements of a Grounded Theory method, this coding process may be poorly structured, and there is no standardised or widely acknowledged method of conducting analysis using 'Grounded Theory Methods'. This makes it difficult for a researcher to confidently apply a subset of the grounded theory techniques to their particular application. In contrast, the structured, and self-contained methodology for conducting high quality thematic analysis on a variety of data sources makes it an accessible qualitative tool.

Finally, unlike Grounded Theory approaches which attempt to generate theory in tandem with the data collection process, Thematic Analysis is intended to be suitable for analysing existing data, as it does not rely on specific collection techniques such as theoretical sampling. While the specialised data collection techniques of more involved methods might be powerful analytic tools, they can nevertheless demand an unspecified amount of time, continued access to participants, and sustained funding for an on-going research process.

The Thematic Analysis conducted for this study followed the detailed six stage process outlined by Braun and Clarke (2006):

**Familiarisation with data** Familiarisation with the data is an essential aspect of TA. The process of familiarisation usually begins with transcription, however the researcher is encouraged to read the whole data set several times before moving to the next stages. The familiarisation process can include note taking. Repeated reading of the data can be very time consuming.

**Generating Initial Codes** The initial coding phase entails the generation of codes based on interesting features within the data. Despite the time it takes, this should proceed systematically through the entire data set. Relevant data for each code should also be collated at this stage.

**Searching for Themes** Grouping codes into potential or candidate themes enables a higher level view of the data to be generated. Data extracts to support each theme should also be gathered at this stage.

**Reviewing Themes** Themes should be reviewed and compared against the entire data set and against the initial codes, to evaluate their strength and relevance. At this stage, a Thematic Map can be generated to aid in assessing how the themes relate to one another.

**Defining and Naming Themes** Clear names and definitions should be produced for each theme. These definitions should serve to delimit the themes from one another.

**Producing a report** Writing the report is a fundamental aspect of the TA method. At this point, the themes should be described and discussed in relation to examples from the data, and related to literature and the research questions of the study.

### 7.12.2 Data for Thematic Analysis

The Thematic Analysis was driven primarily by the dialog recorded during the participants' interactions together and with the software. It is important to stress that these exchange between participants occurred prior to the post-test questionnaires and interviews, so were less guided or led by the research objectives of the experiment. This matters, as a common problem for analysis such as this is producing results which simply re-phrase the questions or issues put to participants by the researcher. Focusing on the conversations participants had while interacting also offered an opportunity to study the interaction as it unfolded. Data collected during the post-test interviews was, however, incorporated into the analysis during the later stages to elaborate upon the thematic descriptions produced through analysis of the participant interactions.

Video editing software was used to divide footage of the study sessions into forty-eight separate files, one for each condition and one file for each interview (2 conditions x 16 interaction sessions + 1 x 16 interviews). Although each experiment condition was presented for exactly twenty minutes (timed automatically by the software) the video footage was cut to include participants' discussions before and after the conditions. The average length of each interaction session video file was twenty-one minutes and thirty-two seconds, and footage equated to a total of eleven hours, twenty-nine minutes and twenty three seconds. The interviews consisted of four hours twenty-two minutes and fifty eight seconds of video footage. The videoed interviews ranged from seven minutes twenty-nine seconds to twenty-six minutes and eight seconds in length. The average interview time was seventeen minutes and thirty-one seconds.

All the footage was transcribed manually using the ELAN video annotation software[2]. The transcription was a verbatim account of all verbal utterances, but did not adhere to more complex transcription techniques such as those used in Discourse or Conversation Analysis. Back-channeling[3] was transcribed where possible, and non-speech utterances were approximated in the transcription where appropriate. Non verbal communication such as hand gestures was not coded

[2]http://www.lat-mpi.eu/tools/elan/

[3]In verbal communication, back-channelling refers to the the utterances made by an interlocutor who is not the current speaker. Examples of back-channel utterances in the English language include sounds such as 'uh-huh' and 'hmm', or words such as 'yes' or 'right'.

or transcribed in any way. This level of transcription was deemed appropriate as the analysis sought to identify themes and patterns in the topics of verbal exchange between participants during their interaction. In addition to the transcriptions, the logged text-chat messages exchanged by participants via the software interface were included in the data set to provide a more complete set of all communication during the interaction. Although the interviews remained for the most part on-topic, sections of dialogue which fell outside the remit of the research study were not transcribed. For instance one participant talked at length about his/her life history and anecdotal experiences working in the music industry.

The interview transcripts were divided into two sets, based on the experimental conditions the participants had been presented with. These are referred to as the M0 and M1 transcriptions, in reference to Module Soloing, which was presented to half the participant groups. The transcripts and text-chat logs were imported into a spreadsheet for the purpose of coding and analysis. Combined, the interaction session transcriptions and text-chat accumulated in approximately 2000 lines of dialogue, ranging from single words or non-speech sounds to long sentences. In total, there were approximately ten thousand words of transcribed speech in the interaction data, while the interview transcripts contained over five thousand lines of text, and in excess of forty-thousand two-hundred individual words.

### 7.12.3 Coding

Familiarisation with the data forms the first step in a Thematic Analysis. As noted by Braun and Clarke (2006), the process of transcription should engender a deep familiarity with the data, although to ensure familiarity, the entire corpus of dialogue was read through several times before coding. An initial coding process was performed by reading each line of data and textually coding it with a word of phrase which captured the meaning or significance of the data item. This generated an extensive list of codes which loosely described the topics of conversation and other instances of dialog during the interaction, including participants vocalising melodies, talking to themselves or asking the researcher questions. Table 7.7 shows an example of coded transcription from the initial coding stage of the analysis. At this stage in the process new codes were created as required to describe each line of data. Consequentially, as the right column in table 7.7 shows, the initial codes ranged from very generalised concepts such as ‘Approach’ to more specific codes such as ‘discussion of tempo’. A full list of the initial codes can be found in Appendix D. Coding the dialog during interaction generated fifty eight codes (see Appendix

| Participant | Transcription | Initial Codes |
|---|---|---|
| Participant 3 | what now, genre? | |
| Participant 2 | err, lets try Trance | discussion of style/genre, approach, planning |
| Participant 3 | trance, ok | discussion of style/genre |
| Participant 1 | trance | discussion of style/genre |
| Participant 1 | four to the floor | discussion of style/genre |
| Participant 2 | yeah, hundred and thirty or something (laughs) | discussion of tempo, modification |
| Participant 3 | OK | |
| Participant 3 | same, with regards, to drums, synths and chords? stuff like that roles | approach, planning |
| Participant 3 | umm, do same? | |
| Participant 1 | I'm just gonna do bits and piece, like that piece is I was just putting in little bits of drums, and little bits of bass, just little, flourishes, do you know what I mean? | roles, statement of intent, asserting |
| Participant 2 | cool | |
| Participant 1 | cos erm, my head's not, into, I haven't worked out the, the actual notes on the bass, the step sequencer, it'll take me a little while to sort it out | problems with the software, software, constraints |
| Participant 2 | cool | |
| Participant 3 | ok | |
| Participant 1 | and I don't want it to sound shit, so | problems with software, statement of intent, asserting |
| Participant 3 | (laughs) | |
| Participant 2 | I'll do, err, well, I'll do the drums | statement of intent, roles, planning |

Table 7.7: Example of initial coding for thematic analysis of participants during interaction

D.3. Coding the M0 condition interviews generated 566 codes (see Appendix D.3), and coding the M1 condition interviews generated 787 codes (see Appendix D.3).

At this stage the two sets of interview transcripts were initially coded by reading each line of transcript text and manually adding a code to describe the significance or meaning of the statement made by participants. There codes were then collated into loosely defined categories and visually grouped, resulting in two densely populated maps, shown in Figures 7.9 and 7.10.

Figure 7.9: Initial Codes for M0 Condition Interviews. Digital copy available on request.


Figure 7.10: Initial Codes for M1 Condition Interviews. Digital copy available on request.

### 7.12.4 Candidate Themes

The initial codes were then grouped and organised into a set of *candidate themes*, which attempted to loosely capture similarities between the codes and form higher level categories. At this stage it was assumed that further coding iterations and comparison would radically alter the initial candidate themes. The candidate themes were then organised into a candidate Thematic Map to visually group similar themes and provide an overview of how different themes and codes related to one another.

Figure 7.11, shows the first generation of this process. At this stage, some of the more general codes were elevated to form higher level groupings of similar codes. The spatial layout of the themes in this figure is arbitrary and does not represent any form of high-level groupings or categorisation. Such groupings had not taken place at this stage in the analysis. This figure therefore represents an initial attempt to identity the themes present in the data. Examples from this stage included:

**Identification** : Discussion over the identification of sounds, instruments, authors or activities within the collaboration. For example where a participant asked the other group members who was making a certain sound.

**Modification** : Discussion over making changes to the music. For instance the Tempo control was a common source of modification related discussion.

**Ownership** : The ownership candidate theme gathered together instances of participants referring to the ownership of specific contributions. For example where a participant talked about ‘my synth’, or ‘your bass’.

**Assistance** : The assistance candidate theme collected together codes related to participants asking for or providing assistance in the use of the software. This candidate theme included codes attributed to discussion over how certain features worked, as well as instances where participants described how they had created certain sounds.

*Revision to the Candidate Thematic Map*

The candidate thematic map was revised through iteration by means of comparison with the data set and previous coding. This process entailed attempting to apply the candidate themes to each line of data to assess how well they ‘fitted’ the data , and identifying instances where the themes


Figure 7.11: First generation of Candidate Themes

and codes were not sufficiently nuanced to categorise, describe or capture the intricacies of the data. These instances of mismatch between themes and data resulted in refinement to existing themes and the generation of new themes based on pre-existing or newly formed codes. This process was repeated several times, and each time resulted in iterative modifications to the candidate thematic map. For instance, comparing the first and second generation candidate thematic maps presented in Figures 7.11 and 7.12, the second generation introduced new themes which placed more emphasis on ideas such as ‘Reflection’ or ‘Feedback’, and several themes were also re-named, for instance changing ‘Limitations’ to ‘Constraints’. These subtleties illustrate the effect the comparative process had on shaping and developing the themes.

### 7.12.5 Thematic Map

Once the iterative coding and comparison process ceased to yield substantial changes to the codes or themes previously applied to the data, a finalised thematic map was produced (Figure 7.13). The final Thematic Map contains three high level themes of ‘Identification’, ‘Software’ and ‘Working Together’. Within these main themes are a number of more nuanced sub-themes which describe individual aspects of the discussions and exchanges between the participants. For

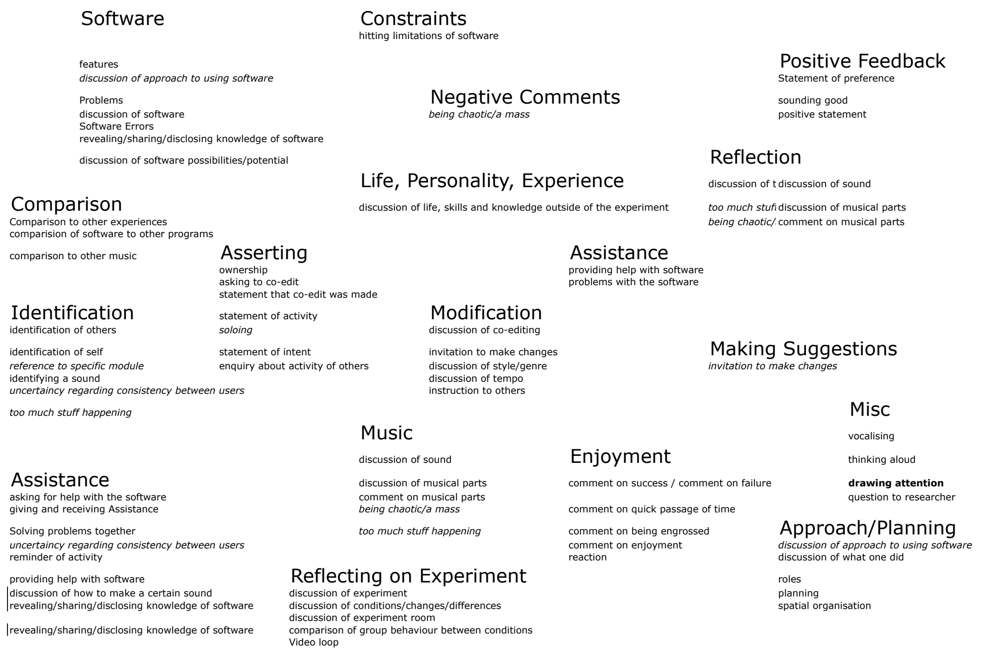



Figure 7.12: Refined second generation of Candidate Themes

instance the thematic map groups themes such as 'Problems', 'Constraints' and 'Features' under the wider category of 'Software', while discussion about changes to the music ('Modification') and 'Positive Feedback' fall under the theme of 'Working Together'. In addition, a number of commonly recurring extraneous themes were placed into an 'Other' category. The following sections draw on extracts from the data to describe, demonstrate and support the themes contained within in the Thematic Map[4]. The 'Other' category is excluded from this discussion, due to irrelevance.

### 7.12.6 Software Theme

The software theme relates to discussion over the software itself, rather than practical aspects of the music-process (These are addressed in the Working Together theme). The Software theme captures participants experiences using the software, as well as highlighting the kinds of problems which were encountered, and participants' thoughts about how the software might be improved. The following sub-themes were identified within the Software category: *Features, Use, Constraints, Comparison, Problems, Assistance*.

[4]Thematic maps are zoomable in digital form of this thesis, or can be downloaded in .pdf format from http://www.eecs.qmul.ac.uk/~robinfencott/thesis_resources/


Figure 7.13: Final Thematic Map

*Software: Features*

The Features theme captures discussion about specific features within the software, and participant's statements about them. It also applies to discussion about potential or suggested features and expressed views on possible changes that could be made to the software. Some participants discussed positive features in the software design, particularly regarding the audio quality and available musical/sonic resources.

PARTICIPANT 2: there's some quite good bass synths in that
(Condition M0C1, Session 9)

PARTICIPANT 3: ah, I kind of can't believe it took me so long to find the filter
PARTICIPANT 2: (laughs)
PARTICIPANT 3: it makes everything so much more interesting
PARTICIPANT 2: yeah yeah, it's good (Condition M0C1 Session 11)

However, it was more common for participants to make critical statements about the software and discuss what they identified as shortcomings or failings in the design. In some cases participants suggested ways in which certain features could be improved to streamline the music-making process. Some of these exchanges were related to music-making functionality, for instance modifications to the instruments, interaction design and mixing capabilities.

PARTICIPANT 1: it's ok though isn't it
PARTICIPANT 2: yeah
PARTICIPANT 1: well there's lots of, yeah few things I'd like to, I'd, I'd like to just have a, spend just a little time, so I could get used to the err, step sequencer
PARTICIPANT 3: yeah
PARTICIPANT 1: not the drum one, I know how to use the drum one, the er, the notes, just so I can, I hate putting in wrong notes, I fucking hate it you know, I don't mind when I do it on my own
PARTICIPANT 2: you know the problem here, it's too small, the actual, the increments
PARTICIPANT 1: the increments, I know, it's really, you're going really too, far I know,
PARTICIPANT 2: yeah
PARTICIPANT 1: I mean, yeah, if I was on my own, I'd just go through, it'd take me an hour and I'd go through and I'd find the, you know, the, I'd like a controller as well
PARTICIPANT 1: and also, the drop down menu where the drums the drums are
PARTICIPANT 2: yeah
PARTICIPANT 1: it should be just click, you have to go into it every time, you couldn't just like, you know if you want to audition sounds as they come
PARTICIPANT 2: yeah
PARTICIPANT 1: but I just keep on going, just changing things really fast
PARTICIPANT 2: yeah yeah
PARTICIPANT 1: just couldn't do it fast enough, that'd be really easy to fix
PARTICIPANT 1: yeah
PARTICIPANT 2: I think it's pretty cool (Condition M0C1, Session 16)

In this extract a participant raised several issues relating to the interface design for specific features of the software; that it is difficult to find the correct note using the sliders for the Step Synth, and that selecting and auditioning the drum samples in the Drum Sequencer could be made easier by reducing the amount of times the user needs to click on the menu bar. Similarly in the following extract participants suggest a mixer feature and note that a longer sequence would be beneficial.

PARTICIPANT 2: you need a sort of mixer or something
PARTICIPANT 3: you've gotta
PARTICIPANT 3: yeah
PARTICIPANT 2: and plus I would say, longer sequence
PARTICIPANT 3: yeah (Condition M0C1, Session 16)

*Software: Use*

The Software Use theme categorises discussion between participants related to how they used the software. This theme differs from Features as it does not necessarily relate to specific features or aspects of the software. Some participants reflected on their approach to interacting with it, for instance considering the collaborative environment as an instrument, or discussing the experience of encountering the software for the first time. The following extracts show how some participants talked about the software in a more general sense, for instance stating that the software could be used an 'instrument', while another group reflected on encountering the software, and how the group worked together.

PARTICIPANT 1: we can use it as a instrument
PARTICIPANT 2: yeah (Condition M1C1, Session 8)

PARTICIPANT 3: but it's true that, you, we're learning a new tool, and
PARTICIPANT 1: mm
PARTICIPANT 2: mmm
PARTICIPANT 3: and we're, it, we're, but, I know, that, we're not proficient, err, enough to know to, how to cope with a progression and then kuu!
PARTICIPANT 2: yeah
PARTICIPANT 3: but still, it, it's, I just wanted to push our limits, to see,
PARTICIPANT 2: yeah
PARTICIPANT 3: to the max, we, what we can do, up, up there
PARTICIPANT 2: yeah
PARTICIPANT 3: but, I think we
PARTICIPANT 2: yeah, yeah, it was very rulesy, it was interesting cos like yeah, I, yeah, I not only did we settle into a groove, but we started to do some like on the fly like, mixing as it were, as well
PARTICIPANT 1: mm
PARTICIPANT 2: which was, which was really interesting
PARTICIPANT 3: hmm
PARTICIPANT 3: and, I don't know if it, because it was, we knew which person did what, I think it's, it's good to know, umm, what colour, and you know what, where we're at because
PARTICIPANT 2: yeah
PARTICIPANT 1: yeah (Condition M0C1, Session 14)

*Software: Constraints*

Participants discussed the constraints imposed by the software. Commonly discussed constraints included the four beats to the bar time signature and the limited length of the musical loop, which participants often expressed as imposing a limitation on their ability to realise certain musical contributions.

PARTICIPANT 1: I find it quite tricky as well with, the erm, you know, trapped inside the four-four
PARTICIPANT 3: yeah
PARTICIPANT 1: can't get anything in between
PARTICIPANT 3: yeah (Condition M0C0, Session 15)

Some participants discussed the limitations of the sounds and sonic pallet offered by the software.

PARTICIPANT 1: there's no kind of like a pad sound, or, you know, chords,
PARTICIPANT 2: yeah, that's true, it's it's all like midi sounds
PARTICIPANT 1: that's, that's the problem I have (Condition M0C1, Session 13)

*Software: Comparison*

Some participants made comparisons between the study software and other music technologies. These comparison where frequently related to the graphical patching metaphor of the software, which shares characteristics with popular computer music environments such as Max/MSP [5] and Native Instruments Reaktor [6].

PARTICIPANT 1: I don't know if you've ever worked with Reactor, it's very similar to that kind of thing
PARTICIPANT 2: not a
PARTICIPANT 1: but it's obviously not so
PARTICIPANT 2: I've, I've used err, Max/MSP
PARTICIPANT 1: is that similar, is it?
PARTICIPANT 2: yeah that, that's all kind of connecting synths, to output, inputs and
PARTICIPANT 1: hmm
PARTICIPANT 2: samplers and that
PARTICIPANT 1: have, you've all used Reason?
PARTICIPANT 1: no?
PARTICIPANT 2: long time ago
PARTICIPANT 2: I just use Logic and Protools
(Condition M0C0, Session 13)

*Software: Problems*

This theme describes situations where participants discussed problems they encountered with the software, such as bugs, perceived bugs, misunderstandings about how aspects worked, or confusion over the shared nature of the interface. Within the category, no distinction is made between genuine problems (such as a lack of functionality, software bugs) and perceived problems (for

[5] http://cycling74.com/products/max/ accessed 6 March 2012
[6] http://www.native-instruments.com/#/en/products/producer/reaktor-55/ accessed 6 March 2012

instance where participants misinterpreted an aspect of the software, or encountered an error due to a misunderstanding about how something was designed to work).

This Theme differs from the Features Theme as some problems are not related explicitly to particular features of the software, but may entail more general issues. For instance the following extract describes a situation in which the participants believed the software was malfunctioning (not producing sound), however in reality the participants had not entered any musical sequences or drum patterns.

PARTICIPANT 3: do you have to connect it?
PARTICIPANT 2: oh
PARTICIPANT 2: there's still no sound
PARTICIPANT 1: what is there no sound?
PARTICIPANT 2: we have no sound
RESEARCHER: oh really?
RESEARCHER: um
RESEARCHER: put some things in here?
PARTICIPANT 2: oh, fine
PARTICIPANT 2: stupid
PARTICIPANT 1: what happened?
RESEARCHER: have you all got sound?
PARTICIPANT 3: yeah
PARTICIPANT 2: we didn't put in tick boxes ( Condition M1C0, Session 3)

In the previous example the researcher was called by the participants to assist in their trouble, however in many cases the participants discussed or dealt with problems by themselves:

PARTICIPANT 2: I think we overloaded it, like
PARTICIPANT 1: no
PARTICIPANT 2: I think someti, well sometimes I feel it
PARTICIPANT 3: it sounded like it was glitching
PARTICIPANT 1: yeah
PARTICIPANT 2: yes (Condition M1C0, Session 4)

*Software: Assistance*

A strong theme was for participants to provide assistance to one another during the interaction. This is similar to the Technology Confusion category identified in Olson et al. (1993). In the majority of cases assistance was related to the use of specific features within the software or the ways in which certain parts of the interface functioned. In some cases participants assisted others when problems arose:

PARTICIPANT 3: I couldn't make much distortion, I tried a few times and erm, I wasn't hearing

PARTICIPANT 3: I think I need to, there is the window, and you to click this one to the one I want to distort, and then this one to the middle?
PARTICIPANT 1: you need to connect the sound to the distortion, then the distortion to the main sound
PARTICIPANT 3: yeah so this is the, this is the instrument I want to distort, this is the distortion, I need to connect this to this, and then this to this
PARTICIPANT 1: yes
PARTICIPANT 3: yeah, I did that, and it was just not, yeah, OK
(Condition M0C1, Session 9)

While in other cases participants offered help and shared their knowledge of how the software worked without there being a specific problem at hand:

PARTICIPANT 2: I'll have a go at a melody
PARTICIPANT 1: I've found for fine tuning, you can use the arrows, like if you hold the thing, the cursor over the circle, tell you what number it is out of a hundred, and can use the arrows to move it
PARTICIPANT 2: really?
PARTICIPANT 1: one at a time
PARTICIPANT 2: how do you tell that?
PARTICIPANT 1: if you set your cursor over the little circles, under the step synth,
PARTICIPANT 2: yeah
PARTICIPANT 1: for each of the beats, and just leave it there for a second it should tell you what number it's at
PARTICIPANT 2: OK (Condition M1C9, Session 17)

In one instance, a participant had produced an interesting sound using a complex arrangement of modules, and when questioned by one of the other participants, provided a detailed description of how this had been achieved.

PARTICIPANT 1: how did you get that swoosh?
PARTICIPANT 3: sorry?
PARTICIPANT 1: hooweewwowww
PARTICIPANT 3: sorry?
PARTICIPANT 1: the swoosh, that's really good?
PARTICIPANT 3: swoosh thing?
PARTICIPANT 3: that is
PARTICIPANT 3: sounds like The Tardis
PARTICIPANT 1: yeah
PARTICIPANT 3: erm, basically, I got a cymbal, sound
PARTICIPANT 1: uh uh
PARTICIPANT 3: and just put it though a reverb, ne, m, put everything on the reverb to max so it sounds like it's at the very far end of a cathedral, then filtered it so it goes hwwweeeoooww
PARTICIPANT 1: uh hu
PARTICIPANT 1: really?
PARTICIPANT 1: that's amazing, that's a really cool sound

PARTICIPANT 3: basic, I was basically trying to get something that, was as close to white noise as I could get and then filter it
PARTICIPANT 1: yeah yeah
PARTICIPANT 1: very good, that's really, I really like that
PARTICIPANT 3: there we go
PARTICIPANT 1: mmm?
PARTICIPANT 3: I found the right cymbal for that now (Condition M1C1, Session 5)

### 7.12.7 Identification Theme

Identification was a key theme within the discussion between participants during the interaction, and pertains to the way participants ascertained who was responsible for particular musical contributions and the ways in which they talked about the source of particular sounds within the mix. This theme was partially informed by the research questions and aims of the study (see Section 7.2), however within the data, discussion over identification and awareness formed a large proportion of all dialog between participants. Identification was divided into three sub themes: *Authorship, Sounds, and Self.*

*Authorship*

The authorship theme describes discussion by participants about who was responsible for which sounds, ideas or contributions with the music:

PARTICIPANT 2: which one are you working on now? the, out of the, out of the three,which one are you working on, the top, the middle or the bottom?
PARTICIPANT 3: those, those three on the left
PARTICIPANT 2: oh right (M0C0 Session 16)

PARTICIPANT 1: how did you get that swoosh?
PARTICIPANT 3: sorry?
PARTICIPANT 1: hooweewwowww
PARTICIPANT 3: sorry? (Condition M1C1, Session 5)

*Sounds*

Discussion or questions about the source of a sound were frequently identified. For instance some participants used vocal mimicry to aid in establishing reference to specific sounds:

PARTICIPANT 3: what is this, what is the grrrrrrr!, what is that, who is that? who is doing that? (Condition M0C1, Session 9)

In other cases, participants talked about using the interface features to discover the source of sounds, for instance re-appropriating the Client Soloing feature as a means of narrowing the range of sounds in the mix:

PARTICIPANT 3: I think the, I only ever used it just to basically, s, solo like like like, mute a whole load if, when it was getting pretty hectic
RESEARCHER: yeah
PARTICIPANT 3: like just like, just use it to just, regardless of who's I was muting I would, I was like well I'll mute that one and that just cuts all that shit out
RESEARCHER: yeah yeah yeah
PARTICIPANT 3: so it was just more of a, kind of like a, a mute of stuff that you don't want (Session 11 Interview)

*Self*

Discussion about identification of one's self within the mix:

PARTICIPANT 2: there's a bit of confusion with the synths though
PARTICIPANT 1: yeah, it's
PARTICIPANT 2: cos you don't know which notes are playing
(Condition M1C0, Session 8)

### 7.12.8 Working Together Theme

Working Together captures themes related to how the participants went about creating music together. The data coded with the Working Together theme relates to suggested modifications, statements about the one's activity, and reflection on the progress made. The following sub-themes were identified within the 'Working Together' category *Suggestions, Modification, Comparison, Planning, Reflection, Positive Feedback*.

*Suggestions*

Instances of participants suggesting changes to the music

PARTICIPANT 1: it's very dark, should we lighten it up a little?
(Condition M0C0, Session 15)

PARTICIPANT 2: why don't you, do a, try and make a, like a sine curve?
PARTICIPANT 1: again?
PARTICIPANT 1: oh OK (Condition M1C0, Session 3)

*Modification*

A key sub-theme within the Working Together theme group was that of Modification. It was common for participants to discuss the tempo of the music, which was a global control that applied uniformly to all participants using the software.

PARTICIPANT 1: do you guys mind if I see how it sounds at one forty?
PARTICIPANT 2: (laughs), no
PARTICIPANT 1: mind if I change the tempo for a bit, is that?

PARTICIPANT 3: maybe in the middle something (Condition M1C1, Session 17)

It was also common to see participants announce modifications they were performing to other people's work.

PARTICIPANT 1: I'm gonna move the kicks up to the first beat
PARTICIPANT 2: mmm?
PARTICIPANT 1: I'm gonna move the kicks in the top right corner up to the first beat
PARTICIPANT 2: yeah, you can
PARTICIPANT 1: cos it's really confusing
PARTICIPANT 3: yeah
PARTICIPANT 1: it's, it's, it's all on the
PARTICIPANT 3: it's yeah, it's making the first beat off-beat
(Condition M0C1, Session 12)

The Modification theme was also applied to instances of participants discussion the ability to edit modules created by others.

PARTICIPANT 1: I suppose you can change other people's bits, seems a bit rude
PARTICIPANT 3: yeah
PARTICIPANT 2: (laughs)
PARTICIPANT 1: but I was just thinking yeah, hang on
PARTICIPANT 2: yeah
PARTICIPANT 2: um yeah, yeah yeah, no no, I know,
PARTICIPANT 1: did you? (laughs)
PARTICIPANT 2: I noticed, well, I I wasn't doing too much of it, but ab
PARTICIPANT 2: I noticed like a couple of people stepped in and tweaked bits of mine, and that, that was perfectly fine, though, I was,
PARTICIPANT 3: yeah
PARTICIPANT 2: well especially, since we didn't have the opportunity to solo I was just listening to the whole thing and just like going along with it, so
PARTICIPANT 1: yeah yeah yeah
PARTICIPANT 3: yeah
PARTICIPANT 1: cool, but yeah, tweak mine if
PARTICIPANT 3: OK
PARTICIPANT 1: (laughs) that's fine (Condition M0C1, Session 14)

Some participants discussed the introduction of new elements

PARTICIPANT T: what's better...this...
PARTICIPANT T: or this
PARTICIPANT L: 1st one
PARTICIPANT T: the one that's on now then?
PARTICIPANT T: not a fan of the piano sound (Condition M0C1, Session 15)

Others noted the deletion of modules, and acknowledged the accidental modifications they made:

PARTICIPANT 2: I deleted the synth by accident
PARTICIPANT 3: ahh (laughs) (Condition M0C1, Session 16)

PARTICIPANT 1: that, that one just annoying tone, the the one that goes duun duun (clicks fingers)
PARTICIPANT 1: where is that one? did it disappear? did it disappear?
PARTICIPANT 3: yeah it disappeared
PARTICIPANT 2: the one that was down there?
PARTICIPANT 2: yeah I got rid of it
PARTICIPANT 1: you got rid of it? why did you get rid of it?
PARTICIPANT 2: because it was just so annoying (Condition M1C0, Session 3)

*Comparison to other music*

The comparison theme captures incidents where participants drew on their existing musical knowledge and referred to their musical background to describe the music they were creating. This Theme shows how participants used musical genres, styles and cliches as a means of finding common ground with the group members and aiding in discussion of the collaboration.

PARTICIPANT 3: some death metal drumming going on there
(Condition M1C0, Session 5)

PARTICIPANT 1: bit, bit of Pink Floyd (laughs)
PARTICIPANT 2: (laughs) yeah (Condition M0C1, Session 13)

PARTICIPANT 3: think we managed to cover most music genres in that
PARTICIPANT 2: ah?
PARTICIPANT 3: think we managed to cover most of the electronic music genres in that one
PARTICIPANT 2: yeah, hu hu (laughs) (Condition M0C1, Session 11)

Some groups discussed the musical style they would adhere to, often with reference to specific sub-genres of electronic music such as garage, techno, and trance.

PARTICIPANT 2: go, go for something a bit happier now?
PARTICIPANT 3: yeah
PARTICIPANT 2: it's like dark industrial techno otherwise (M0C1 Session 13)

These discussions in some instances lead to discussion over musical parameters such as tempo.

PARTICIPANT 3: what now, genre?
PARTICIPANT 2: err, lets try Trance
PARTICIPANT 3: trance, ok

PARTICIPANT 1: trance
PARTICIPANT 1: four to the floor
PARTICIPANT 2: yeah, hundred and thirty or something (laughs)
PARTICIPANT 3: OK ( Condition M0C1, Session 16)

*Planning*

Discussions of planning were thematically grouped to describe how the participants approached spatial organisation, musical roles, and the musical direction of the piece. The majority of groups did not engage in extensive planning, although some groups attempted, at least initially, to impose structure onto their interaction through the delegation of roles and the allocation of shared screen space between each other.

Some groups of participants attempted to assign musical roles to members,

PARTICIPANT 2: shall we, gonna say, divide, who's gonna do, what instrument, for example?
PARTICIPANT 1: is it, it's just a drum and a synth isn't it?
PARTICIPANT 3: yeah thats
PARTICIPANT 2: yeah yeah
PARTICIPANT 1: yeah
PARTICIPANT 2: but I mean
PARTICIPANT 3: but then, but then like one person can do bass,
PARTICIPANT 2: yeah
PARTICIPANT 3: for example, and the other person can do like
PARTICIPANT 1: lead or something
PARTICIPANT 3: lead or chords, we can even like, I can try, if you guys want, I can try to do like, the middle range
PARTICIPANT 1: OK
PARTICIPANT 3: and try to chords in the middle range
PARTICIPANT 1: Ok
PARTICIPANT 3: and somebody do the bass, I can like, match in
PARTICIPANT 1: OK
PARTICIPANT 3: the bass
PARTICIPANT 2: cool
PARTICIPANT 3: so, either one, of you, want to do drums, or
PARTICIPANT 2: cool,. I'll do it
PARTICIPANT 1: which do you want to do?
PARTICIPANT 2: I'll do, err, drums then
PARTICIPANT 1: ok, yeah
PARTICIPANT 2: yeah
PARTICIPANT 3: so if you do like a bass, yeah
PARTICIPANT 1: I'll try bass (Condition M0C0, Session 16)

Others delimited the shared screen space to help enforce these musical roles.

PARTICIPANT 2: I'll take the, the top left corner for now
PARTICIPANT 3: (laughs), ok that's, that's a good idea

PARTICIPANT 2: yeah
PARTICIPANT 1: so that's, the
PARTICIPANT 3: top left, OK, close to the video
PARTICIPANT 1: left yeah
PARTICIPANT 1: I'll take, I'll take
PARTICIPANT 3: I'm taking
PARTICIPANT 3: a, the
PARTICIPANT 3: I just created the step synth
PARTICIPANT 2: yeah
PARTICIPANT 3: down on the bottom left
PARTICIPANT 1: oh right so,
PARTICIPANT 3: thta's going to be my area, lets say
PARTICIPANT 1: top, top right
PARTICIPANT 2: yeah Ok, cool
PARTICIPANT 1: right, I'm just gonna probably have one bass anyway so
PARTICIPANT 3: unless you want to put effects on the right, like on the, the remaining corner, the, the bottom right (Condition M0C0, Session 16)

*Reflection*

The reflection theme responds to participants' reflection on the music they were making. In many cases this formed a criticism of specific sounds, or a critique of the music as a whole.

PARTICIPANT 2: it's just the end that doesn't sound right
(Condition M1C0, Session 3)

PARTICIPANT 2: huh, there's so many stuff going on
PARTICIPANT 1: huh?
PARTICIPANT 2: there's so much stuff going on
PARTICIPANT 1: yeah, you can't
PARTICIPANT 1: can't, chaotic (Condition M1C0, Session 8)

PARTICIPANT 1: chaos (Condition M1C0, Session 8)

However in some cases there was a more of a positive emphasis on how well the group had performed.

PARTICIPANT 3: we did it all together, our first run, I was, just had my own sound
PARTICIPANT 1: mmm
PARTICIPANT 3: but then when we had it all together I didn't do that once, it was all about making the song, right? (Condition M0C1, Session 10)

*Positive Feedback*

As well as the more negative critique, participants frequently made positive statements made about the music, or about specific features within it. Making positive statements was a common use for the text chat feature.

PARTICIPANT S: sounds cool! (Condition M1C1 Session 2 text-chat)

PARTICIPANT 3: there was some nice, umm, bits
PARTICIPANT 1: nice, yeah (Condition M1C0 Session 6)

PARTICIPANT 2: the beat is nice now
PARTICIPANT 1: hmm?
PARTICIPANT 2: it's a better beat (Condition M1C0, Session 8)

## 7.13 Discussion

This study investigated how interface mechanisms for authorship awareness and contribution awareness impacted on the process of collaborative digital musical interaction. This section begins by addressing the research questions and hypotheses presented at the start of this chapter. A more general discussion follows, drawing on the Thematic Analysis to highlight a number of additional issues raised by the study. The chapter concludes with a number of design guidelines and implications for future collaborative systems.

### 7.13.1 Hypothesis H1: Source Awareness

The concept of Source awareness refers to awareness of where sounds are coming from. Hypothesis H1 posited that an interface which incorporated additional Source awareness would increase the tendency for participants to edit each others' contributions (co-editing), have a positive impact on reported preferences, and increase participants' reported levels of awareness of one another. Source Awareness was supported within the interface through the Module Soloing function, which allowed participants to aurally solo and visually highlight specific musical contributions. The Module Soloing function did not however provide any indication of *who* had created the modules; this information was presented to some groups via the Client Soloing feature explored by Hypothesis H2. The experimental design investigated the effects of Source awareness using a between-subjects comparison.

Analysis of the interaction logs showed that incorporating Module Soloing into the interface had no statistically significant effects on the way participants used the software. This was true for the groups of participants who were presented with Client Soloing and the groups who were not presented with Client Soloing. This evidence refutes the hypothesis that increased Source awareness information would increase the tendency for participants to perform co-editing. Furthermore, in groups for whom the interface did not feature Client Soloing, the inclusion of Module Soloing had no significant effects on the self report data collected by the post-test questionnaire. This implies that where the interface did not feature Client Soloing, the inclusion of Module Soloing made no difference to how participants used the software, or to how they rated the group interaction, the music they made or their preferences towards the interface.

When Client Soloing was present within the interfaces, a between subjects comparison of the post-test questionnaire data comparing Module Soloing revealed that participants reported working *most effectively*, *editing the music together*, and being *influenced by the other group members* when the interface *did not* feature Module Soloing. This is an unexpected finding, and provides evidence that for participants who had access to the Client Soloing feature, the addition of the Module Soloing feature had a detrimental effect on the participants ability to work together. This finding is surprising as it challenges the assumption that increased awareness is always beneficial.

The results of these two comparisons have several implications. Firstly, that the Module Soloing function by itself had little effect on the way participants worked, or the way they felt about the interaction. Secondly, that when presented in tandem with the Client Soloing feature, the inclusion of Module Soloing made certain aspects of collaborative more difficult. This evidence entirely refutes hypothesis H1, and suggests either that Source Awareness is not a useful form of awareness within CDMI, or that the Module Soloing feature was not suitably designed to provide Source Awareness to participants in a useful fashion.

Thematic Analysis of group interviews and discussion during the interaction suggests that inadequate interface design was a more likely reason the Module Soloing function was unsuccessful. In particular, turning to the Thematic Analysis, the *identification of sounds* theme indicates that participants were concerned with discovering the source of sounds within the mix. This implies that although Source Awareness was important and useful during the interaction, the Module Soloing function may not have been adequately designed to support the acquisition of this information

Furthermore, during the interviews, some participants talked about an alternative use for the Client Soloing feature as a means of gathering Source Awareness. Participants who talked about this appropriation of the Client Soloing function noted being less concerned with who was initially responsible the creation of certain contributions, and therefore ignoring the Authorship Awareness provided by the function. Instead, the Client Soloing feature was used as a means of identifying the source of sounds by selectively eliminating some of the music modules from the audible mix. This use of the Client Soloing feature is illustrated by the following interview excerpt (repeated from Section 7.12.7:

PARTICIPANT 3: I think the, I only ever used it just to basically, s, solo like like like, mute a whole load if, when it was getting pretty hectic
RESEARCHER: yeah
PARTICIPANT 3: like just like, just use it to just, regardless of who's I was muting I would, I was like well I'll mute that one and that just cuts all that shit out
RESEARCHER: yeah yeah yeah
PARTICIPANT 3: so it was just more of a, kind of like a, a mute of stuff that you don't want (Session 11 Interview)

In this case the participant used interfaces which did not feature the Module Soloing mechanism in both conditions, although this extract demonstrates the importance for the participant of discovering the source of sounds. The fact that some participants used the Client Soloing feature as a means of gathering Source Awareness also suggests that a re-designed Module Soloing mechanism might have been more successful.

Finally, Module Soloing, both alone and in combination with Client Soloing did not impact on reported preferences, and did not influence the participants' reported reflections about the quality of music, the best contributions or their tendency to explore the interface. This is an unanticipated result, as in the case of the previous studies increased awareness often contributed towards increased preferences towards an interface.

### 7.13.2 Hypothesis H2: Authorship Awareness

Hypothesis H2 posited that an interface which incorporate additional *Authorship* awareness would decrease instances of co-editing, discourage discussion over authorship management (e.g., spatial partitioning, discussion over who is doing what, and have a positive impact on preferences. Authorship awareness information was incorporated into the interface through the Client Soloing function. The Client Soloing feature allowed users to aurally solo and visually highlight the musical contributions from particular people within the group.

The study results provide evidence to support the hypothesis. For both sets of groups (those with Module Soloing and those without), the post-test questionnaire results provided a strong (statistically significant) indication that the Client Soloing feature increased reported levels of awareness. This is based on responses to the questionnaire statement *I knew what other people were doing*. This finding indicates that the Client Soloing feature functioned as intended as a source of awareness information. However, as anticipated, the inclusion of the Client Soloing feature also effected the degree to which participants shared and edited the music they made.

For groups who were presented with an interface which also featured Module Soloing, the inclusion of Client Soloing discouraged co-editing, and encouraged group members to work more in isolation from one another. Analysis of the interaction logs shows a trend for less co-editing to occur when the interface provided Authorship Awareness information through the Client Soloing function. While not quite at the level of statistical significance (p=0.0537), it must be stated that this finding is based on a sample size of only 8 groups. Moreover, the trend for reduced co-editing identified in the log file analysis is supported by the post-test questionnaire responses which show that for groups who were presented with the Module Soloing option in both conditions, a significant number of participants reported that they 'worked mostly on [their] own' in the interface which featured Client Soloing. This suggests that the Client Soloing feature encouraged the participants to work individually. Participants also noted being more influenced by others in the group when presented with the interface which excluded Client Soloing. Finally, there was a trend for the participants who were presented with Module Soloing to select the interface without Client Soloing as the one in which they 'edited the music together'. Although this finding is slightly above the significance level (p=0.0505), it supports the overall trend for the participants to report working more individually when the interface included the Client Soloing feature. Overall, these results suggests that while the Client Soloing feature may have enabled more awareness information to be gathered, this did not necessarily lead to members of the group drawing influence from one another, especially in the case when Module Soloing was also present within the interface.

For groups who were not given Module Soloing, the inclusion of Client Soloing had no effect on the amount of co-editing which took place. Participants did however make significantly more deletions when they had neither Client Soloing or Module Soloing, although the number of Module Creations was not significantly influenced by the inclusion of Client Soloing. This suggests that while a similar number of contributions were created regardless of the awareness features, participants were more likely to delete modules when the interface provided no means of identifying who had created them.

In the comparison of Client Soloing, more participants selected the condition which included Client Soloing as the condition they preferred the most. For groups who were not presented with the Module Soloing feature, the preference result was statistically significant, although for groups who had both Client Soloing and Module Soloing this result was slightly above the level of statistical significance (p=0.0516). This suggests that the Client Soloing feature had a strong influence of reported preferences for the interfaces, despite it not influencing the reported quality of the music, and despite causing the participants to work more individually.

In summary, the inclusion of the Client Soloing feature had three key effects on interaction.

- It increased participants levels of awareness of one another
- It discouraged the co-editing of musical contributions, and discouraged participants from deleting music modules.
- It encouraged participants to work more individually, and this idea is addressed more fully in relation to the concept of Privacy in Section 7.13.4.

### 7.13.3 Congruence of Audio Delivery

The study focused primarily on the way audio representations delivered to users can be filtered and manipulated to provide auditory indicators of the relationship between sounds and on-screen graphical interface components. While the participants in this study interacted via a shared graphical interface which visually followed a strictly congruent WYSIWIS design (Stefik et al., 1987), congruence of audio presentation was not preserved between all participants. This was because the Source and Authorship awareness mechanisms developed for this experiment operated at the individual level, meaning each participant could be listening to a unique and personalised combination of music modules, depending on which soloing functions they were using. This situation places the interface more in line with the 'relaxed WYSIWIS' model of distributed groupware (see Section 2.2.3), which allows individual users to see different representations of a shared workspace. One key distinction between research based on 'relaxed WYSIWIS' designs (Greenberg, 1996) and the research presented here is that this study relaxes the *auditory presentation congruence*.

Results of the study presented in this chapter show that in some circumstances the inclusion of an Authorship awareness mechanism decreases collaborators' tendency to edit contributions created by others within the group (referred to as co-editing), and this was reflected in self-report data collected from participants. Studying the effects of Source awareness shows little effects on the way the software was used, however participants who did not use this feature often suggested that it would be an important feature to add to the software. Video analysis of the participants' interaction suggests that allowing each user to be presented with an individualised audio mix did not introduce problems for the participants, although according to the self-reported data from the participants, some combinations of awareness mechanisms did appear to negatively impact on certain aspects of the groups' collaboration. Finally, a qualitative analysis of interviews and discussion during the interaction points to the perceived constraints of the software, the ways in which the interface was interpreted and utilised by the participants, and an account of the participants experiences while collaborating.

### 7.13.4 Privacy

Unlike the previous studies (see Chapters 5 and 6), which both featured some form of privacy mechanism, the interfaces used in this study did not feature a privacy mechanism to support the creation of musical contributions in isolation from the other group members. The decision to exclude privacy options from the interface was primarily driven by practical issues such as the need to reduce training time with the software, although it was also felt that the inclusion of privacy might reduce the effects of the Awareness features the study sought to investigate.

Interestingly, during interviews participants repeatedly suggested that some form of privacy function would have been useful. This was particularly the case for participants working in the conditions without Module Soloing, although it was a strong theme within both sets of interview codes. Participants often described features similar to the Private Space investigated in Study One (Chapter 5) or the Personal audio mix found in the Study Two interface (see Chapter 6), however none of the participants in Study Three had participated in the previous studies.

### 7.13.5 Summary of key findings

- Client Soloing (Authorship Awareness) encouraged more personal work and less co-editing for groups who were also given Source Awareness information
- Client Soloing (Authorship Awareness) had a positive influence on preferences, especially when it was the only awareness mechanism in the interface
- Module Soloing (Source Awareness) had no effects on quantitative measures of interaction with the software
- Module Soloing had limited effects on reported preferences
- A privacy function was repeatedly suggested by participants as a key feature to include
- Some interaction between the awareness mechanisms was noted, in particular that the absence of Module Soloing appeared to heighten the effects of Client Soloing
- In addition to providing Authorship-awareness, the Client Soloing function was also used to ascertain as the source of sounds regardless of who create them (what-awareness)
- Authorship Awareness information detracted from tendency to co-edit, suggesting a design tradeoff between providing knowledge of who is doing what, and encouraging co-authorship
- Individuals used the Client Soloing function as a means of developing contributions in isolation from other group members, by soloing only themselves. This activity would have been visible to other users. Privacy is therefore a desirable feature.
- The way Authorship was presented may have encouraged participants to work individually, rather than being a mechanism for quickly determining the source/author of a sound

## 7.14 Summary

This chapter presents the third and final study undertaken within the PhD. The study used a four condition design to investigate the implications of support for Authorship and Source awareness in a shared interface for CDMI. It was discovered that the inclusion of Source awareness within the interface did not significantly alter the way the software was used by participants, whilst the inclusion of Authorship awareness tended to reduce the instances of co-editing, and caused the participants to work more in isolation from one another. Unlike studies One and Two, the interface did not feature a privacy mechanism, although this was frequently requested by participants. It also appears that the Authorship awareness mechanism was re-appropriated by some participants to served the dual function of providing Source and Authorship awareness. The following chapter steps back from the minutiae of the studies and discusses more broadly the implications of this thesis.

# Chapter 8

# Discussion and Implications

---

*This chapter draws together the findings of the three studies to reflect more broadly on what has been discovered and how this relates to previous research within collaborative music and Computer Supported Cooperative Work. Following this, Section 8.3 proposes a series of design implications for future collaborative music environments. These implications are then demonstrated through the proposal of a hypothetical system for CDMI. The chapter concludes with a discussion of how the findings of this thesis might relate to other modalities such as touch interfaces.*

## 8.1 Designing for Collaborative Digital Musical Interaction

Designing new tools to support an activity may result in a re-design or alteration of the activity itself. Within the field of CSCW, it has previously been stressed that creating computer systems to support collaborative work engenders a process of 'work redesign' (Rouncefield et al., 1994), whereby the tasks and activities change in consequence to their new means of execution. Similarly, the influence of technology is a central feature in the development of music and musical thought (Boulez, 1978, Cascone, 2000, Chadabe, 1996, Makelberge, 2010), and although the reasons people create music or engage in Musicking (Small, 1998) may not have changed over time, the ways and means by which music is created have been radically altered by technology (Makelberge, 2010). The development of new technology has allowed people to create, store and manipulate sounds in new ways, in turn influencing the ways people conceive of, create and experience music. Within the context of CDMI, by constructing a novel system to support musical collaboration, the designer is modifying the ways in which musicians can communicate, contribute, share ideas and participate in acts of music making. Creating such interfaces, by definition, alters the nature and process of music making, and potentially restructures the relationships between musicians. Designers therefore have a responsibility to consider the implications that new tools or systems have on the way people work, share, create or interact, as the choices a designer makes may vastly alter the way people experience the collaboration, or use an interface.

As computers are becoming more important for the creation of many forms of music, the design of systems to support CDMI has huge potential to alter the way in which music is made by many people. As previously discussed, systems for CDMI can open up a range of new possibilities for group musical creativity. However, compared with single-user systems for musical interaction, CDMI is a relatively uncharted area for designers and researchers. Little is known about how collaborative features in an interface impact on the way groups or individuals interact, and it is important to remember that design choices often have unintended consequences, which are difficult or impossible to predict (Norman, 2002, Wania et al., 2006). Due to this, it is vital for designs to be tested and evaluated to understand how they might affect the way people collaborate. Such testing and evaluation allows us to gain an understanding of which features are useful to include in an interface, how these features should function from a usability standpoint, and what characteristics they may impose on the group interaction.

With respect to methods for evaluating interface designs and studying the effects of different interface features, CDMI presents a number of interesting challenges. In particular, the nature of music as a process centred, open ended and problem seeking activity (as discussed in Chapter 3) makes it an especially difficult domain within which to apply the methodological approaches commonly associated with HCI and CSCW. Within CSCW, the majority of studies follow either an experimental approach which focuses on the product a group creates (Gutwin and Greenberg, 1999), or an observational approach in a real-world work setting, where the group processes are studied more closely (Gutwin and Greenberg, 1999). As a methodology for studying CDMI, both approaches have limitations. The purely observational methodology is difficult to adopt as there are currently very few examples of collaborative software in existence, and therefore few ensembles of electronic musicians using collaborative music software as part of their artistic practice. The experimental approach taken within CSCW is also problematic, as it often focuses on the evaluation or rating of task outcomes, such as the quality of a document or design solution. However, the subjective nature of music means that the outcome of a musical interaction cannot be objectively compared with any other. Furthermore, for the musicians involved, the process of engaging in the interaction may be at least as important as the arriving at a final outcome (Sawyer, 2003), and in certain kinds of musical interaction (such as some forms of improvisation), there may not be a final outcome at all, as all emphasis is placed on the process as it unfolds.

As discussed more thoroughly in Chapter 4, this research opted for an experimental approach, where groups of participants were exposed to different interface designs, and a variety of measures were taken to discover how the different interface designs influenced the way they used the software. In addition, more observational and qualitative methods were also applied to gain additional insight into the qualitative findings of the experimental studies. A central focus of these studies was the design of interfaces which allow musicians to edit each others contributions in real-time via a shared software interface. This form of interaction is not possible using either acoustic instruments or existing software environments designed for single users. This thesis has demonstrated that even small changes in the user interface design can have a dramatic effect on the way individuals choose to use a shared software interface, and on their reported feelings towards such interfaces. Providing privacy, revealing information about who contributed, and altering the way sound is presented, all had effects on the way groups of musicians worked together. This finding underlines the responsibilities designers of musical interfaces have in shaping the collaborative processes musicians might engage in, and indeed one contribution this thesis hopes to make is engendering an increased concern for the issues of privacy and awareness in systems for CDMI.

## 8.2 Awareness

The literature survey of CSCW research (Section 2.1) identified the concept of Awareness as a central design concern for collaborative software in the workplace (Benford et al., 1994, Dourish and Bellotti, 1992, Endsley, 1995, 2000, Gutwin and Greenberg, 1998, 1999, 2002, Gutwin et al., 2011, Heath et al., 2002b, Hornecker et al., 2008). Within the context of this thesis, Awareness was defined as knowledge of others' actions and of ones surroundings within a shared workplace;

a definition drawn from Gutwin and Greenberg (2002). The design and evaluation of mechanisms to support awareness is an established and recognised avenue of research within CSCW. Concerning CDMI, facets of Awareness have been touched upon by previous research in a variety of interaction modalities and contexts, including performer-audience dynamics (Gates et al., 2006), table-top interaction (Klügel et al., 2011, Laney et al., 2010), web-based jamming (Bryan-Kinns and Hamilton, 2009, Gurevich, 2006, Miletto et al., 2006) and networked collaboration (Merritt et al., 2010). Despite the existence of this previous research, there is no clear picture of how to design awareness mechanisms within systems for CDMI, and there has been limited attention directed towards understanding which forms of awareness information are important to provide within shared musical interfaces.

Within this thesis, the Workspace Awareness (WA) framework (Gutwin and Greenberg, 2002) was used as a foundation for considering awareness in CDMI. Despite the shared concern for co-located collaboration interaction, the nature of music as an open ended, problem seeking activity distinguishes it in some ways from the types of task and work oriented activities the WA framework was based on. For instance the studies within this thesis have shown that within musical interaction, including too much awareness information causes participants to make fewer contributions and discourages editing of each others work.

The ephemeral, invisible and pervasive qualities of sound require careful consideration in the design of a collaborative system. The pervasive and ephemeral properties of sound make it difficult to provide passive awareness mechanisms which function in a similar way to the forms of feedthrough awareness described in the WA framework. For instance in pen and paper collaboration with visual media, sound has been identified as an important source of ‘feedthrough’ about the status of objects, and consequential communication about the actions of others in the workplace (Gutwin and Greenberg, 2002, Heath et al., 2002b) (see Section 2.2). However, in CDMI, sounds may not provide implicit information about the environment or activities of others. Primarily, this is because in a computer-based system the way audio is presented may displace individuals’ actions from their physical location in the environment. This was explored in Study Two (Chapter 6), where the use of speakers and headphones to separate personal and public audio proved to be an unsuccessful design, as participants found it difficult to concentrate on both the Personal and Public channels when they were split across headphones and speakers, and spatialising the personal channel using a personal speaker next to each participant proved to be unsuccessful as a means of providing awareness about the activities of others.

In contrast to awareness mechanisms which attempted to provide information passively, analysis of the interviews and questionnaires suggested that awareness mechanisms which required additional effort from the participants appeared to be more successful. Such mechanisms included the ‘Views’ tab in Study One, which allowed participants to choose which personal space to listen to (their own or someone else’s), and the Client Soloing mechanism in Study Three. These mechanisms provide participants with control over the information they were presented with, allowing them to selectively attend to specific aspects of the shared music, rather than being simultaneously presented with the contributions from all participants. To a greater or lesser degree, these mechanisms presented incongruous or personalised representations of the group music to each participant, allowing the individual participants to choose what they were attending to. This is perhaps especially important in the context of auditory collaborations, where a surplus of sonic information might quickly turn into a cacophony.

This suggests that within systems for CDMI, a balance should be struck between presenting appropriate levels of awareness about the activities of others, and causing confusion by delivering too much detail at once. Providing user controllable mechanisms for managing awareness can circumvent this issue, although care must be taken to design awareness mechanisms so as to make them easy to use without distracting from the musical activities. For instance whilst the ‘views’ mechanism in Study One provided participants with an overview of what their collaborators were working on, it did not allow them to easily compare their own work with that of other people, and therefore required participants to remember what each personal space sounded like when returning to work within their own personal space.

### 8.2.1 Authorship Awareness

This thesis proposed in Chapters 6 and 7 that the WA framework is not suitably nuanced to account for certain aspects of awareness within CDMI. In particular, interaction with sound via a distributed graphical interface can dislocate participants’ actions and contributions from their sonic manifestation, and can disassociate sound sources from their visual representations in the interface. This can lead to a disconnection between *who* contributed something and *what* within the interface is responsible for controlling this contribution. Such confusion is less likely to occur in collaboration over primarily visual media, due to the ease with which referential identify can be established through gaze and pointing with fingers. As argued in Chapter 3, the properties of sound make these actions less effective in collaboration over auditory media. In response, the categories of Authorship Awareness and Source Awareness are proposed within this thesis to account for awareness of which individual contributed within the musical collaboration, and awareness of which elements within the interface are responsible for which sounds. These distinctions have not previously been proposed within research into the design of collaborative interfaces for musical interaction.

The concept of Authorship Awareness was developed to account for awareness of who contributed which musical elements. Study Three (Chapter 7) suggested that increasing the availability of information about the authorship of musical material within the shared workspace decreased the tendency for individuals to edit one-another's contributions. This could be due to participants attaching a greater sense of ownership to other people's contributions when they are able to ascertain who created them. Studying the Daisyphone, a similar idea was suggested by Bryan-Kinns and Hamilton (2009), who used colour to indicate the creator of notes in a shared sequencer. This thesis extends the finding of Bryan-Kinns and Hamilton (2009) in a number of ways. Firstly, the software used for this research was more complex, offered more opportunities for participants to create personalised musical contributions, and was designed with experienced musical users in mind. Secondly, concerning the awareness mechanism itself, the interface in this thesis allowed individuals to *listen* to each other, and sonically isolate the sounds created by others as a means of attributing authorship. In comparison, the Daisyphone interface used only a visual indicator of authorship, although this may have been sufficient for the simplified interface.

Territorial approaches to laying out the interface were also identified as influencing the tendency for participants to edit each others' contributions. In Study Two an analysis of the way participants chose to lay out the interface revealed evidence of highly territorial behaviour in around half the groups, and this behaviour reduced the likelihood that individuals would edit contributions which they had not initially created. As it seems the groups adopted these territorial arrangements so as to scaffold awareness of ownership, this also follows the finding that increased authorship awareness can discourage collective editing of musical material.

The idea that increased awareness can impact on the way people contribute in a shared environment has also been identified in other collaborative tools (Pinsonneault et al., 1999, Rogers et al., 2009). For instance within tasks such as group mind-mapping (Pinsonneault et al., 1999), increased authorship awareness has been noted to reduce the creativity of the group by making members more self-conscious and less inclined to edit or critique contributions which are not their own. Whilst CSCW research often advocates an increase in awareness information provided by an interface (Gates et al., 2006, Gutwin and Greenberg, 1999, 2002, Gutwin and Penner, 2002, Gutwin et al., 2011), within CDMI the inclusion of authorship awareness might be considered as a design choice or tradeoff between the benefits of allowing participants to know who contributed what, and potentially reducing the degree to which participants will edit collectively. Future studies could investigate the degree to which this is influenced by the time groups spend working together, and possible interactions with other awareness mechanisms within the interface.

### 8.2.2 Source Awareness

In all the studies within this thesis, participants discussed issues surrounding knowledge of which interface elements were responsible for creating which sounds. This is not an issue which is exclusive to collaborative interfaces for music. In a single user musical interface, users may need a means of identifying where a sound is coming from. However, the multi-user nature of CDMI can make matters more complicated, as a given user is not necessarily responsible for all the items in the sounds in the shared interface, so may require additional support in discovering their source. In Study Two, some participants used the Personal audio channel as a means of identifying the source of sounds, by temporarily routing sounds to their headphones. This re-appropriation of the privacy feature prompted the design of a Module Soloing function in Study Three, which allowed individuals to isolate the sound of specific music modules within the shared workspace.

Study Three investigated the concept of Source Awareness through an investigation of the Module Soloing function. The results suggested that this awareness mechanism did not alter the way the participants used the software, and had only minimal influence on self-reported statements about the software, the music or the quality of the group interaction. Analysis of the groups' discussions during the interaction suggested that Source Awareness was, however, useful within the collaboration, suggesting that the Module Soloing function could be redesigned to better support the presentation of Source Awareness.

### 8.2.3 Privacy

Chapter 3 identified CDMI as a mixed focus activity in which participants work both individually and as a group. Within a CSCW context, support for mixed focus collaboration often entails the provision of private workspaces (Dourish and Bellotti, 1992, Stefik et al., 1987), or other mechanisms to support parallel working within a shared interface (Ishii et al., 1993, Morris et al., 2004, Vogel and Balakrishnan, 2004). However this concept has not been given much attention in previous research on group musical interaction. For instance, the majority of collaborative music environments previously developed have not featured private work spaces or offered users control over the access other people have to their contributions (Bryan-Kinns, 2004, Burk, 2000, Fels and Vogt, 2002, Gurevich, 2006, Jordà et al., 2007, Newton-Dunn et al., 2003). Although systems for CDMI can potentially enable groups of people to collectively edit, share and co-create musical works in a democratic fashion (Xambó et al., 2011), the findings of this thesis show the importance of individuality, privacy and ownership within collaborative music making. In fact, findings from all the studies pointed to aspects of Privacy as being central to musical collaboration, and this section therefore highlights the sensitivity a designer must place on the design of features to support these concepts.

Study One used both visual and auditory privacy within the shared interface by presenting a Private Space for each user, in addition to a Public area which was available to all users. The privacy feature was heavily exploited by participants to facilitate the creation and preparation of musical ideas in private before sharing via the Public space. The inclusion of the Private space also discouraged Public editing and co-editing, suggesting that groups not only used the Private space extensively, but were discouraged from using the Public space for performing modifications. The extent to which privacy features were exploited by the participants contradicts the findings of Dourish and Bellotti (1992), who reported that participants in their study of text based collaboration neglected to use their individual workspaces whilst working in real-time on a design task. Dourish and Bellotti (1992) explained this in part because of the time constraints of the task they were assigned during the experiment, however the studies presented in this thesis were of a similar duration to those conducted by Dourish and Bellotti (1992), perhaps suggesting that the nature of musical interaction requires more sensitivity to privacy.

In Study Two, visual privacy was relaxed through the use of a single shared workspace. Within the workspace, public contributions were visible and editable by all users, whilst grey boxes were used to indicate the presence and location of contributions which were held in private by a single participant. The auditory aspect of privacy was manipulated in the study through changing the way audio was presented. Presenting private audio through individual speakers placed next to the users' screens reduced the extent to which participants used their private audio channels, when compared to presenting audio through headphones or a combination of headphones and speakers. This finding indicates that the ability to listen to audio in isolation from others is a determining factor in the usefulness of a privacy feature within CDMI.

Finally, whilst the interface in Study Three did not feature any privacy functions, some participants (non of which had used the previous interfaces) discussed using the soloing functions as a method of isolating their own sounds. Some participants also suggested that privacy features similar to those studied in the previous experiments would be an important addition to the interface.

In summary, the absence of privacy features was often noted as a problem by the participants, whilst the inclusion of privacy features and options for allowing incongruous audio presentation often resulted in higher preference ratings for the interfaces. The inclusion of such features had a measurable impact on the way participants collaborated in all studies. A strong implication is that collaborative musical systems should feature support for individual work, therefore allowing users to formulate their contributions away from the group before sharing them. According to the participants' statements, this allows collaborators to contribute more high quality contributions, and provides an alternative to working exclusively in public, although it may consequentially discourage group members from editing each others work. This was particularly apparent in Study One, where participants were exposed to the condition which did not feature a Private workspace.

### 8.2.4 Balancing Awareness and Privacy

The concepts of Privacy and Awareness are often tightly interwoven, as an increase in Privacy may necessitate additional awareness mechanisms to facilitate effective collaboration (Gutwin and Greenberg, 1999, 2002, Gutwin et al., 2011), whilst providing awareness mechanisms by definition reduces the level privacy a user is afforded. Consequentially, the inclusion of privacy within a collaborative interface complicates the design. Discussing impromptu face to face collaboration over tabletops, Shen et al. (2003) stresses that a binary notion of public and private is too simplistic a model for most forms of collaboration. The findings of this thesis suggest that within CDMI this position also applies. Crucially, it is important for users to be made aware of the activities of others, even whilst working alone, otherwise users may find that their new contributions do not fit with those in the public area, or those being created by others in private.

Study One proposed that whilst such an awareness mechanism was used by participants to gain an overview of who was doing what, it did not influence the way the software was actually used. Similarly, the Module Soloing function in Study Three did not have a measurable impact on the way participants used the software. This suggests that such awareness mechanisms can change the subjective quality of the interaction, although they may not elicit a strong effect on measurable features such as the amount of editing which takes place or the tendency for people to edit contributions made by others.

Another important issue raised in the analysis of the studies was the presentation of awareness information about the privacy status of musical contributions. If systems for CDMI provide mechanisms for supporting individual work, they need to also provide users with information about which aspects of the music product are currently private, and which are currently shared. This is a separate issue to providing other users with awareness of the activities of others. In physically spatial forms of workspace collaboration, information about the status of contributions might be performed through proximity and territory (Scott et al., 2004, Shen et al., 2003), however in sound, such distinctions between public and personal contributions are not so easily made. In Study Two, using audio delivery as a means of separating public and personal audio by means of individual speakers and headphones met with limited success, as participants tended to prefer conditions where all audio was presented through a single device. This could have been related to the way the speakers and headphones were physically organised, however a more likely explanation is that the pervasive nature of sound made the separation of Personal and Private sounds problematic. Future research might uncover ways of successfully providing information about the privacy status of musical contributions through audio delivery mechanisms such as highly directional speakers, however a more practical approach might be to provide visual indicators combined with soloing mechanisms which allow users to individually interrogate the audio mix for the purpose of determining which sounds are public and which are private.

Taking this into consideration, it could be argued that designers should not aim to create interfaces which promote entirely shared and democratic means of music making, but should instead focus on providing opportunities for people to explore their own ideas and make personal contributions. The following sections explore in more detail aspects of the support for private working, awareness and issues surrounding the congruence in the presentation of audio.

### 8.2.5 Audio Delivery Congruence

In addition to privacy features, mixed-focus interaction was explored in this thesis by considering audio presentation congruence. Within HCI the concept of congruence has commonly been used to refer to visual interfaces, where it concerns the degree to which users see and interact through the same representation or view of a shared artefact (Stefik et al., 1987) (see Section 2.2.3). Within shared visual interfaces for group collaboration it has been noted that allowing interfaces to relax their congruence can aid in collaboration by allowing individuals to work in isolation from other group members, or work on separate aspects of a shared artefact (Greenberg, 1996, Stefik et al., 1987). Support for such activities also parallels the kinds of interaction which take place in pen and paper based collaboration (Healey and Peters, 2007), and across tabletops (Scott et al., 2004).

This thesis contributes to an understanding of congruence in collaborative interfaces by considering how audio can be presented to groups of users during musical interaction via a shared interface. Specifically, the ability for users to work in private, and the ability for participants to choose which audio is presented to them all break the congruence of the shared interface by allowing individuals to listen to discrete audio representations. Although privacy features are the most common way in which congruence of audio presentation was effaced within the studies presented in this thesis, it is important to make a distinction between interfaces which incorporate privacy, and interfaces which feature other forms of incongruence. As an example, some of the interfaces in Study Three allowed participants to audition different combinations of music modules simultaneously on their own headphones, thus breaking the congruence of audio presentation between users, without supporting privacy by preventing others from listening to specific contributions.

Interestingly, the most commonly preferred interfaces in all the studies were the ones which provided opportunities for incongruence, either through privacy features or through awareness mechanisms. Fully congruent interfaces were not favoured by self-report measures in any of the studies, and during interviews participants regularly talked about the advantages of being able to work in isolation, and also of being able to choose what they were listening to. Such support for incongruence, even without privacy, may have been useful as a means of allowing participants to focus on specific aspects of the music, and selectively shut out elements which they did not wish to hear. This finding supports the assertion made by others (Greenberg, 1996, Stefik et al., 1987) that relaxing congruence can aid collaboration if supported by appropriate awareness mechanisms.

## 8.3 CDMI: Implications for Design

This section draws together the findings of this thesis and synthesises a number of general design implications for systems to support Collaborative Digital Musical Interaction. In summary these implications are:

- Support for Individual Work
- Increased Authorship Awareness Leads to Less Collective Editing
- Momentary Provision of Awareness Information
- Avoidance of text-chat
- Single Device for Audio Delivery
- Ability to Offload Coordination and Organisation onto the Interface
- Auditory Displays of Information

In addition, the final part of this section considers implications for other devices and platforms.

### 8.3.1 Support for Individual Work

As discussed previously, the findings of the studies within this thesis have consistently pointed to the advantages of providing opportunities for users to work individually during musical interaction. This has been somewhat overlooked in other collaborative music environments, although privacy and mixed focus interaction have been acknowledged in previous CSCW research.

Xambó et al. (2011b) used findings from Fencott and Bryan-Kinns (2010) to propose that collaborative multi-touch music environments should include both 'shared and personal workspaces' to facilitate both the sharing of ideas and developing ideas individually. However the exact form of these workspaces, the way people interact with them, and they way they are represented in the interface are not specified in their 'Candidate Design Pattern'. As the Candidate Design Patterns in Xambó et al. (2011b) are based on a literature survey their apparent lack of specificity could be a consequence of limited available literature to draw upon. The experiments conducted for this thesis contribute to an understanding of how to design for privacy by exploring two different privacy implementations, separate on-screen workspaces, and the use of audio presentation devices to support privacy in a single shared workspace design. Some important considerations are:

- Ease of movement between Public and Private
- Awareness information pertaining to the privacy status of musical contributions
- Support for awareness of private activities for other users

### 8.3.2 Increased Authorship Awareness Leads to Less Collective Editing

Providing Awareness of Authorship is an ambiguous issue within CDMI. Although Study Three (Chapter 7) showed that the interface which incorporated authorship awareness was rated more positively by the participants, there is evidence from within this thesis and elsewhere (Bryan-Kinns and Hamilton, 2009) that such information reduces the tendency for users to edit each others contributions. Designers must therefore take this into account when incorporating such features into collaborative interfaces, and weigh up the benefits of providing this information with the possible implications it has for the way people will use the system. Future studies, as outlined in Section 9.2 could investigate the extent to which this effect of Authorship awareness is influenced by the length of time people use the software, and their familiarity of the group members with each other in a creative context.

### 8.3.3 Identification of Musical Elements (Source Awareness)

In CDMI, where a group of people are simultaneously contributing, editing and removing musical elements from a shared workspace, the mix of audio can quickly become confusing. Individuals may not know the source of all sounds in the mix, and may therefore accidentally remove the wrong thing, or change the wrong parameter. A strong implication for this research is that interfaces should provide a means of allowing participants to easily identify the source of sounds they can hear.

The grounded analysis of interviews conducted as part of Study Two (Section 6) suggested that participants often wanted to discover which elements within the interface were responsible for creating specific sounds within the musical output, and that in comparison to this they were less concerned with who was responsible for the authorship of specific contributions. This suggests that interfaces should provide separate mechanisms for identifying which individuals made which contributions, and in addition, a mechanism for determining which contributions are responsible for creating which sounds.

The interface did not provide mechanisms to easily isolate specific music modules, or to audition the contributions created by a specific user. Participants frequently stated that one use of the Personal audio channel was to identify specific elements or components in the composition by re-routing. Using the Personal channel in this way was not anticipated, and highlights a key deficit in the software design. While using the Personal audio channel to identify specific sounds within the mix would be a rather cumbersome task, it is more important to stress that using the Personal channel in this way would impact on the other group members, as re-routing any of the music modules from the public output (audible by all participants) to a Personal channel (audible by a single participant) would remove this module from the audible mix presented to the other users.

Given that the ability to identify musical components within the mix was clearly a useful feature for users, a more efficient mechanism for performing this task could be incorporated into a single audio delivery design. For instance many mixing consoles and digital audio workstations incorporate a 'solo' button to isolate audio at a particular point in the signal path. In a collaborative environment a solo button could be designed to operate on the audio mix presented to each individual. This would allow an individual to interrogate various aspects of the interface and discover which components are responsible for making which sounds, without disrupting the work of other people in the group. The study presented in Chapter 7 explores this concept in greater detail.

### 8.3.4 Momentary Provision of Awareness Information

Participants in Study One used the 'Views' awareness mechanism for a short amount of time, although in interview the features was repeatedly discussed as beneficial. In Study Two participants had difficulties making use of the spatial audio delivery through multiple speakers to maintain awareness of one-another, and talked about re-appropriating the personal channel as a means of identifying the source of certain sounds. Study Three suggested that while the awareness mechanisms functioned to provide participants with the ability to glean additional information about the interface and the shared music, the Client Soloing mechanism was re-appropriated to function in a similar way to the privacy functions of the first two studies. This suggests that rather than providing continuous Awareness of the activities of others, awareness mechanisms for CDMI should function in a more momentary fashion, so as to enabled users to gain brief auditory glimpses of the activities of others, without being distracting. There mechanisms should be under the control of individual users (i.e., incongruous), rather than operating automatically or at a group level.

### 8.3.5 Avoidance of text-chat

As a feature, a text-chat tool is easy to implement, however it offers only limited benefits within the context of collaboration in time-based media such as music. This is primarily because persistent text messages can quickly lose their context and meaning in relation to the continually evolving and ephemeral musical sounds. Some participants discussed this in the interviews, stating in particular that they did not notice text messages appearing, or that when they checked the chat window, they were unsure which messages referred to which sounds. Secondly, there is the issue of establishing referential identity to sounds via the medium of text chat. Participants sometimes uses vocal mimicking to draw attention to sounds they wished to draw attention to, however this is not possible using a chat-tool due to the text based nature of the communication. The human voice is capable of mimicking many different sounds, however these can not be easily translate into a written form.

If users of a CDMI system are remote, rather than co-located (therefore excluding the possibility of using fingers to point, and vocal mimicry), a more effective solution might be visually co-located annotation windows or text-messaging facilities with each musical contribution, so as to scaffold the establishment of referential identity. If used correctly, such a design would reduce the ambiguity over which messages referred to which musical contributions. One possible problem with this approach is that having multiple points within the interface where dialog can occur might become confusing, and might cause more text chat messages to go unnoticed. The affordances of such a design would therefore need to be evaluated within future work.

### 8.3.6 Single Device for Audio Delivery

While certain musical practices use multiple devices for audio presentation, and in particular, the use of headphones as a personal channel for cueing and preparing material before sharing it on speakers, Study Two suggested that this approach was not beneficial form an awareness perspective. This was because participants found it difficult to switch between using headphones and speakers for listening, whist simultaneously working on musical contributions and attempting to maintain awareness of the contributions made by other people in the group. Therefore, an implication is that audio should be presented on a single device, such as headphones or shared speakers. In particular, the findings of Study Two (Section 6.12.1) indicated that participants reported feeling most involved with the group when audio was presented through a single device, rather than a combination of headphones and speakers. The findings do not indicate that either headphones or speakers was preferred by the participants, although the results do suggest that presenting audio through speakers discouraged the use of the privacy function. As discussed previously (in Section 8.3.2), if a systems is to use a single device, designers may need to balance the benefits of headphones in supporting individual privacy against the benefits of presenting audio through speakers such that it encourages users to become more involved in the group.

In addition, it is important to state that within the scope of this thesis, the design implication of using a single device for audio delivery applies primarily to first time users. Future work may suggest that users with more experience find switching between headphones and speakers less awkward. This is acknowledged in the limitations of this thesis (see Section 9.2).

### 8.3.7 Ability to Offload Coordination and Organisation onto the Interface

Participants were observed to use the spatially reconfigurable properties of the virtual workspace as a support for organisation of the musical materials of their collaboration. A common approach was for each group member to work in a self-allocated area of the screen, although some groups attempted to group similar musical elements (such as drums or melodies) within the same location on the screen. These kinds of activities highlight the importance of the interface as a place to offload organisational and coordinative information. A clear implication for the design of future systems is therefore the provision of mechanisms to support informal forms of coordination such as spatial organisation. These mechanisms should not impact directly on the music being created, but should enable users to arrange, configure and organise their musical contributions so as to support and scaffold their collaboration. These mechanisms should also be user configurable, rather than prescribed by the system, so as to support a plurally of approaches by different users.

### 8.3.8 Auditory Displays of Information

As music is an auditory domain, it makes sense for interfaces to provide awareness mechanisms which allow for investigation and interrogate of the being presented. For instance Study Three investigated the implications of being able to solo the audio from specific participants. This extends previous research which has looked at colour identification within collaborative musical interfaces (Bryan-Kinns, 2004), yet without the ability to *hear* other people's contributions.

## 8.4 Hypothetical System Design

This section proposes a hypothetical system for CDMI, as a means of illustrating the design implications contributed by this thesis. This hypothetical system is a paper prototype and has not been implemented in software. The design is based on the software interfaces described in Sections 5.1, 6.2 and 7.4. In particular, it follows the same shared workspace and relaxed WYSIWIS metaphor, and is a modular environment within which users create and edit virtual instruments. The design does however propose a number of changes and deviations from the designs which were implemented for the practical component of this thesis. These changes are largely focused around support for collaboration, the provision of awareness information, and means by which the interface can be used to support organisation during the collaboration. The design does not expand on the music-making or sound generation capabilities of the software interface used in the practical sections of this thesis, and as with the systems developed for this thesis, the hypothetical system is designed for computer literate and musically experienced users, rather than novice or non-musician users.

Figure 8.1 presents the hypothetical design. The key features are numerically annotated, and described bellows. Subsequent sections discuss and justify these features in greater depth, with reference to the design implications presented in Section 8.3.

1. The User Identification with Peek, Solo and Mute options for each user
2. Module Annotation/History pop-up
3. Module Peek, Solo and Mute menu
4. Workspace colour panels
5. Public audio output
6. Public audio Peek, Solo and Mute menu
7. Personal Audio Outputs

8. Personal Audio Output Peek, Solo and Mute menu

9. Personal Audio Outputs can be re-positioned by the users

10. Music Modules can be collapsed to save on-screen space

11. Modules can be connected to multiple outputs

12 Modules connected to the a Personal output are represented as grey boxers for other users

### 8.4.1 Public and Personal audio channels

As with the interface featured in Study Two, the proposed design includes Public and Personal audio outputs. These allow users to share musical contributions with everyone via the Public interface, or enable individuals to work on contributions in private by routing the audio to their Personal output. This provides support for individual work, as advocated by Section 8.3.1. The Public and Personal outputs are represented as blocks within the interface. Routing audio to the Public output makes the module editable by all users. Annotation number 12 in Figure 8.1 shows that modules which have been routed to a Personal output are represented as greyed out boxes for all users except the owner of the Personal output.

Unlike the interface used for Study Two, modules can also be routed to *multiple* outputs, so as to provide an intermediate level of privacy, whereby a module can be made audible to everyone in the group whilst still remaining protected and non-editable. This is shown by annotation point 11 in Figure 8.1. This additional control over access is intended to provide users with an intermediate step between concealing and sharing, whereby a musical contribution can be introduced without it being fully shared with the group.

### 8.4.2 Peek, Solo and Mute options

Music Modules, and other items in the interface, feature a Peek, Solo and Mute menu. This provides a variety of controls over the audibility of sound sources. The Mute functions prevents a sound source of channel from being heard. The Solo function causes one or more items to be made audible, whilst all items which have not been soloed become inaudible.

The Peek function allows a currently inaudible item to be played momentarily, so as to provide the user with a quick indication of what it sounds like. When called on a currently audible item, the Peek function temporarily causes all other sound sources to become inaudible, so as to

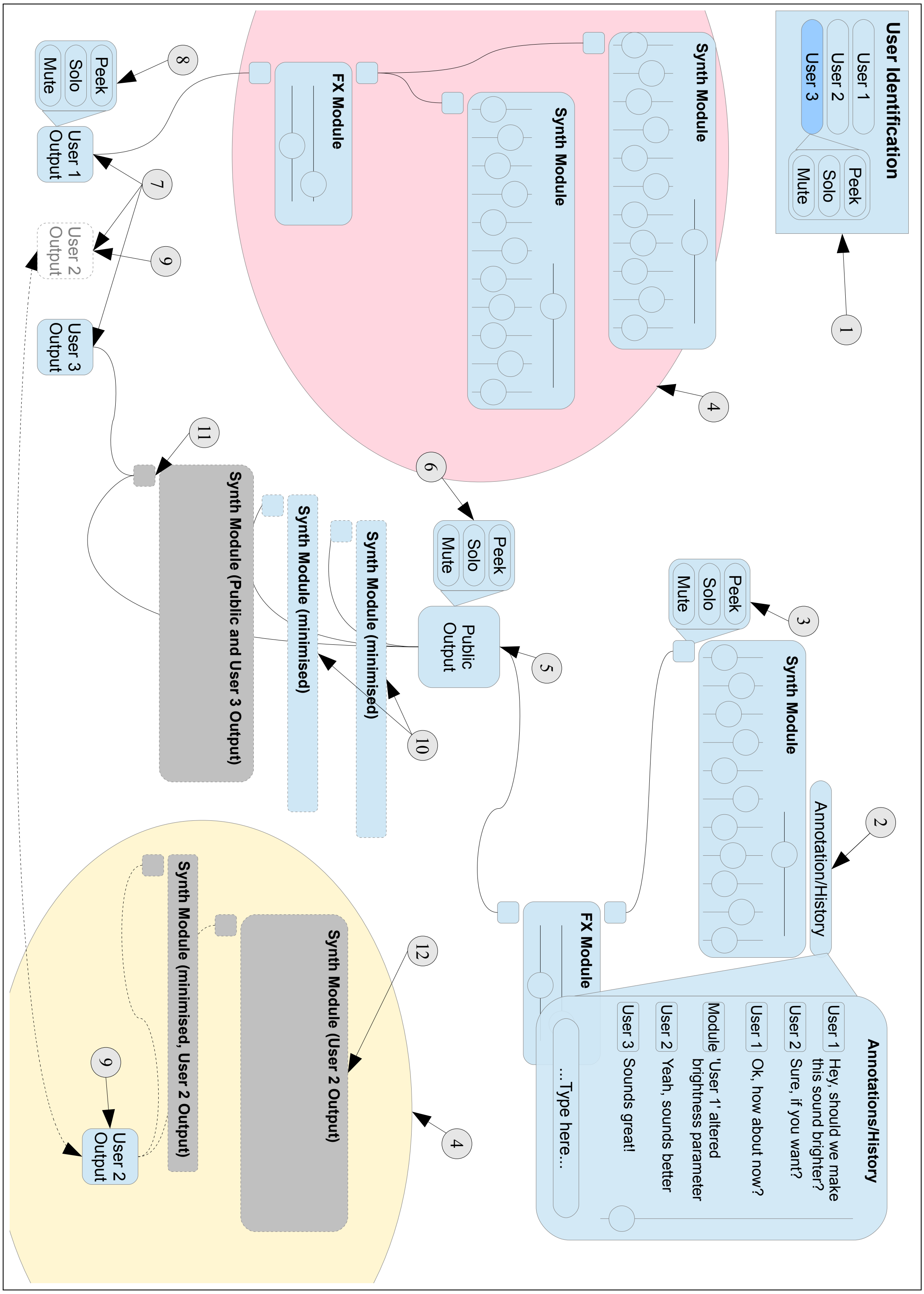



Figure 8.1: A hypothetical CDMI environment, embodying many of the design implications proposed in this thesis. The interface is presented here as though it were viewed from the perspective of User 1

enable the quick identification of sounds within the mix. For instance a user might use the Peek function when attempting to discover the source of a sound which is playing out of time.

The Peek, Solo and Mute options can also be used to gain an overview of what other users are currently working on. For instance it would be possible to call the Peek function on another user's private channel to hear a quick snippet of the sounds they are creating. This Momentary Provision of Awareness Information provides a similar function to the 'Views' mechanism presented in Study One, but would not disrupt the activities or workflow of a user to the same extent.

### 8.4.3 Workspace Colour Panels

The coloured areas to the left and bottom right of Figure 8.1 represent user definable colour panels which can be placed anywhere in the workspace. These allow users to colour code areas of the workspaces for any purpose they see fit. Users can create as many coloured panels as they wish, and would be offered a range of colours. The panels do not affect the sounds or music in any way, but instead serve as a flexible means of organising the Music Modules within the shared workspace.

For instance users might choose to colour code separate areas as personal workspaces for each member of the group, or they might choose to colour code designated areas for particular kinds of musical contribution, such as drums or melodies. Another possible use of the colour coded zones might be to separate contributions which have been finalised from contributions which are regarded as experimental or unfinished.

The inclusion of the Colour Panels is based on the design recommendation for the 'Ability to Offload Coordination and Organisation onto the Interface'.

### 8.4.4 Movable Personal and Public Channel Terminals

The Public and Personal output terminals are the points at which users connect music modules so as to route their sound publicly or privately. They are annotated as 5 and 7 in Figure 8.1. In the Study Two and Study Three interfaces the position of the output terminals was fixed, making it impossible for users to move them. In the proposed design, the output terminals can be positioned anywhere within the workspace. The ability to move the output terminals in this way follows the design recommendation for offloading coordination and organisation onto the interface by allowing users to organise the shared workspace in accordance with their needs. For instance annotation 9 in Figure 8.1 shows that User 2 has moved their Personal channel to the bottom right of the screen, so that it is closer to the modules they are working on.

### 8.4.5 Annotation History Dialog

The Annotation and History Dialog is labelled number 2 in Figure 8.1. This replaces the text chat tool found in previous interfaces, whilst still allowing for textual information to be exchanged between users. The Annotation History dialog allows users to attach named and time-stamped text messages to any music module within the collaboration. The dialog also logs all parameter changes made to a module, and allows users to step back through the editing history of the module to recall a previous state or 'undo' changes which have been made. In 8.1 the Annotation History dialog is only shown associated with one Module, however if the interface were to be implemented, the dialog would be available as a pop-up panel for all modules.

The Annotation dialog can be used for instant messaging between users, however it has several benefits over a simple text-chat window. Firstly, where users are talking about a specific musical contribution, the Annotation History dialog co-locates the text message with the music module itself, thus aiding the users in establishing reference to a particular aspect of the interface. Secondly, by coupling time-stamped text messages with the undo feature, the dialog provides additional contextual information to the message history by allowing users to recall what the module sounded like at the time messages were being exchanged about it.

### 8.4.6 Hypothetical Design as Future Work

As well as illustrating the application of a number of the design implications proposed by this these, the hypothetical design also provides the starting point for future work in the field. The implementation of such an interface would be in itself a challenging task, however more interestingly, studying how the proposed features impact on the collaborative and creative process of group musical interaction would be a valuable contribution to the field of CDMI research.

## 8.5 Implications for Other Platforms and Modalities

This section addresses implications this research may have for other types of devices, and other forms of interaction.

### 8.5.1 Touch Interaction, Shared Screens, Tablets and Tabletops

The studies presented in this thesis used a separate computer for each participant. This decision was motivated by the desire to develop collaborative software which could potentially work with everyday computing devices such as laptop computers, as opposed to more esoteric devices such as multi-touch tables. Devices such as laptops and computer monitors can be orientated so as to afford users with visual privacy over their screen, whilst using headphones can afford auditory privacy for users. Such affordances were exploited by the studies presented in this thesis. In particular, the inclusion of the privacy features in Study One and Two allowed participants to view and audition musical contributions in isolation from the group, whilst in Study Three the awareness mechanisms allowed individual users to listen to personalised audio mixes. Incongruence of this kind has previously been identified as beneficial in collaborative systems (Stefik et al., 1987).

However, other forms of interface may not offer this degree of separation between each users' view into the system. For instance an interface running on a screen which is shared by several people will provide all users with visual access to all the information within the system. An example of such a system could be a games console connected to a large TV or projector. In computer games, applications such as Guitar Hero divide the screen into smaller sections for each user, however this does not prevent players from viewing each others' screens should they so wish. Collaborative interfaces designed for shared screens can be regarded as examples of Single-Display Groupware (Stewart et al., 1999). Although this thesis has not concerned the design of SDG, there are a number of ways in which the finds may still be applicable. In particular, although the screen may be shared, SDGs can still feature individual audio channels (Morris et al., 2004), making the design implications related to audio presentation relevant. Presenting audio through a combination of headphones and speakers is likely to be at least as difficult and cumbersome when users orient around a shared screen as when they have individual screens, and presenting audio through headphones as opposed to speakers might be expected to discourage conversation and cause feeling of isolation between group members, although future studies might investigate if these feelings are mitigated by the inclusion of a shared screen.

The introduction of touch interaction is another interesting case. Multi-touch screens are not yet commonplace, however they represent an active field of research, and may soon become common everyday computing devices. Musical collaboration on multi-touch interfaces has been investigated by a number of researchers. As well as reducing the level of visual privacy afforded by the shared screen, the inclusion of touch potentially reduces the level of anonymity an interface can support, as it is possible for collaborators to who is manipulating which aspects of the shared interface, in much the same way as people can see others whist working together at a tabletop. Similarities to tabletop interaction may also encourage heightened territorial behaviour, as people may claim a physical area of the table (Scott et al., 2004).

Touch also provides collaborators with consequential information about the activities of others, potentially reducing the need for awareness mechanisms to be designed into the software. For instance a person's orientation around the screen, and their gestures with the screen can be seen by others, possibly making awareness information about authorship less useful than when people are working on separate screens. This consequential information naturally reduces the extent to which users' actions can be hidden from others, and may make the inclusion of personal audio channels more important as a means of allowing users to work in isolation from the group if so required.

Unlike multi-touch tables, tablet computers such as the Apple iPad[1] are already commonplace computing devices. The opportunities these devices offer for musical collaboration are still to be explored. A collaborative tablet interface designed such that multiple users share a single device would be similar in many ways to small touch table. In comparison, a collaborative interface which runs on multiple tablet computers, each held by an individual in a group, would share many of the challenges faced within this thesis. Such an interface would provide each user with their own screen, thus allowing for visual privacy, and the designer is still faced with similar tradeoffs regarding the way audio is presented. In the case of users working on separate tablets, the inclusion of touch interaction is perhaps less informative as a source of awareness, as it may not provide detailed information about where a user is touching.

The mobility of tablet devices opens up yet more possibilities. For instance a number of tablets could be re-arranged and moved during the interaction, potentially offering the opportunity to combine the openness and visual access provided by a large touch table with the privacy options afforded by individual computer screens. These kinds of interfaces and interactions have yet to be fully explored in the literature, and may be an exciting avenue of future work.

Emphasis is often placed on the affordances and advantages of touch interaction in supporting collaboration and providing awareness. Touch provides a direct indicator (through feedthrough

[1]http://www.apple.com/ipad/ Last Accessed 15 October 2012

awareness) of who is working on a particular aspect of a shared representation. However the findings of this thesis, and in particular Study Three suggest that too much information about authorship can discourage individuals from engaging in co-authorship or co-editing of contributions. Furthermore, designers should consider that in a single screen touch table context a musical parameter may be represented by a single graphical object (rather than duplicated across multiple personal screens, as in this thesis), and that this may further contribute to people choosing to work individually.

### 8.5.2 Remote Collaboration

Remote online collaboration has been the goal for a number of researchers and commercial ventures in the past. However, this thesis has demonstrated that there is still much to be learnt about co-located musical interaction using computers. A central issue for remote collaboration is the support for communication between musicians. Unless there is a speech communication channel, remote collaborators will not be able gain awareness information through verbal enquiries such as identified in Study Three, Section 7.12. Interfaces for remote collaboration must therefore work harder to provide such information to users, and must also incorporate flexible mechanisms for communication. As noted above, text-chat is not an especially suitable method of communication for real-time interaction over a time based medium such as music. Interfaces should therefore provide mechanisms to support planning, and should allow users to gain an overview of what their collaborators are doing, so as to support anticipation of future contributions.

## 8.6 Summary

This chapter has drawn together the findings of the empirical studies, and discussed them in relation to the overarching themes of privacy and awareness within CDMI. This has led to a series of design implications which are rooted in the findings of this research. The final chapter of this thesis concludes with a brief summary, a review of the limitations within this research and an outline of future work.

# Chapter 9

# Conclusion

---

*This chapter summarises the key findings and restates the contributions of the thesis. Limitations are presented and the thesis concludes with suggestions for future work.*

Considering the extent to which music is a social and collective activity, it is surprising that there are so few examples of collaborative music software, and it is surprising that the design and evaluation of collaborative music systems remains an under explored field of academic research. The subject domain of this PhD was characterised as Collaborative Digital Musical Interaction (CDMI). The emphasis this term places on the collaborative nature of the interaction distinguishes it from previous definitions of musical interaction identified within the literature. The activity of CDMI was examined through a review of technological constraints, a discussion of the properties of sound as a medium, and a concern for the open ended, creative and problem seeking nature of music-making.

Taking as a starting point the more academically established field of Computer Supported Cooperative Work, the practical component of this PhD explored the implementation of mechanisms to support facets of privacy and awareness within systems for CDMI. This work makes several novel contributions to the field. Firstly, there have been no previously published studies concerned with the evaluation of awareness mechanisms for co-located group collaborative musical interaction supported via shared software interfaces. Secondly, the results of these studies point to two key implications; that privacy is a highly valued feature within CDMI, and that increased awareness of authorship can lead to a reduced tendency for people to edit musical contributions they did not make. Finally, based on the empirical data gathered, a series of design implications have been synthesised to aid designers in the creation of future systems for CDMI. Implications have also been proposed to extend beyond the shared workspace metaphor studied in this research.

The following sections present a brief synopsis of the empirical studies and a summary of the key findings. The chapter draws to a close with suggestions for future work.

## 9.1 Overview and Major Findings

Several prototype music making environments were developed to facilitate this research. The motivations for creating software as part of the research methodology are discussed in Section 4.2.1. The prototype music environments allow people working on separate computer workstations to interact via a shared graphical interface. The software allows groups of people to collectively make 'electronica' style music together using multiple networked computers. Users share an onscreen workspace which is mirrored across all computer screens, and which is based on the 'What you see is what I see' (WYSIWIS) abstraction for collaborative desktop interfaces discussed in Section 2.2.3. Users compose by creating 'virtual instruments' within an on-screen workspace which is shared across all currently connected computers. The virtual instruments provided in the software are represented as small windows containing various controls, and the instruments can be used to create rhythmic parts such as drum beats, as well as melodies, bass lines and ambient textures. Later versions of the software (such as used in Chapters 6 and 7) allow the virtual instruments to be routed through audio effects to add sonic variety, complexity and richness.

Participants for each study were recruited with the requirement of having musical knowledge, experience making music or an understanding of music technologies. In total, one hundred and five participants took part in studies over the course of the research. While the three studies looked at distinct aspects of interface design for collaborative musical software, they each held a number of common properties:

- Three people per session
- Participants with musical background
- Participants were co-located

- Interaction was co-temporal
- Interaction via a shared software workspace running on desktop computers
- The music software participants used had similar sounds and musical possibilities

### 9.1.1 Study One: Privacy and Awareness (Chapter 5)

Study One investigated the effects of introducing privacy into a collaborative music making environment. Three interface variations were used, one in which all music was shared and public, one interface which allowed participants to work on musical contributions in privacy via a Private space, and one in which participants were able to listen to the contributions held in each others' private spaces via an additional awareness mechanism.

Nine groups of three participants interacted with all three interface designs. A range of quantitative measures were automatically collected during the interaction to study how the different interface variations affected the way the participants used the software. The participants also completed a questionnaire and took part in a group interview. Quantitative analysis showed that when the Private Space was made available to participants, they created almost all their contributions in it, before sharing them with the rest of the group in the Public Space. The inclusion of the additional awareness mechanism did not affect the way participants used the software, although it did alter the reported quality of the interaction. The participants indicated that the additional awareness it provided was useful for formulating contributions, however some participants found it difficult to keep in mind what each person was doing in their Private Space.

The study pointed to several avenues of future work; the ability for individual users to mute the spaces or define their own auditory representation of the shared interface, the use of passive awareness mechanisms to push information to users, and the use of different audio delivery devices to separate the audio from different workspaces.

### 9.1.2 Study Two: Audio Delivery (Chapter 6)

Study Two looked at the way audio could be delivered to users of a collaborative environment as a potential means of presenting awareness information. A three condition experimental design was used, where each condition delivered audio using a different combination of speakers and headphones.

Ten groups of three participants took part in the study. Analysis showed that when audio was delivered using only speakers, participants made less use of their Personal audio channel. Analysis was performed on the way participants visually and spatially organised the shared workspace, leading to the identification of territorial behaviour among around half the groups of participants. Participants in the more territorial groups were less inclined to edit contributions made by other group members. This form of spatial organisation was not affected by the way audio was delivered, and groups often used a similar spatial layout in each of the conditions they were presented with.

Interviews and discussion between participants suggested that playing audio through a combination of headphones and speakers was not effective in supporting collaboration due to disrupted awareness between participants. More generally, the way audio was presented did not engender awareness of who was doing what (Authorship Awareness), although some participants noted using the personal audio channel as a means of determining which sounds were created by which interface element (Source Awareness). Combined, these findings suggested that additional awareness should be provided by the interface to scaffold understanding of where musical sounds were originating, and who had created them.

### 9.1.3 Study Three: Identification Mechanisms (Chapter 7)

Study Three identified two concepts, Source Awareness and Authorship Awareness. These referred to notions of which interface component was creating a particular sound (Source Awareness), and awareness of who was responsible for creating a particular contribution (Authorship Awareness). Mechanisms to support the gathering of Source and Authorship awareness were implemented in the software. These mechanisms provided users with the ability to aurally solo and visually highlight specific musical contributions within the interface (Module Soloing), and the ability to aurally solo and visually highlight all the musical contributions by a particular user or set of users (Client Soloing).

A four condition study was conducted to investigate how these mechanisms effected the way participants used with the software during musical collaboration. Sixteen groups of three participants were recruited. All participants used two interfaces; the first included a feature which enabled them to solo the contributions made by a particular user or set of users (including themselves). The second interface excluded this feature. In addition, half the groups were given an additional interface feature which enabled them to solo specific musical contributions within the interface regardless of who created them.

A key finding was that for participants whose interface included the ability to solo any music module, the addition of the Client Soloing feature reduced the likelihood of participants editing contributions made by other users. This was identified through analysis of the interaction logs, and also through trends in the participant self reported data which were statistically significant. The results also showed that the ability to solo specific contributions within the interface regardless of who created them (the Module Soloing feature) had no measurable effects on the way the participants used the software, although the presence or absence of Module Soloing appeared to influence the effect of the Client Soloing feature. Interfaces which included the additional awareness mechanisms were regularly preferred by participants.

## 9.2 Limitations

This section acknowledges several key limitations to the contributions of this thesis, and where appropriate proposes future work which might address these limitations. The final section of this chapter then goes on to discuss future work more broadly.

### 9.2.1 Participant Groups

Although section 4.1.2 made a strong justification for the recruitment process adopted within this research, there are several drawbacks to using groups of participants who have not previously worked together. In particular, issues of social anxiety, evaluation apprehension and conflicting musical taste may have inhibited the participants. Future studies could investigate this through the observation of groups of musicians who have previously worked together. Studying smaller or larger groups would be another interesting avenue of exploration. The studies in this thesis all used groups of three participants, however the software developed to run the studies can support a much larger number of simultaneous users.

### 9.2.2 Participant Personalty Types

Personality type has been previously identified as influencing musical preferences (Rentfrow and Gosling, 2003) and has been discussed in relation to playing style, musical aptitude and commitment to learning (Marcus, 2012). Within the context of this work, a ‘Myers-Briggs Type Indicator’ (Quenk, 2009) or similar measure of personality type may have been a useful tool in gathering information about the participants used for the studies, and may have provided additional insight into the personality makeup of the groups under observation. Such tests attempt to categorise the personality type of an individual

However, for primarily practical reasons, personality type was not controlled for or factored into the experimental designs within this thesis. Crucially, the participant recruitment process made it difficult to know what type of people would be arriving to participate prior to a study session taking place. This was for the most part due to the ethical constraints of using e-mail communication to contact participants. Such lack of control made it impossible to systematically organise groups of participants featuring specific combinations of personality types without collecting such information prior to running the session, and attempting to control for personality type in this way would have made the recruitment and session organisation significantly more complex and time consuming. For practical reasons participants were instead placed into groups as and when they contacted the researcher to take part in the study. Furthermore, ethical and financial constraints limited the duration of the study sessions, and collecting personality type information would therefore have caused there to be less time for collecting interaction data or running interviews.

Future studies could investigate the influence of personality type within CDMI. Particularly interesting areas to explore would include the way groups with different combinations of personalty types assign roles, discuss changes within the music, and reach decisions about the musical style they intended to explore. Studies could also investigate the use of awareness mechanisms by people of different personality types. For instance, it may be the case that people who are classified as being more introvert may be more likely to use awareness mechanisms within the interface than to verbally ask other people in the group for information. Such studies remain as future work.

### 9.2.3 First-time user experience

A third limitation of this research concerns the experience of the participants with the software itself. Without exception, all the participants experienced the software for the first time when taking part in the study. This step was justified in Section 4.2.1 as a means of controlling for familiarity with the software by ensuring that none of the participants has previous experience using it. Although the participants were given training and free time to explore the software before being presented with the experiment conditions, this feature of the experimental design does limit the scope of the findings, as they are based on first-time experience. As noting in Section 6.13 this may have particular implications for aspects of the software which the participants found confusing or difficult to use, such as attempting to use an unfamiliar combination of headphones and speakers, and as users gain increasing experience with the software, these issues may become less of a concern. One possible means of addressing this issue would have been to run longer experiment sessions.

### 9.2.4 Study Duration

The studies were of a short duration, and although condition ordering was randomised, the short time in which participants used the software may have exaggerated ordering effects, as participants may have still been learning the features of the software during the earlier experiment conditions. Future studies could address this issue by providing more time for participants to become familiar with the software, and providing more time for observed use during the conditions themselves. Financial reimbursement for the longer scale studies would clearly require a larger budget than was available for this research.

### 9.2.5 Software Limitations

The outcomes of this thesis are aimed primarily at the future development of tools to support experienced users in collaborative musical interaction. To this end, the software used in all the studies borrowed from the metaphors of existing music production environments, and used terminology which would be familiar to advanced users. However at the same time, the software presented to participants was designed to support group interaction for the short duration of the user studies. Future studies could explore collaboration using more advanced software environments. This approach is essential if research into musical collaboration is to move forward from this point and begin to address how software can be designed to support advanced and experienced users in working together. Future studies could continue in this direction through the inclusion of more advanced functionality and features which are in keeping with the options provided by real-world music applications, for instance the inclusion of wider sound pallets, and support for plug-in effects (thus allowing users to create music with familiar tools. Such studies would likely require a much larger training period due to the additional complexities of the environment.

### 9.2.6 Musical Style

The software environment used for the studies in this thesis was primarily suited to the creation of electronic music. In all studies the interface featured virtual instruments such as sequencers with on-screen controls to determine their musical output and sonic characteristics. Such a design would have been familiar to users of software such as Reason[1], whist the software in Studies Two and Three also featured a modular patching metaphor which might be familiar to users of programs such as Max/MSP[2], although it may not have been as easy to grasp for people who were more experienced with sequencing environments such as ProTools[3].

Due to these factors, it could be argued that the findings of this thesis apply only to the design of interfaces for electronic music, and in particular modular environments which are designed for the creation and layering of looping musical phrases. If this argument is to be accepted, the findings of this thesis still apply to a wide range of potential music software applications. For instance Section 4.2 highlights a number of existing commercial software environments which employ similar interaction metaphors to the software used in this thesis.

Moving forward, it is important to restate that this thesis set out to study the design and evaluation of *awareness mechanisms* within collaborative music software, rather than investigating the design of features specific a certain musical practice or particular musical style. As such, the design implications related to awareness are not tied specifically to the modular form of music creation which was adopted for the studies. In particular, the way awareness information is presented, concerns for privacy within the interface, and the way personal and public audio is delivered are likely to be pertinent subjects for a variety of collaborative music environments, regardless of how the music is actually represented or created. These concerns could apply equally to software designed for the creation of music which deviates from the loop-orientated styles associated with the software used in the experiments. For instance a multi-user DAW which allows musical contributions from different people to be placed on a timeline might still need to provide an indication of which sounds had been contributed by who, and where users are co-located, the way in which audio is presented would be an equally important factor to consider.

This thesis is however concerned with the design of software interfaces which allow tightly interwoven collaboration between users, and as such, findings may not be applicable to practices such as live electronics, which might employ a range of computer based technologies, combined with electronic devices and acoustic instruments. Performers engaged in such practices will undoubtably rely on awareness information, however due to the wide range of possible setups

[1] http://www.propellerheads.se/products/reason/ Last Accessed 15 October
[2] http://cycling74.com/products/max/ Last Accessed 15 October
[3] http://www.avid.com/US/products/family/pro-tools Last Accessed 15 October 2012

and technologies involved, it is difficult to propose specific implications for these scenarios. Merritt et al. (2010) propose and evaluate a visual awareness system designed for these types of situations, however there is arguably scope for more innovation in this area.

As the number of musicians worldwide, and as computers become increasingly integrated into music production the scope and impact of collaborative computer music software will only increase. Future research could look in more detail at systems which support the creation of a broader range of musical styles and genres, and delve more deeply into the relationship between interface metaphors and awareness mechanisms. Related concerns in this case would also be the use of physical controllers and live instruments.

## 9.3 Future Work

The previous section identified future work in relation to the limitations of this thesis. This section considers future work more generally, and points to avenues of research which are beyond the scope of this thesis.

### 9.3.1 Touch Interaction, Tablets and Studies of Other Modalities

Alongside sound, haptic feedback is an under-explored modality within CSCW research, and touch based interaction is an expanding field of research. The design implications suggested for other devices and platforms (Section 8.5) require evaluation to assess their merit. Such research may provide interesting opportunities for the presentation of awareness information in future CDMI systems. The use of non-aural modalities such as haptic feedback has not been investigated within this thesis, and remains a subject for future work.

### 9.3.2 Multiple Device Collaboration

The participants within this thesis all used identical computers, however in the real-world, people have widely different computing devices and hardware setups. For instance ranging from desktop computers to tablets and mobile phones. Future studies could investigate the affordances of systems which support collaboration over a range of different devices.

### 9.3.3 Interaction Between Privacy and Congruence

Study Three (Chapter 7) explored the presentation of incongruent audio to participants, however it did this without incorporating a privacy feature. A future study could explore the interaction between these two issues.

### 9.3.4 Remote Interaction

The studies conducted for this thesis were all based on co-located interaction. A clear avenue for future work would be to evaluate the success of the current software environment when used remotely. This would likely lead to a number of usability issues related to the absence of awareness which collocated users would take for granted. Such a study would therefore be highly informative for designers of both remote and co-located systems.

*Digital Rights Management*

The use of shared interfaces for musical collaboration and creativity potentially opens up issues surrounding copyright and ownership. Such concerns may require legal scrutiny and clarification.

*Performance*

The systems developed for this research have a clear application in musical performance, a context in which groups of musicians frequently engage in real-time musical interaction. Collaborative computer supported musical performance has been addressed in laptop ensemble and laptop performance projects (Collins et al., 2003, Dannenberg et al., 2007, Trueman et al., 2006, Wang et al., 2009), as well as earlier projects such as The Hub (Gresham-Lancaster, 1998), although there are few examples of research focusing on the study of collaborative performance systems in use, or concerning the design of awareness mechanisms for live performance within ensembles of computer musicians.

In a performance context, where music is being created in front of an audience, rather than in a more relaxed studio environment, features such as privacy may be especially important, as a means of testing and cueing musical contributions before they are made audible to an audience. In addition, on-stage monitoring may require particular attention, as it may be necessary for musicians to hear a mixture of audio which is private to them, sounds which are shared with the group, and sounds which are being broadcast to the audience.

Concerning live performance, designers must not only consider awareness and communication between the musicians, but also their relationship with the audience, and the ways in which an audience might relate what they are seeing on stage to the sounds being created. In the case of computer interfaces which abstract the gestures and actions of the musicians away from the sonic results, such a system might include awareness mechanisms designed specifically to aid the audience in understanding what is happening. CDMI for live performance is clearly a complex issue, and therefore studying performance, performers, audiences and performer-audience interaction has been left as future work.

*Education*

A number of participants expressed an interest in using the software as an educational or classroom tool. This is another interesting and practical application of the software development conducted as part of this thesis. Within an educational context, studies based on awareness and privacy mechanisms could be performed to explore creativity in children, or to teach aspects of team-working at an early age, and there are already a number of publications discussing the use of collaborative music software by school-age children (Bryan-Kinns, 2004, Truman, 2011a, Weinberg et al., 2002).

*Collaboration with Artificial Agents*

Previous work has demonstrated the ability for artificial agents to collaborate together without human intervention (Bown et al., 2009). Such generative processes and live algorithms for computer improvisation (Blackwell and Bentley, 2002) are beyond the scope of this thesis, although incorporating autonomous musical agents into a collaborative environment could be an exciting avenue of future exploration from both creative and research perspectives. It would, for example, be possible to conduct studies where human and artificial collaborators worked together. In this context, a deeper understanding of how human musicians collaborate could inform the design of new intelligent agents.

## Closing Remarks

The potential for future work outlined above suggests that the field of CDMI is still in an exciting stage of development. It is hoped that the findings of this thesis will aid the development of new and exciting technological advancements for the musicians and composers of the future.

# Bibliography

Demosthenes Akoumianakis, George Vellis, Ioannis Milolidakis, Dimitrios Kotsalis, and Chrisoula Alexandraki. Distributed collective practices in collaborative music performance. In *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*, DIMEA '08, pages 368–375, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-248-1. doi: http://doi.acm.org/10.1145/1413634.1413700. URL http://doi.acm.org/10.1145/1413634.1413700.

Christopher Alexander. *A Pattern Language: Towns, Buildings, Construction*. OUP USA, 1978.

Chrisoula Alexandraki and Ioannis Kalantzis. Requirements and application scenarios in the context of network based music collaboration. In *Proceedings of the AXMEDIS 2007 Conference. i-maestro workshop*, Barcelona, Spain, 2007.

David L. Altheide. Ethnographic content analysis. *Qualitative Sociology*, 10(1), 1987.

Kazuhiro Jo Adam Parkinson Atau Tanaka, Andreia Cavaco. Creative workshopping methodologies: Its like musicking! In *Proceedings of CHI 2012*, 2012.

Curtis Bahn. Physicality and feedback: a focus on the body in the performance of electronic music. In *Proceedings of the International Computer Music Conference*, pages 44–51, 2001.

Curtis Bahn and Dan Trueman. Interface - electrionc chamber ensemble. In *HI 2001 Workshop on New Interfaces for Musical Expression*, Seattle, 2001.

Liam J. Bannon and Kjeld Schmidt. *CSCW: four characters in search of a context*, pages 3–16. North-Holland Publishing Co., Amsterdam, The Netherlands, The Netherlands, 1991. ISBN 0-444-88811-X. URL http://dl.acm.org/citation.cfm?id=117730.117731.

Álvaro Barbosa. Displaced soundscapes: A survey of network systems for music and sonic art creation. *Leonardo Music Journal*, 13:53–59, 2003.

Álvaro Mendes Barbosa. *Computer-Supported Cooperative Work for Music Applications*. PhD thesis, Universitat Pompeu Fabra, 2006.

Stuart A Battersby and Patrick G T Healey. Head and hand movements in the orchestration of dialogue. *Communication*, pages 1998–2003, 2008. URL http://www.eecs.qmul.ac.uk/~stuart/battersby_healey_cogsci_2010.pdf.

Steve Benford, John Bowers, Lennart E. Fahlén, and Chris Greenhalgh. Managing mutual awareness in collaborative virtual environments. In *VRST '94: Proceedings of the conference on Virtual reality software and technology*, pages 223–236, River Edge, NJ, USA, 1994. World Scientific Publishing Co., Inc. ISBN 981-02-1867-2.

Hrvoje Benko, Andrew D. Wilson, and Patrick Baudisch. Precise selection techniques for multi-touch screens. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1263 – 1272. Montréal, Québec, Canada, 2006.

Ilias Bergstrom and R. Beau Lotto. Harnessing the enactive knowledge of musicians to allow the real-time performance of correlated music and computer graphics. *Leonardo*, 42(1), 2009.

Zafer Bilda, Ernest Edmonds, and Linda Candy. Designing for creative engagement. *Design Studies*, 29(6):525 – 540, 2008. ISSN 0142-694X. doi: DOI:10.1016/j.destud.2008.07.009. Interaction Design and Creative Practice.

Tim M. Blackwell and Peter Bentley. Improvised music with swarms. In *Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on*, volume 2, pages 1462 –1467, 2002. doi: 10.1109/CEC.2002.1004458.

Tina Blaine and Sidney Fels. Contexts of collaborative musical experiences. In *NIME '03: Proceedings of the 2003 conference on New interfaces for musical expression*, pages 129–134, Singapore, Singapore, 2003. National University of Singapore.

Tina Blaine and Clifton Forlines. Jam-o-world: evolution of the jam-o-drum multi-player musical controller into the jam-o-whirl gaming interface. In *NIME '02: Proceedings of the 2002 conference on New interfaces for musical expression*, pages 1–6, Singapore, Singapore, 2002. National University of Singapore. ISBN 1-87465365-8.

Tina Blaine and Tim Perkis. The jam-o-drum interactive music system: a study in interaction design. In *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, DIS '00, pages 165–173, New York, NY, USA, 2000.

ACM. ISBN 1-58113-219-0. doi: http://doi.acm.org/10.1145/347642.347705. URL http://doi.acm.org/10.1145/347642.347705.

Jan O. Borchers. Worldbeat: designing a baton-based interface for an interactive music exhibit. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 131–138, New York, NY, USA, 1997. ACM. ISBN 0-89791-802-9. doi: http://doi.acm.org/10.1145/258549.258632.

Pierre Boulez. Technology and the composer. *Leonardo*, 11(1):59–62, 1978.

John Bowers. Improvising machines: Ethnographically informed design for improvised electro-acoustic music. Master's thesis, Masters in Music by Research, University of East Anglia, Norwich, UK, 2002.

Oliver Bown, Alice Eldridge, and Jon McCormack. Understanding interaction in contemporary digital music: from instruments to behavioural objects. *Organised Sound*, 14(2), 2009.

Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

Albert Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, 1990.

Harry Brignall and Yvonne Rogers. Enticing people to interact with large public displays in public spaces. In *Proceedings of INTERACT'03*, pages 17–24, 2003.

Andrew R. Brown. Network jamming: Distributed performance using generative music. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, 2010.

Nick Bryan-Kinns. Daisyphone: The design and impact of a novel environment for remote group music improvisation. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 135–144, 2004.

Nick Bryan-Kinns. Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, 2012. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2012.02.004. URL http://www.sciencedirect.com/science/article/pii/S1071581912000328.

Nick Bryan-Kinns and Fraser Hamilton. Identifying mutual engagement. *Behaviour & Information Technology*, 2009.

Nick Bryan-Kinns, Patrick G. T. Healey, and Joe Leach. Exploring mutual engagement in creative collaborations. In *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, C&C '07, pages 223–232, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-712-4. doi: http://doi.acm.org/10.1145/1254960.1254991. URL http://doi.acm.org/10.1145/1254960.1254991.

Philip L. Burk. Jammin' on the web - a new client/server architecture for multi-user musical performance. In *Proceedings of the ICMC*, 2000.

Linda Candy. Constraints and creativity in the digital arts. *Leonardo*, 40(4), 2007.

Peter H. Carstensen and Kjeld Schmidt. Computer supported cooperative work: New challenges to systems design. In Kenji Itoh, editor, *Handbook of Human Factors*, pages 619–636. Asakura Publishing, 1999.

Kim Cascone. The aesthetics of failure: "post-digital" tendencies in contemporary computer music. *Computer Music Journal*, 24(4):12–18, 2000.

Joel Chadabe. *Electric Sound: the Past and Promise of Electronic Music*. Pearson, 1996.

Kathy Charmaz. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. SAGE Publications Ltd, 2006.

Herbert H. Clark and Susan E. Brennan. Grounding in communication. In L.B. Resnick, J.M. Levine, and S.D. Teasley, editors, *Perspectives on socially shared cognition*. APA Books, Washington, 1991.

Eric F. Clarke. Making and hearing meaning in performance. *The Nordic Journal of Aesthetics*, 18(33-34), 2006.

Andrew L. Cohen. Awareness, planning and joint attention in collaborative writing: From fieldwork to design. Technical report, IBM Watson Research Center, 2008.

Nick Collins, Alex McLean, Julian Rohrhuber, and Adrian Ward. Live coding in laptop performance. *Organised Sound*, 8(3):321–330, 2003.

Perry Cook. Principles for designing computer music controllers. In *Proceedings of the 2001 conference on New interfaces for musical expression*, NIME '01, pages 1–4, Singapore, Singapore, 2001. National University of Singapore. URL http://dl.acm.org/citation.cfm?id=1085152.1085154.

Gregory W. Corder and Dale I. Foreman. *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. Wiley, 2009.

Brigid Costello and Ernest Edmonds. A study in play, pleasure and interaction design. In *DPPI '07: Proceedings of the 2007 conference on Designing pleasurable products and interfaces*, pages 76–91, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-942-5. doi: http://doi.acm.org/10.1145/1314161.1314168.

Tim Coughlan and Peter Johnson. Support for creativity in musical collaboration and creator / technologist interaction. In *Position Paper for CHI 2007 Workshop: HCI and New Media Arts*, 2007.

Ian Cross. Music and meaning, ambiguity and evolution. In Dorothy Meill, Raymond MacDonald, and David J. Hargreaves, editors, *Musical Communication*. Oxford University Press, 2005.

Mihalyi Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper & Row, 1990.

Roger B. Dannenberg, Sofia Cavaco, Eugene Ang, Igor Avramovic, Barkin Aygun, Jinwook Baek, Eric Barndollar, Daniel Duterte, Jeffrey Grafton, Robert Hunter, Chris Jackson, Umpei Kurokawa, Daren Makuck, Timothy Mierzejewski, Michael Rivera, Dennis Torres, and Apphia Yu. The Carnegie Mellon Laptop Orchestra. In *Proceedings of the 2007 International Computer Music Conference*, San Francisco, 2007.

Raymond Macdonald David J. Hargreaves and Dorothy Meill. How do people communicate using music. In Dorothy Meill, Raymond MacDonald, and David J. Hargreaves, editors, *Musical Communication*. Oxford University Press, 2005.

Philip L. Davidson and Jefferson Y. Han. Synthesis and control on large scale multi-touch sensing displays. In *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06),*, Paris, France, 2005.

Hugh Davies. Gentle fire: An early approach to live electronic music. *Leonardo Music Journal*, 2001.

José Díaz, Juan Márquez, Miguel Sánchez, José Sánchez-Aguilera, Miguel Sánchez, and Javier Bajo. Diami: Distributed intelligent environment for blind musicians. In Sigeru Omatu, Miguel Rocha, José Bravo, Florentino Fernández, Emilio Corchado, Andrés Bustillo, and Juan Corchado, editors, *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*, volume 5518 of *Lecture Notes in Computer Science*, pages 475–482. Springer Berlin / Heidelberg, 2009.

Christopher Dobrian and Daniel Koppelman. The 'E' in NIME: musical expression with new computer interfaces. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, pages 277–282, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3.

Paul Dourish and Victoria Bellotti. Awareness and coordination in shared workspaces. In *CSCW '92: Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 107–114, New York, NY, USA, 1992. ACM. ISBN 0-89791-542-9. doi: http://doi.acm.org/10.1145/143457.143468.

Steven P. Dow, Blair MacIntyre, and Michael Mateas. Understanding engagement: A mixed-method approach to observing game play. In *Conference on Computer-Human Interaction (CHI'08) Workshop on Evaluating User Experience in Games*, Florence, Italy, 2008.

Matthew Duignan, James Noble, Pippin Barr, and Robert Biddle. Metaphors for electronic music production in reason and live. In Masood Masoodian, Steve Jones, and Bill Rogers, editors, *Computer Human Interaction*, volume 3101 of *Lecture Notes in Computer Science*, pages 111–120. Springer Berlin / Heidelberg, 2004.

Mica R. Endsley. Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1):65–84, 1995.

Mica R. Endsley. Theoretical underpinnings of situation awareness: A critical review. In Mica R. Endsley and Daniel J. Garland, editors, *Situation Awareness Analysis and Measurement*, pages 3–32. Lawrence Erlbaum Associates, 2000.

Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. G*power 3:a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 2007.

Sidney Fels. Designing for intimacy: creating new interfaces for musical expression. *Proceedings of the IEEE*, 92(4):672 – 685, apr 2004. ISSN 0018-9219. doi: 10.1109/JPROC.2004.825887.

Sidney Fels and Florian Vogt. Tooka: explorations of two person instruments. In *Proceedings of the 2002 conference on New interfaces for musical expression*, 2002.

Sidney Fels, Linda Kaastra, Sachiyo Takahashi, and Graeme McCaig. Evolving Tooka: from experiment to instrument. In *NIME '04: Proceedings of the 2004 conference on New interfaces for musical expression*, pages 1–6, Singapore, Singapore, 2004. National University of Singapore.

Robin Fencott. Interactive music using multi-touch cellular automata. In *(re)Actor3 the third international conference on digital live art*, Liverpool, September 2008.

Robin Fencott and Nick Bryan-Kinns. Sensory Threads: Sonifying imperceptible phenomena in the wild. In *Proceedings of the 6th Sound and Music Computing Conference*, 2009.

Robin Fencott and Nick Bryan-Kinns. Hey man, You're invading my personal space! Privacy and awareness in collaborative music. In *Proceedings of the 10th International Conference on New Interfaces for Musical Expression (NIME 2010)*, Sydney, Australia, 2010.

Robin Fencott and Nick Bryan-Kinns. 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 '12, pages 69–78, Swinton, UK, UK, 2012. British Computer Society. URL http://dl.acm.org/citation.cfm?id=2377916.2377925.

Robin Fencott and Nick Bryan-Kinns. HCI, CSCW and Collaborative Digital Musical Interaction. In Simon Holland, Katie Wilkie, Paul Mulholland, and Allan Seago, editors, *Music and Human-Computer Interaction*, Springer Cultural Computing Series. Springer, 2013.

Robin Fencott and John Dack. An interactive surface realisation of Henri Pousseur's 'Scambi'. In *SMC Conference 2011*, 2011.

Geraldine Fitzpatrick, Simon Kaplan, and Tim Mansfield. Applying the locales framework to understanding and designing. In *Proceedings of the Australasian Conference on Computer Human Interaction*, OZCHI '98, pages 122–, Washington, DC, USA, 1998. IEEE Computer Society. ISBN 0-8186-9206-5. URL http://dl.acm.org/citation.cfm?id=552884.825513.

Jason Freeman. Technology, real-time notation, and audience participation in Flock. In *Proceedings of the International Computer Music Conference*, 2008.

Dominic Furniss and Ann Blandford. Understanding emergency medical dispatch in terms of distributed cognition: a case study. *Ergonomics Journal*, 2006. URL http://eprints.ucl.ac.uk/5114/.

Dominic Furniss, Ann Blandford, and Paul Curzon. Confessions from a Grounded Theory PhD: Experiences and lessons learnt. In *CHI*, Vancouver, BC, Canada, 2011.

Carrie Gates, Sriram Subramanian, and Carl Gutwin. Djs' perspectives on interaction and awareness in nightclubs. In *DIS '06: Proceedings of the 6th conference on Designing Interactive systems*, pages 70–79, New York, NY, USA, 2006. ACM. ISBN 1-59593-367-0. doi: http://doi.acm.org/10.1145/1142405.1142418.

Andrew Gerzso. Paradigms and computer music. *Leonardo Music Journal*, 2(1):73–79, 1992.

Barney G. Glaser and Anselm L. Strauss. *The discovery of grounded theory: strategies for qualitative research*. Aldine, Chicago, 1967.

Maya Gratier. Grounding in musical interaction: Evidence from jazz performances. *Musicae Scientiae*, 12:71–110, March 2008.

Saul Greenberg. A fisheye text editor for relaxed-WYSIWIS groupware. In *Conference companion on Human factors in computing systems: common ground*, CHI '96, pages 212–213, New York, NY, USA, 1996. ACM. ISBN 0-89791-832-0. doi: http://doi.acm.org/10.1145/257089.257285. URL http://doi.acm.org/10.1145/257089.257285.

Saul Greenberg, Carl Gutwin, and Mark Roseman. Semantic telepointers for groupware. *Computer-Human Interaction, Australasian Conference on*, 0:0054, 1996. doi: http://doi.ieeecomputersociety.org/10.1109/OZCHI.1996.559988.

Judith Greene and Manuela D'Oliveira. *Learning to Use Statistical Tests in Psychology (Open Guides to Psychology)*. Open University Press, 1999.

Scot Gresham-Lancaster. The aesthetics and history of the hub: The effects of changing technology on network computer music. *Leonardo Music Journal*, 8:39–44, 1998.

Michael Gurevich. Jamspace: designing a collaborative networked music space for novices. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, pages 118–123, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3.

Michael Gurevich and Cavan A. Fyans. Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound*, 16:166 –175, 2011.

Carl Gutwin and Saul Greenberg. Design for individuals, design for groups: tradeoffs between power and workspace awareness. In *CSCW '98: Proceedings of the 1998 ACM conference on Computer supported cooperative work*, pages 207–216, New York, NY, USA, 1998. ACM. ISBN 1-58113-009-0. doi: http://doi.acm.org/10.1145/289444.289495.

Carl Gutwin and Saul Greenberg. The effects of workspace awareness support on the usability of real-time distributed groupware. *ACM Trans. Comput.-Hum. Interact.*, 6(3):243–281, 1999. ISSN 1073-0516. doi: http://doi.acm.org/10.1145/329693.329696.

Carl Gutwin and Saul Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work*, 11(3):411–446, 2002. ISSN 0925-9724. doi: http://dx.doi.org/10.1023/A:1021271517844.

Carl Gutwin and Reagan Penner. Improving interpretation of remote gestures with telepointer traces. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*, CSCW '02, pages 49–57, New York, NY, USA, 2002. ACM. ISBN 1-58113-560-2. doi: 10.1145/587078.587086. URL http://doi.acm.org/10.1145/587078.587086.

Carl Gutwin, Reagan Penner, and Kevin Schneider. Group awareness in distributed software development. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, CSCW '04, pages 72–81, New York, NY, USA, 2004. ACM. ISBN 1-58113-810-5. doi: 10.1145/1031607.1031621. URL http://doi.acm.org/10.1145/1031607.1031621.

Carl Gutwin, Oliver Schneider, Robert Xiao, and Stephen Brewster. Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, CSCW '11, pages 85–94, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0556-3. doi: http://doi.acm.org/10.1145/1958824.1958838. URL http://doi.acm.org/10.1145/1958824.1958838.

Rikke Hadrup, Pelle Svane Jakobsen, Mathilde Schytz Juul, Dan Lings, and Ásta Olga Magnúsdóttir. Designing an auditory w-lan based game. In Karenza Moore and Jason Rutter, editors, *MOBILE ENTERTAINMENT: User-centred Perspectives*. ESRC Centre for Research on Innovation and Competition, 2004.

Patric G T. Healey, Joe Leach, and Nick Bryan-Kinns. Inter-play: Understanding group music improvisation as a form of everyday interaction. In *Proceedings of Less is More — Simple Computing in an Age of Complexity*. Microsoft Research Cambridge, 2005.

Patrick G. T. Healey and Charlotte R. Peters. The conversational organisation of drawing. *Pen-Based Learning Technologies, International Workshop on*, 0:1–6, 2007. doi: http://doi.ieeecomputersociety.org/10.1109/PLT.2007.25.

Christian Heath, Marina Jirotka, Paul Luff, and Jon Hindmarsh. Unpacking collaboration: the interactional organisation of trading in a city dealing room. *Comput. Supported Coop. Work*, 3(2):147–165, 1995. ISSN 0925-9724. doi: http://dx.doi.org/10.1007/BF00773445.

Christian Heath, Paul Luff, Dirk Vom Lehn, Jon Hindmarsh, and Jason Cleverly. Crafting participation: designing ecologies, configuring experience. *Visual Communication*, 1(1):9–33, 2002a.

Christian Heath, Marcus Sanchez Svensson, Jon Hindmarsh, Paul Luff, and Dirk vom Lehn. Configuring awareness. *Comput. Supported Coop. Work*, 11(3):317–347, 2002b. ISSN 0925-9724. doi: http://dx.doi.org/10.1023/A:1021247413718.

Jim Hollan and Scott Stornetta. Beyond being there. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '92, pages 119–125, New York, NY, USA, 1992. ACM. ISBN 0-89791-513-5. doi: 10.1145/142750.142769. URL http://doi.acm.org/10.1145/142750.142769.

Simon Holland. Artificial intelligence in music education: a critical review. In E. Miranda, editor, *Readings in Music and Artificial Intelligence*, pages 239–274. Harwood Academic Publishers, Amsterdam, 1999.

Eva Hornecker and Matthias Stifter. Learning from interactive museum installations about interaction design for public settings. In *OZCHI '06: Proceedings of the 18th Australia conference on Computer-Human Interaction*, pages 135–142, New York, NY, USA, 2006. ACM. ISBN 1-59593-545-2. doi: http://doi.acm.org/10.1145/1228175.1228201.

Eva Hornecker, Paul Marshall, Nick Sheep Dalton, and Yvonne Rogers. Collaboration and interference: awareness with mice or touch input. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, CSCW '08, pages 167–176, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-007-4. doi: http://doi.acm.org/10.1145/1460563.1460589. URL http://doi.acm.org/10.1145/1460563.1460589.

William Hsu and Marc Sosnick. Evaluating interactive music systems: An hci approach. In *NIME '09: Proceedings of the 2009 conference on New interfaces for musical expression*, 2009.

Scott E. Hudson and Ian Smith. Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, CSCW '96, pages 248–257, New York, NY, USA, 1996. ACM. ISBN 0-89791-765-0. doi: 10.1145/240080.240295. URL http://doi.acm.org/10.1145/240080.240295.

Joh Hughes, Dave Randall, and Dan Shapiro. CSCW: discipline or paradigm? a sociological perspective. In *Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work*, pages 309–323, Norwell, MA, USA, 1991. Kluwer Academic Publishers. ISBN 0-7923-1439-5. URL http://portal.acm.org/citation.cfm?id=1241910.1241933.

John Hughes, Val King, Tom Rodden, and Hans Andersen. Moving out from the control room: ethnography in system design. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, CSCW '94, pages 429–439, New York, NY, USA, 1994. ACM. ISBN 0-89791-689-1. doi: http://doi.acm.org/10.1145/192844.193065. URL http://doi.acm.org/10.1145/192844.193065.

Andy Hunt, Marcelo M. Wanderley, and Ross Kirk. Towards a model for instrumental mapping in expert musical interaction. In *International Computer Music Conference*, 2000.

Charles Céleste Hutchins. *Portfolio of Compositions*. PhD thesis, University of Birmingham, 2011.

Edwin Hutchins. *Cognition in the Wild*. MIT Press, 1996.

Hiroshi Ishii, Minoru Kobayashi, and Jonathan Grudin. Integration of interpersonal space and shared workspace: Clearboard design and experiments. *ACM Trans. Inf. Syst.*, 11:349–375, October 1993. ISSN 1046-8188. doi: http://doi.acm.org/10.1145/159764.159762. URL http://doi.acm.org/10.1145/159764.159762.

Toshio Iwai. Composition on the table. In *ACM SIGGRAPH 99 Electronic art and animation catalog*, SIGGRAPH '99, pages 10–, New York, NY, USA, 1999. ACM. ISBN 1-58113-104-6. doi: http://doi.acm.org/10.1145/312379.312383.

David James and Jeffrey Stanton. Beyond being (t)here: the social and personal implications of making music at a distance. In *Proceedings of the 2011 iConference*, iConference '11, pages 686–687, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0121-3. doi: http://doi.acm.org/10.1145/1940761.1940870. URL http://doi.acm.org/10.1145/1940761.1940870.

Vera John-Steiner. *Creative Collaboration*. Oxford University Press, USA, 2000.

S. Jordà. *Digital Lutherie Crafting musical computers for new musics' performance and improvisation*. PhD thesis, Ph.D. Dissertation. UPF, 2005a.

Sergi Jordà. Multi-user instruments: Models, examples and promises. In *Proceedings of the 2005 conference on New Interfaces for Musical Expression (NIME'05*, pages 23–26, 2005b.

Sergi Jordà, Martin Kaltenbrunner, Günter Geiger, and Ross Bencina. The reactable. In *Proceedings of the International Computer Music Conference*, 2005.

Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 139–146, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-619-6. doi: http://doi.acm.org/10.1145/1226969.1226998.

Kari Kallinen and Niklas Ravaja. Comparing speakers versus headphones in listening to news from a computer - individual differences and psychophysiological responses. *Computers in Human Behavior*, 23(1):303 – 317, 2007. ISSN 0747-5632. doi: DOI:10.1016/j.chb.2004.10.014. URL http://www.sciencedirect.com/science/article/B6VDC-4DTTBC0-H/2/27db6b8d29042c2e80e0198f1084c6d9.

Chris Kiefer. A malleable interface for sonic exploration. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, 2010.

Chris Kiefer, Nick Collins, and Geraldine Fitzpatrick. Evaluating the wiimote as a musical controller. In *Proceedings of the International Computer Music Conference*, Belfast, 2008.

Niklas Klügel, Marc René Frieß, Georg Groh, and Florian Echtler. An approach to collaborative music composition. In Alexander Refsum Jensenius, Anders Tveit, Rolf Inge Godøy, and Dan Overholt, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 32–35, Oslo, Norway, 2011. URL http://www.nime2011.org/proceedings/papers/B02-Klugel.pdf.

Stefan Koelsch, Erich Schrger, and Mari Tervaniemi. Superior pre-attentive auditory processing in musicians. 1999. URL http://edoc.mpg.de/239449.

Robin Laney, Chris Dobbyn, Anna Xambó, Mattia Schirosa, Dorothy Miell, Karen Littleton, and Nick Dalton. Issues and techniques for collaborative music making on multi-touch surfaces. In *7th Sound and Music Computing Conference*, Barcelona, 2010.

Kurt Luther and Amy Bruckman. Leadership in online creative collaboration. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, CSCW '08, pages 343–352, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-007-4. doi: http://doi.acm.org/10.1145/1460563.1460619. URL http://doi.acm.org/10.1145/1460563.1460619.

Robin Maconie. *Stockhausen on Music*. Marion Boyars, 1989.

Thor Magnusson. Screen-based musical interfaces as semiotic machines. In *Proceedings of the 2006 conference on New interfaces for musical expression*, NIME '06, pages 162–167, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3. URL http://dl.acm.org/citation.cfm?id=1142215.1142254.

Thor Magnusson. Of epistemic tools: musical instruments as cognitive extensions. *Organised Sound*, 14(02):168–176, 2009. doi: 10.1017/S1355771809000272. URL http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=5882372&fulltextType=RA&fileId=S1355771809000272.

Thor Magnusson and Enrike Hurtado Mendieta. The acoustic, the digital and the body: a survey on musical instruments. In *Proceedings of the 7th international conference on New interfaces for musical expression*, NIME '07, pages 94–99, New York, NY, USA, 2007. ACM. doi: 10.1145/1279740.1279757. URL http://doi.acm.org/10.1145/1279740.1279757.

Nicolas Makelberge. 'Le concert, c'est moi'. In *Proceedings of the 1st DESIRE Network Conference on Creativity and Innovation in Design*, DESIRE '10, pages 150–153, Lancaster, UK, UK, 2010. Desire Network. URL http://portal.acm.org/citation.cfm?id=1854969.1854991.

Gary Marcus. *Guitar Zero: The New Musician and the Science of Learning*. Penguin Press HC, 2012.

Adnan Marquez-Borbon, Michael Gurevich, A. Cavan Fyans, and Paul Stapleton. Designing digital musical interactions in experimental contexts. In Alexander Refsum Jensenius, Anders Tveit, Rolf Inge Godøy, and Dan Overholt, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 373–376, Oslo, Norway, 2011. URL http://www.nime2011.org/proceedings/papers/K02-Marquez-Borbon.pdf.

Rangarirai Matavire and Irwin Brown. Investigating the use of "Grounded Theory" in information systems research. In *SAICSIT '08: Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries*, pages 139–147, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-286-3. doi: http://doi.acm.org/10.1145/1456659.1456676.

James McCartney. Rethinking the computer music language: Super Collider. *Computer Music Journal*, 26(4), 2002.

Tim Merritt, Weiman Kow, Christopher Ng, Kevin McGee, and Lonce Wyse. Who makes what sound?: supporting real-time musical improvisations of electroacoustic ensembles. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, OZCHI '10, pages 112–119, New York, NY, USA,

2010. ACM. ISBN 978-1-4503-0502-0. doi: http://doi.acm.org/10.1145/1952222.1952245. URL http://doi.acm.org/10.1145/1952222.1952245.

Oussama Metatla. *Collaborating Through Sounds: Audio-only Interaction With Diagrams*. PhD thesis, Queen Mary University of London, 2010.

Paul Stapleton Michael Gurevich and Adnan Marquez-Borbon. Style and constraint in electronic musical instruments. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, 2010.

Evandro Miletto, Marcelo Pimenta, Francois Bouchet, Jean-Paul Sansonnet, and Damian Keller. Music creation by novices should be both prototypical and cooperative: Lessons learned from codes. In *12th Brazilian Symposium on Computer Music (SBCM)*, 2009.

Evandro Manara Miletto, Luciano Vargas Flores, Jérôme Rutily, and Marcelo Soares Pimenta. CODES: supporting awareness in a web-based environment for collective music prototyping. In *IHC '06: Proceedings of VII Brazilian symposium on Human factors in computing systems*, pages 50–55, New York, NY, USA, 2006. ACM. ISBN 1-59593-432-4. doi: http://doi.acm.org/10.1145/1298023.1298031.

Roger Mills. Dislocated sound: A survey of improvisation in networked audio platforms. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, Sydney, Australia, 2010.

Meredith Ringel Morris, Dan Morris, and Terry Winograd. Individual audio channels with single display groupware: effects on communication and task strategy. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 242–251, New York, NY, USA, 2004. ACM. ISBN 1-58113-810-5. doi: http://doi.acm.org/10.1145/1031607.1031646.

Ann J. Morrison, Peta Mitchell, and Margot Brereton. The lens of ludic engagement: evaluating participation in interactive art installations. In *Proceedings of the 15th international conference on Multimedia*, MULTIMEDIA '07, pages 509–512, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-702-5. doi: http://doi.acm.org/10.1145/1291233.1291358. URL http://doi.acm.org/10.1145/1291233.1291358.

Ann J. Morrison, Peta Mitchell, and Stephen Viller. Evoking gesture in interactive art. In *HCC '08: Proceeding of the 3rd ACM international workshop on Human-centered computing*, pages 11–18, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-320-4. doi: http://doi.acm.org/10.1145/1462027.1462030.

Jos Mulder. The loudspeaker as musical instrument. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010), Sydney, Australia*, 2010.

Michael J Muller and Sandra Kogan. Grounded Theory Method in HCI and CSCW. Technical Report 10-09, IBM Watson Research Center, 2010.

Alexander Müller-Rakow and Jochen Fuchs. The human skin as an interface for musical expression. In *NIME '12: Proceedings of the 2012 conference on New interfaces for musical expression*, 2012.

Tim Murray-Browne, Di Mainstone, Nick Bryan-Kinns, and Mark D. Plumbley. The medium is the message: Composing instruments and performing mappings. In Alexander Refsum Jensenius, Anders Tveit, Rolf Inge Godøy, and Dan Overholt, editors, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 56–59, Oslo, Norway, 2011. URL http://www.nime2011.org/proceedings/papers/B08-Murray-Browne.pdf.

Shahin Nabavian. *Distributed cognition in Joint Music Composition: Exploring the role of language and artefacts in multi-session creative collaborative work*. PhD thesis, Queen Mary, University of London, 2009.

Shahin Nabavian and Nick Bryan-Kinns. Analysing group creativity: A distributed cognitive study of joint music composition. In *Proc. of Cognitive Science*, pages 1856–1861, 2006.

Thomas M. Nelson and Thomy H. Nilsson. Comparing headphone and speaker effects on simulated driving. *Accident Analysis & Prevention*, 22(6):523 – 529, 1990.

Henry Newton-Dunn, Hiroaki Nakano, and James Gibson. Block jam: a tangible interface for interactive music. In *NIME '03: Proceedings of the 2003 conference on New interfaces for musical expression*, pages 170–177, Singapore, Singapore, 2003. National University of Singapore.

Donald A. Norman. *The Design of Everyday Things*. Basic Books, 2002.

Jieun Oh, Jorge Herrera, Nicholas J. Bryan, Luke Dahl, and Ge Wang. Evolving the mobile phone orchestra. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, Sydney, Australia, 2010.

Judith S. Olson, Gary M. Olson, Marianne Storrosten, and Mark Carter. Groupwork close up: a comparison of the group design process with and without a simple group editor. *ACM Trans. Inf. Syst.*, 11(4):321–348, 1993. ISSN 1046-8188.

Judith S. Olson, Gary M. Olson, and David K. Meader. What mix of video and audio is useful for small groups doing remote real-time design work? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '95, pages 362–368, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co. ISBN 0-201-84705-1. doi: http://dx.doi.org/10.1145/223904.223951. URL http://dx.doi.org/10.1145/223904.223951.

Dan Overholt. The musical interface technology design space. *Organised Sound*, 14(2), 2009.

James Patten, Ben Recht, and Hiroshi Ishii. Audiopad: a tag-based interface for musical performance. In *NIME '02: Proceedings of the 2002 conference on New interfaces for musical expression*, pages 1–6, Singapore, Singapore, 2002. National University of Singapore. ISBN 1-87465365-8.

Michaela Pfadenhauer. The lord of the loops. observations at the club culture dj-desk. *Forum: Qualitative Social Research*, 10(3), 2009.

Martin Pichlmair and Fares Kayali. Levels of sound: On the principles of interactivity in music video games. In *Situated Play, Proceedings of DiGRA 2007 Conference*, pages 424 – 430, 2007.

Marcelo Pimenta, Evandro Miletto, Luciano Flores, and Aurelio Hoppe. Cooperative mechanisms for networked music. *Future Generation Computer Systems*, 27(1):100 – 108, 2011. ISSN 0167-739X. doi: 10.1016/j.future.2010.03.005. URL http://www.sciencedirect.com/science/article/pii/S0167739X10000506.

Alain Pinsonneault, Henri Barki, R. Brent Gallupe, and Norberto Hoppen. Electronic brainstorming: The illusion of productivity. *Information Systems Research*, 10(2), 1999.

Ilona R. Posner and Ronald M. Baecker. How people write together. In *Hawaii International Conference on System Sciences*, 1992.

Bhavna Prajapati, Mark Dunne, and Richard Armstrong. Sample size estimation and statistical power analyses. *Optometry Today*, 16(07), 2012.

Naomi L. Quenk. *Essentials of Myers-Briggs Type Indicator Assessment*. John Wiley & Sons, 2009.

Pedro Rebelo and Alain B. Renaud. The Frequencyliator: distributing structures for networked laptop improvisation. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, pages 53–56, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3.

Alain B. Renaud, Alexander Carôt, and Pedro Rebelo. Networked music performance: State of the art. In *30th International Conference: Intelligent Audio Environments (March 2007)*, 2007.

Peter J Rentfrow and Samuel D. Gosling. The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84 (6), 2003.

T. Ritchey. Wicked problems: Structuring social messes with morphological analysis, 2005.

Curtis Roads. *Microsound*. MIT Press, 2004.

Andrew Robertson, Mark Plumbley, and Nick Bryan-Kinns. A turing test for a real-time interactive beat-tracker. In *Proceedings of the 8th International Conference on New Interfaces for Musical Expression (NIME 2008)*, pages 319–324, Genova, Italy, 2008.

Dominic Robson. Play!: sound toys for the non musical. In *NIME '01: Proceedings of the 2001 conference on New interfaces for musical expression*, pages 1–3, Singapore, Singapore, 2001. National University of Singapore.

Yvonne Rogers, Youn-Kyung Lim, William R Hazlewood, and Paul Marshall. Equal opportunities: Do shareable interfaces promote more group participation than single users displays? *Human-Computer Interaction*, 24(2-1):79–116, 2009.

Mark Roseman and Saul Greenberg. Building real-time groupware with groupkit, a groupware toolkit. *ACM Trans. Comput.-Hum. Interact.*, 3(1):66–106, March 1996. ISSN 1073-0516. doi: 10.1145/226159.226162. URL http://doi.acm.org/10.1145/226159.226162.

Mark Rouncefield, John A. Hughes, Tom Rodden, and Stephen Viller. Working with "constant interruption": Cscw and the small office. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, CSCW '94, pages 275–286, New York, NY, USA, 1994. ACM. ISBN 0-89791-689-1. doi: http://doi.acm.org/10.1145/192844.193028. URL http://doi.acm.org/10.1145/192844.193028.

Kathy Ryall, Clifton Forlines, Chia Shen, Meredith Ringel Morris, and Katherine Everitt. Experiences with and observations of direct-touch tabletops. In *TABLETOP '06: Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, pages 89–96, Washington, DC, USA, 2006. IEEE Computer Society. ISBN 0-7695-2494-X. doi: http://dx.doi.org/10.1109/TABLETOP.2006.12.

Robert Keith Sawyer. *Group Creativity: Music, Theater, Collaboration*. Psychology Press, 2003.

Stacey D. Scott, M. Sheelagh, T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, CSCW '04, pages 294–303, New York, NY, USA, 2004. ACM. ISBN 1-58113-810-5. doi: http://doi.acm.org/10.1145/1031607.1031655. URL http://doi.acm.org/10.1145/1031607.1031655.

Chia Shen, Katherine Everitt, and Kathleen Ryall. Ubitable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *In Proc. UbiComp 2003*, pages 281–288, 2003.

Jennifer G. Sheridan, Nick Bryan-Kinns, and Alice Bayliss. Encouraging witting participation and performance in digital live art. In *BCS-HCI '07: Proceedings of the 21st British CHI Group Annual Conference on HCI 2007*, pages 13–23, Swinton, UK, UK, 2007. British Computer Society. ISBN 978-1-902505-94-7.

Christopher Small. *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press, 1998.

Dennis Smalley. Spectromorphology: explaining sound-shapes. *Organised Sound*, 2(2):107–126, 1997.

Ian Smith and Scott E. Hudson. Low disturbance audio for awareness and privacy in media space applications. In *Proceedings of the third ACM international conference on Multimedia*,

MULTIMEDIA '95, pages 91–97, New York, NY, USA, 1995. ACM. ISBN 0-89791-751-0. doi: 10.1145/217279.215253. URL http://doi.acm.org/10.1145/217279.215253.

M. Stefik, D. G. Bobrow, G. Foster, S. Lanning, and D. Tatar. WYSIWIS revised: early experiences with multiuser interfaces. *ACM Trans. Inf. Syst.*, 5(2):147–167, April 1987. ISSN 1046-8188. doi: 10.1145/27636.28056. URL http://doi.acm.org/10.1145/27636.28056.

Jason Stewart, Benjamin B. Bederson, and Allison Druin. Single display groupware: a model for co-present collaboration. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 286–293, New York, NY, USA, 1999. ACM. ISBN 0-201-48559-1. doi: http://doi.acm.org/10.1145/302979.303064.

D. Stowell, A. Robertson, N. Bryan-Kinns, and M. D. Plumbley. Evaluation of live human-computer music-making: Quantitative and qualitative approaches. *Int. J. Hum.-Comput. Stud.*, 67(11):960–975, November 2009. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2009.05.007. URL http://dx.doi.org/10.1016/j.ijhcs.2009.05.007.

Dan Stowell, Mark Plumbley, and Nick Bryan-Kinns. Discourse analysis evaluation method for expressive musical interfaces. In *New Interfaces for Musical Expression (NIME'08)*, Genova, Italy, 2008.

Marco Stroppa. Live electronics or live music? towards a critique of interaction. *Contemporary Music Review*, 18(3):41–77, 1999. doi: doi:10.1080/07494469900640341. URL http://www.ingentaconnect.com/content/routledg/gcmr/1999/00000018/00000003/art00008.

Ben Swift, Harry Gardener, and Alistair Riddell. A chasing after the wind: Experience in computer-supported group musicmaking. In *BCS HCI 2011 Workshop - When Words Fail: What can Music Interaction tell us about HCI?*, 2011.

Atau Tanaka. Malleable mobile music. In *Adjunct Proceedings of the 6th International Conference on Ubiquitous Computing (Ubicomp)*, 2004.

Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. Collaborative coupling over tabletop displays. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1181–1190, New York, NY, USA, 2006. ACM. ISBN 1-59593-372-7. doi: http://doi.acm.org/10.1145/1124772.1124950.

Michael Terry, Elizabeth D. Mynatt, Kumiyo Nakakoji, and Yasuhiro Yamamoto. Variation in element and action: supporting simultaneous development of alternative solutions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '04, pages 711–718, New York, NY, USA, 2004. ACM. ISBN 1-58113-702-8. doi: http://doi.acm.org/10.1145/985692.985782. URL http://doi.acm.org/10.1145/985692.985782.

Jean-Baptiste Thiebaut. *Sketching Music: Representation and Composition*. PhD thesis, Queen Mary, University of London, 2010.

Jennifer Thom-Santelli, Dan R. Cosley, and Geri Gay. What's mine is mine: territoriality in collaborative authoring. In *Proceedings of the 27th international conference on Human factors in computing systems*, CHI '09, pages 1481–1484, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: http://doi.acm.org/10.1145/1518701.1518925. URL http://doi.acm.org/10.1145/1518701.1518925.

Daniel Trueman, Perry Cook, Scott Smallwood, and Ge Wang. PLOrk: The Princeton Laptop Orchestra, Year 1. In *Proceedings of the International Computer Music Conference*, 2006.

Sylvia Truman. Soundscape: A music composition environment designed to facilitate collaborative creativity in the classroom. In *SMC Conference 2011*, 2011a.

Sylvia M. Truman. An exploration of creativity in school children's music composition: study, software and framework. In *Proceedings of the 8th ACM conference on Creativity and cognition*, C&C '11, pages 395–396, New York, NY, USA, 2011b. ACM. ISBN 978-1-4503-0820-5. doi: 10.1145/2069618.2069713. URL http://doi.acm.org/10.1145/2069618.2069713.

Edward Tse, Jonathan Histon, Stacey D. Scott, and Saul Greenberg. Avoiding interference: how people use spatial separation and partitioning in sdg workspaces. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 252–261, New York, NY, USA, 2004. ACM. ISBN 1-58113-810-5. doi: http://doi.acm.org/10.1145/1031607.1031647.

M. Unehara and T. Onisawa. Music composition system based on subjective evaluation. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pages 980–986 vol.1, Oct. 2003. doi: 10.1109/ICSMC.2003.1243942.

Daniel Vogel and Ravin Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, UIST '04, pages 137–146, New York, NY, USA, 2004. ACM. ISBN 1-58113-957-8. doi: http://doi.acm.org/10.1145/1029632.1029656. URL http://doi.acm.org/10.1145/1029632.1029656.

Marcelo Mortensen Wanderley and Nicola Orio. Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, 26(3):62–76, 2002.

Ge Wang, Ananya Misra, Philip Davidson, and Perry R. Cook. Co-Audicle: A collaborative audio programming space. In *Proceedings of the International Computer Music Conference*, Barcelona, 2005.

Ge Wang, Ananya Misra, and Perry R. Cook. Building collaborative graphical interfaces in the Audicle. In *Proceedings of the 2006 conference on New interfaces for musical expression*, NIME '06, pages 49–52, Paris, France, France, 2006. IRCAM — Centre Pompidou. ISBN 2-84426-314-3. URL http://dl.acm.org/citation.cfm?id=1142215.1142225.

Ge Wang, Georg Essl, and Henri Penttinen. Mopho: Do mobile phones dream of electric orchestras? In *Proceedings of the International Computer Music Conference (ICMC-08),*, Belfast, Northern Ireland., 2008.

Ge Wang, Nicholas Bryan, Jieun Oh, and Rob Hamilton. Stanford laptop orchestra (Slork). In *Proceedings of the International Computer Music Conference (ICMC 2009)*, 2009.

Christine E. Wania, Michael E. Atwood, and Katherine W. McCain. How do design and evaluation interrelate in hci research? In *Proceedings of the 6th conference on Designing Interactive systems*, DIS '06, pages 90–98, New York, NY, USA, 2006. ACM. ISBN 1-59593-367-0. doi: 10.1145/1142405.1142421. URL http://doi.acm.org/10.1145/1142405.1142421.

Gil Weinberg. *Interconnected Musical Networks - Bringing Expression and Thoughtfulness to Collaborative Group Playing*. PhD thesis, Massachusetts Institute of Technology, 2003.

Gil Weinberg. Interconnected musical networks: Towards a theoretical framework. *Computer Music Journal*, 29(2):23–39, 2005.

Gil Weinberg, Roberto Aimi, and Kevin Jennings. The Beatbug network: a rhythmic system for interdependent group collaboration. In *Proceedings of the 2002 conference on New interfaces for musical expression*, 2002.

Katie Wilkie, Simon Holland, and Paul Mulholland. What can the language of musicians tell us about music interaction design? *Computer Music Journal*, 34(4):34–48, 2010.

Mike Wozniewski, Nicolas Bouillo, Zack Settel, and Jeremy R. Cooperstock. Large-scale mobile audio environments for collaborative musical interaction. In *NIME '08: Proceedings of the 2008 conference on New interfaces for musical expression*, 2008.

Matthew Wright, Adrian Freed, and Ali Momeni. Open Sound Control: State of the art 2003. In *International Conference on New Interfaces for Musical Expression*, pages 153–159, Montreal, 2003. OpenSound Control.

Lonce Wyse and Norikazu Mitani. Bridges for networked musical ensembles. In *The 2009 International Computer Music Conference*, Quebec, 2009.

Anna Xambó, Robin Laney, and Chris Dobbyn. TOUCHtr4ck: democratic collaborative music. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, TEI '11, pages 309–312, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0478-8. doi: http://doi.acm.org/10.1145/1935701.1935773. URL http://doi.acm.org/10.1145/1935701.1935773.

Anna Xambó, Robin Laney, Chris Dobbyn, and Sergi Jordà. Collaborative music interaction on tabletops: an HCI approach. In *25th BCS Conference on Human-Computer Interaction*, 2011a.

Anna Xambó, Robin Laney, Chris Dobbyn, and Sergi Jordà. Multi-touch interaction principles for collaborative real-time music activities: towards a pattern language. In *International Computer Music Conference Proceedings*, Huddersfield, 2011b.

Vivienne M. Young and Andrew M. Colman. Some psychological processes in string quartets. *Psychology of Music*, 7(1):3–11, April 1979.

# Appendix A

# Friedman Test

---

The Friedman Test was used to test for statistical significance within the interaction log data collected for Studies One and Two. Both these studies feature an identical three condition within subjects design where each participant was subjected to three different experimental conditions. The interaction log data contained the quantities of specific actions performed by each participant. The results of these studies are presented in Section 5.10 and Section 6.11.

The Friedman test is an appropriate test for the analysis of data collected in this way, as it is a non-directional (and therefore suitable for two-tailed hypothesis testing), non-parametric statistical test used to identify differences across three test treatments within a repeated measure experimental design (Greene and D'Oliveira, 1999). The Friedman test operates on ordinal (ranked) data, meaning it considers the tendency for a series of scores to follow a particular ranked order, but does not take into account the relative quantities, or numeric differences between data points from different conditions. With reference to the studies in this thesis, using an ordinal test reveals the tendency for a particular condition to encourage more or less of an activity (such as module creations), but does not provide any indication of the relative differences between the scores.

The Friedman test was also used in Study Two to analyse the post-test questionnaire results (Section 6.11.1. The post-test questionnaire in Study Two asked participants to rank the conditions against a range of statements (e.g. 'the best music', 'The interface was most complex'). The Friedman test was appropriate in this case as there were three conditions, and the ranked nature of the data made it suitable for analysis using an ordinal test.

In Studies One and Two, a post-hoc test was applied to determine the statistical effect size of all significant findings identified with the Friedman Test. The post-hoc test performed pairwise comparisons using the Wilcoxon signed-rank test to identify which pairs of conditions out of the three caused a significant effect. A Bonferroni correction was applied at this stage to account for the increased chance of detecting a significant result when multiple pairs are compared. The final part of the post hoc analysis was to calculate the statistical effect size for each pair, taking into account the sample size using the formula:

$$r = \frac{Z}{\sqrt{N}}$$

Where r is the effect size, N is the number of matched pairs in the analysis (Corder and Foreman, 2009). The effect size shows the degree of association between the groups (Corder and Foreman, 2009), and range from 0.0 to 1.0, with 0.10, 0.30 and 0.50 being widely accepted as small, medium and large sizes respectively (Corder and Foreman, 2009).

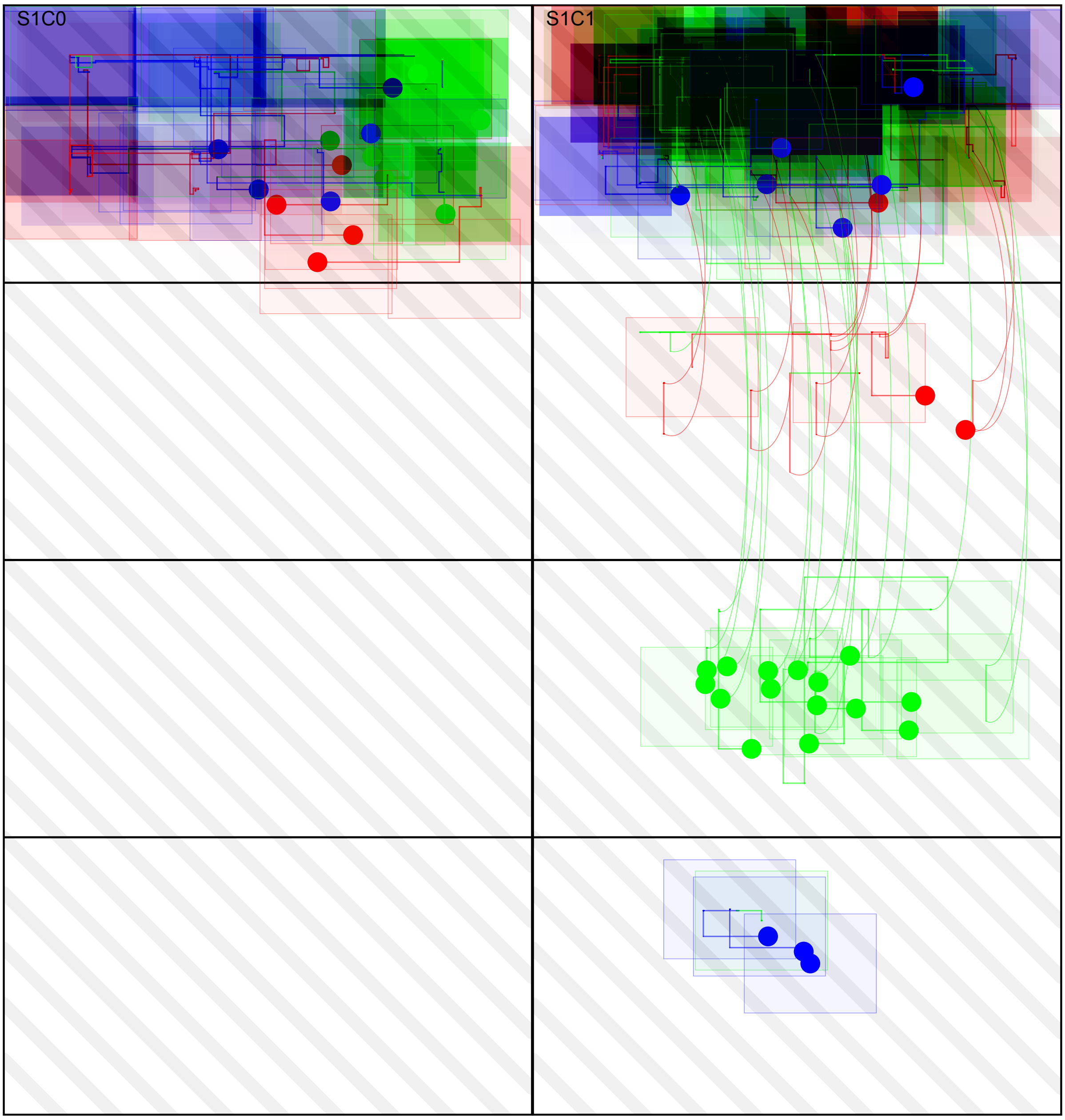
The findings of Studies One and Two, have been been presented in three peer-reviewed publications Fencott and Bryan-Kinns (2010, 2012, 2013), all of which included results produced using the Friedman Test.

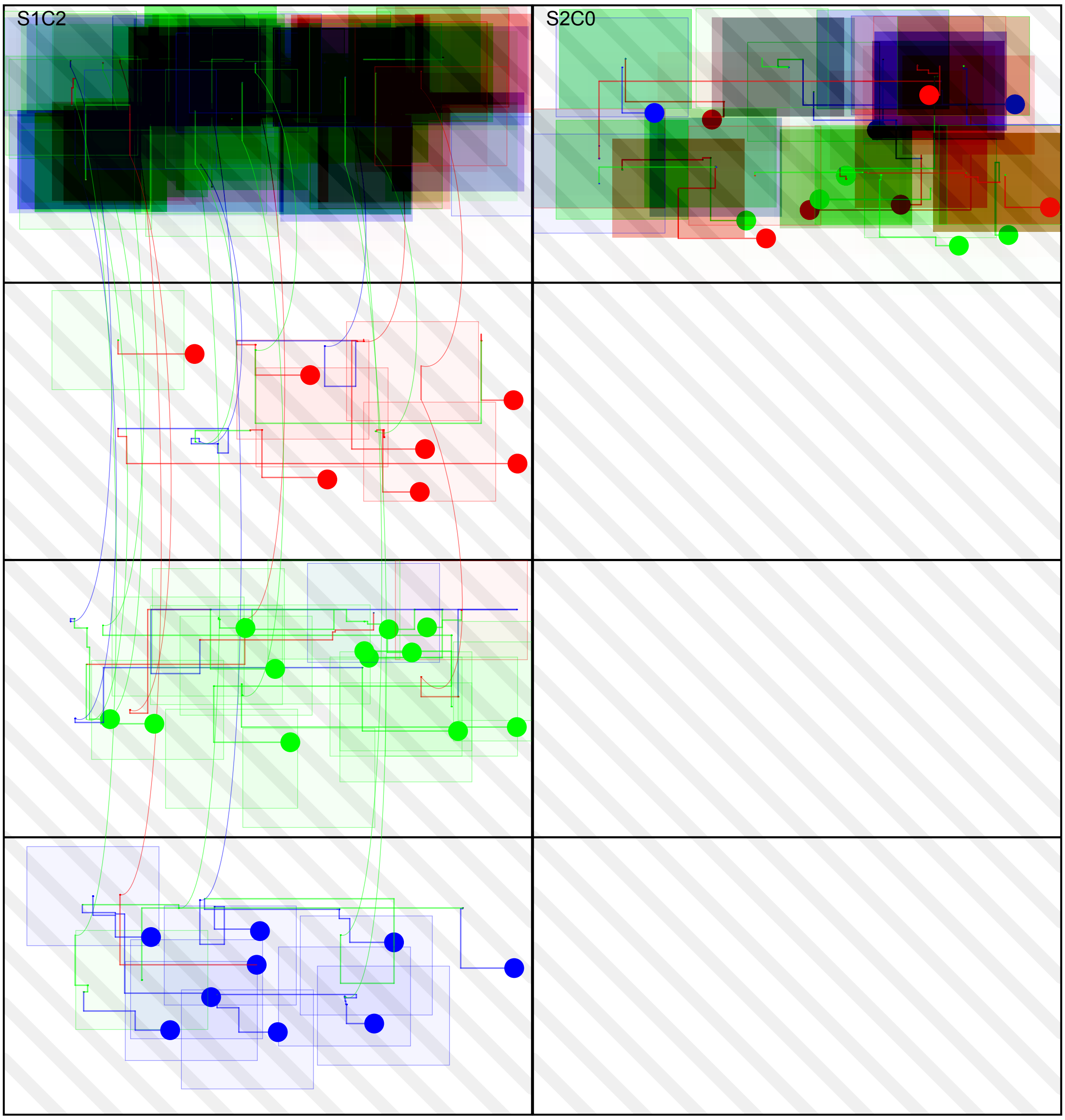
# Appendix B

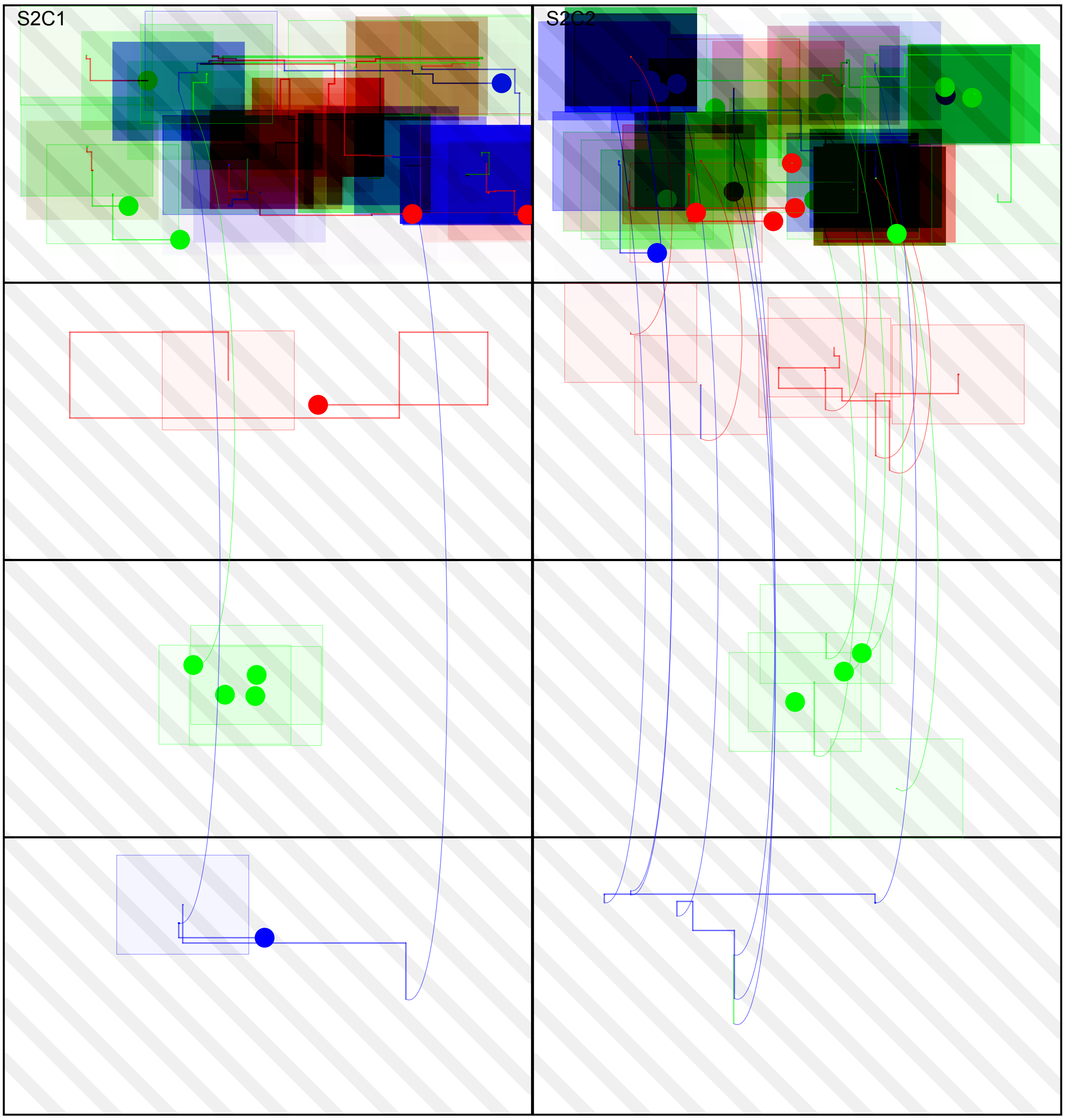
# Study One Materials

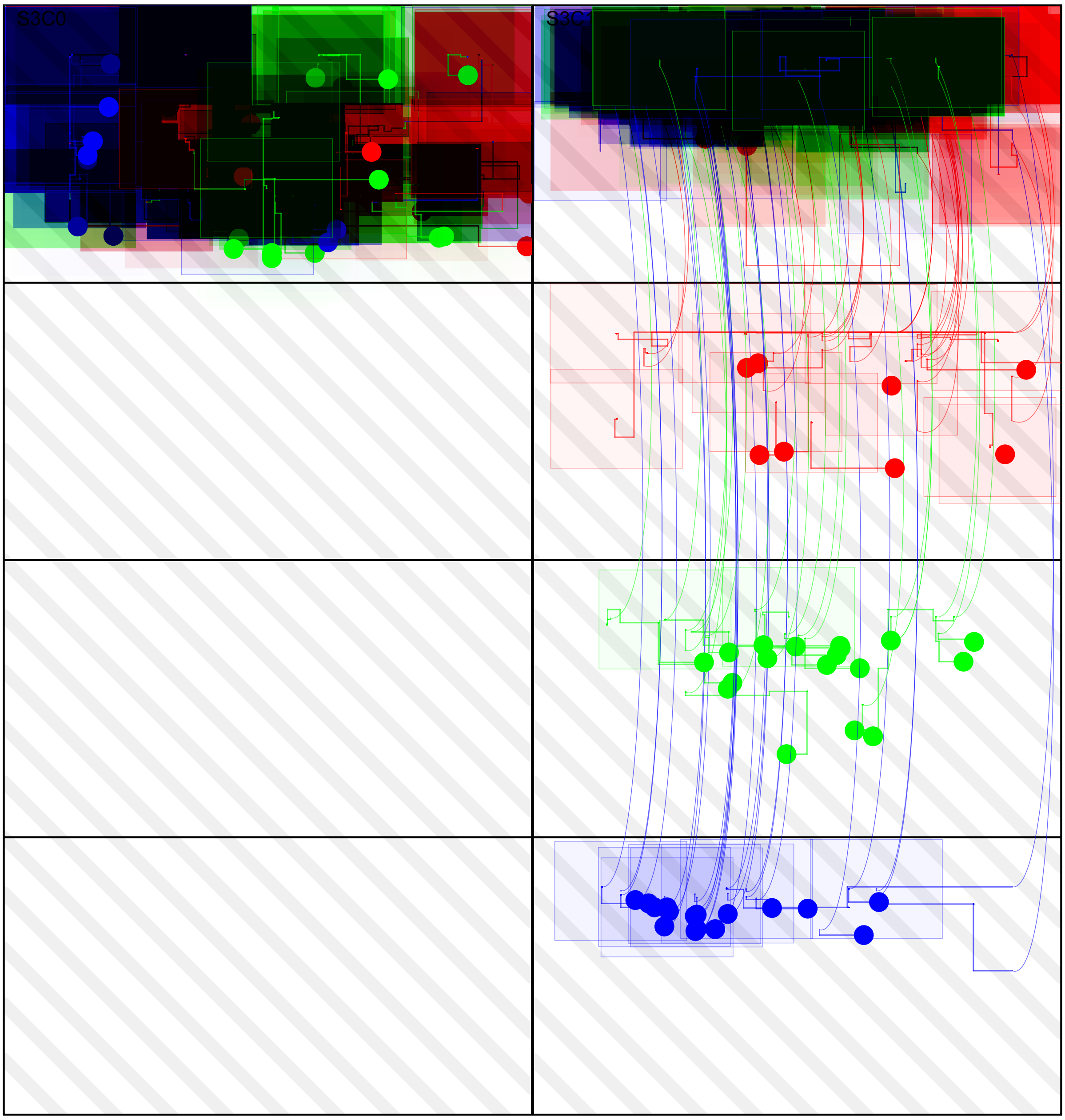
---

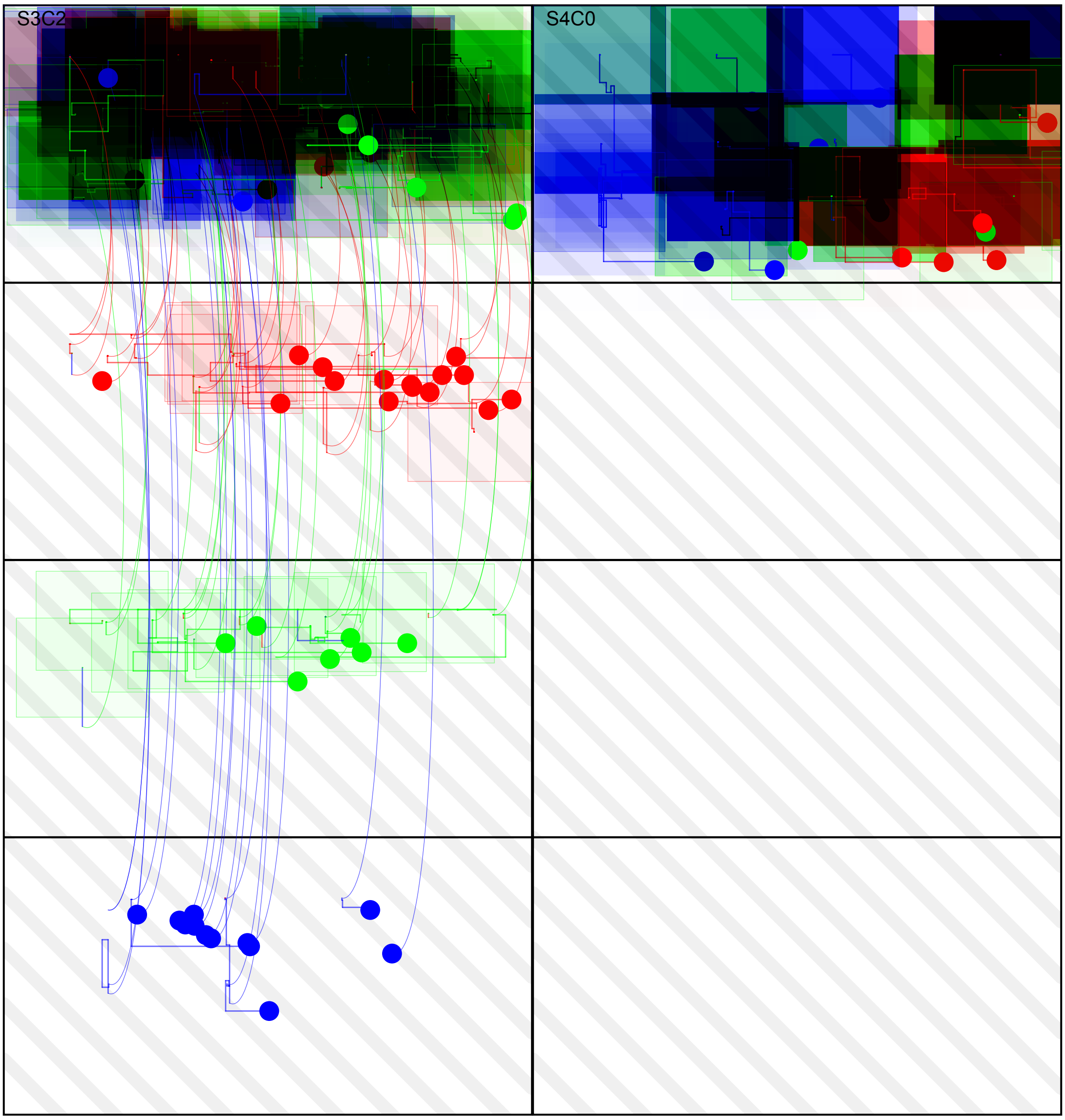
### Study Two Shared Workspace Spatial Visualisations

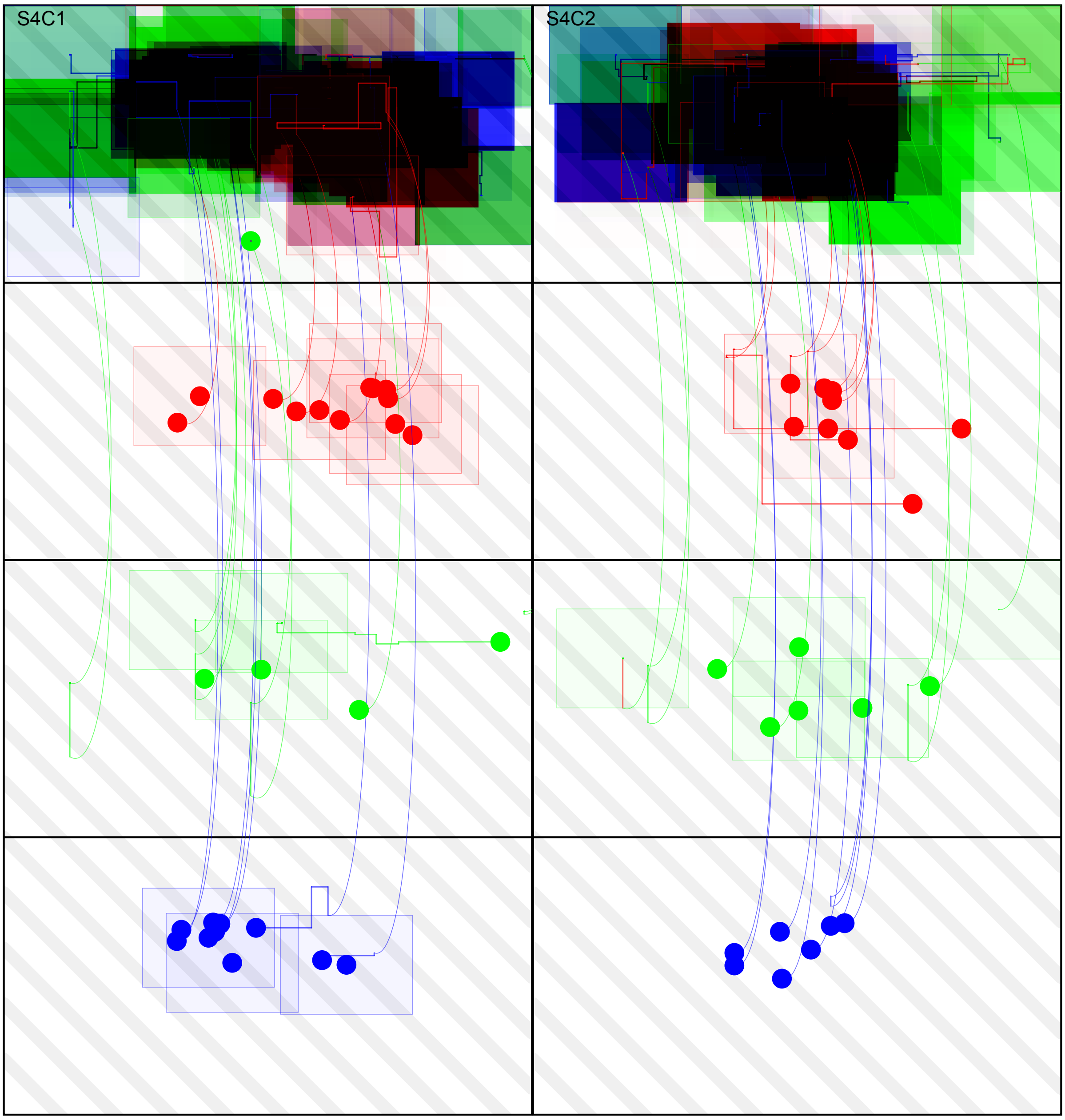

S1C0
S1C1

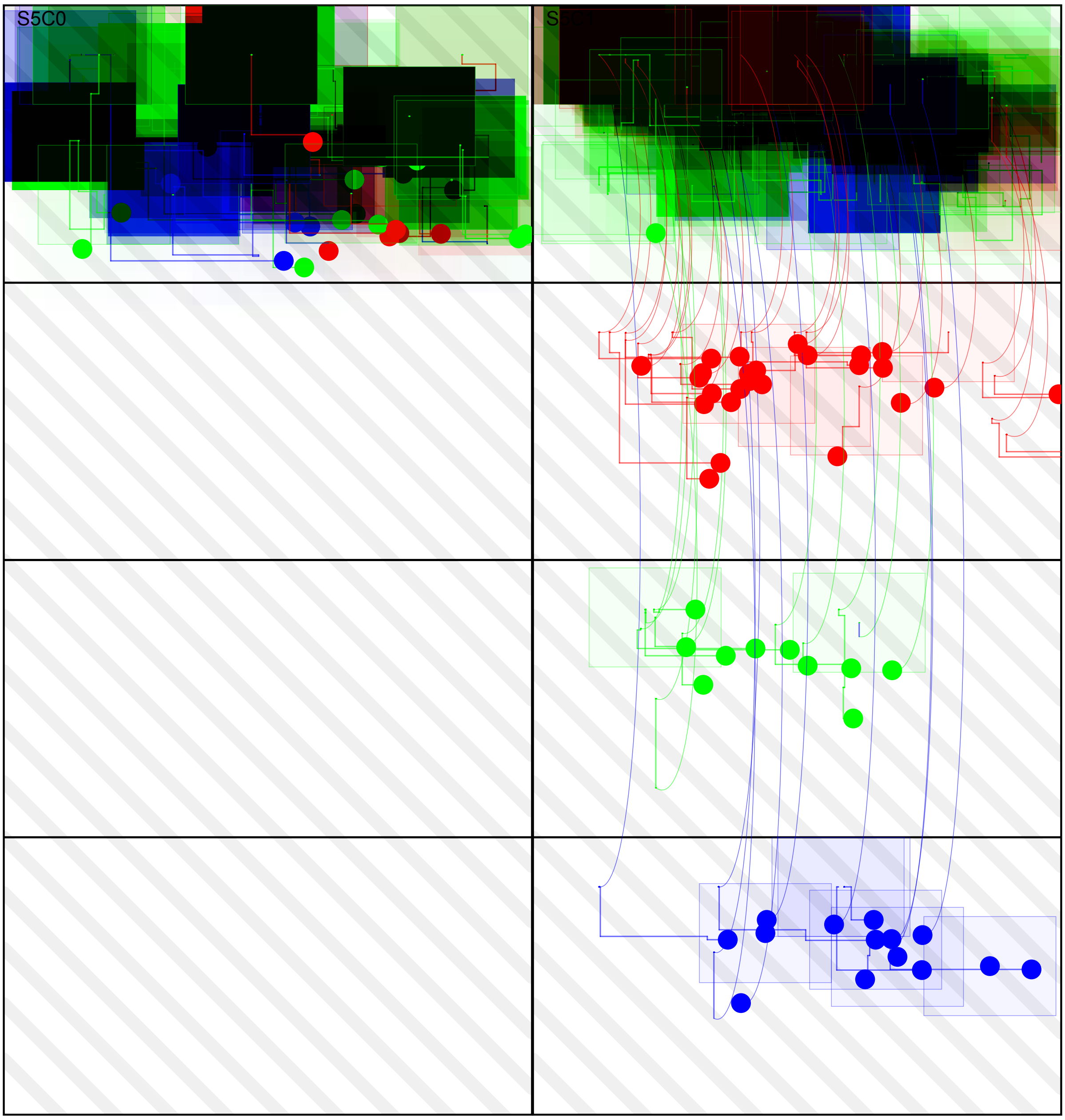

S1C2
S2C0

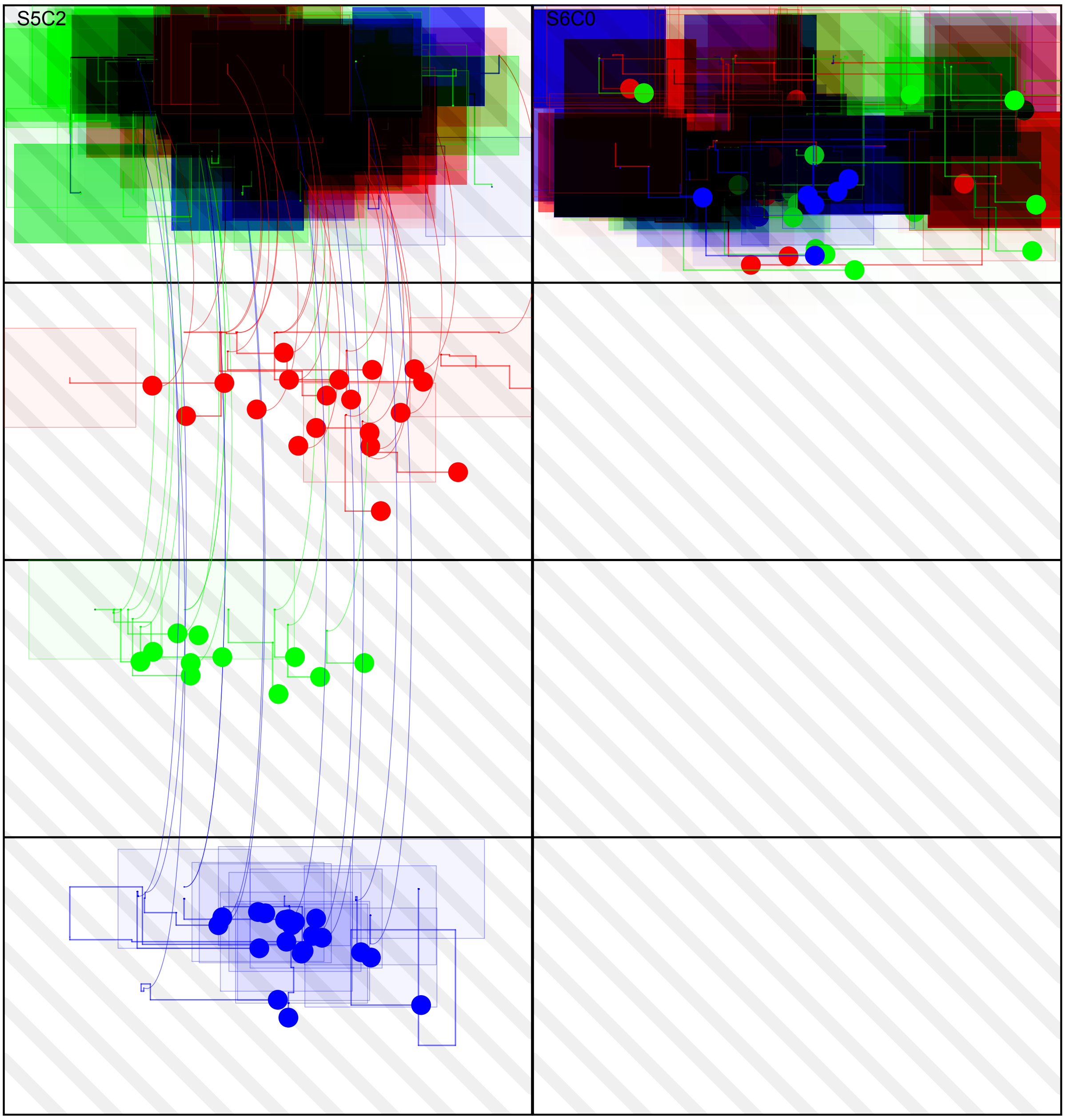

S2C1
S2C2

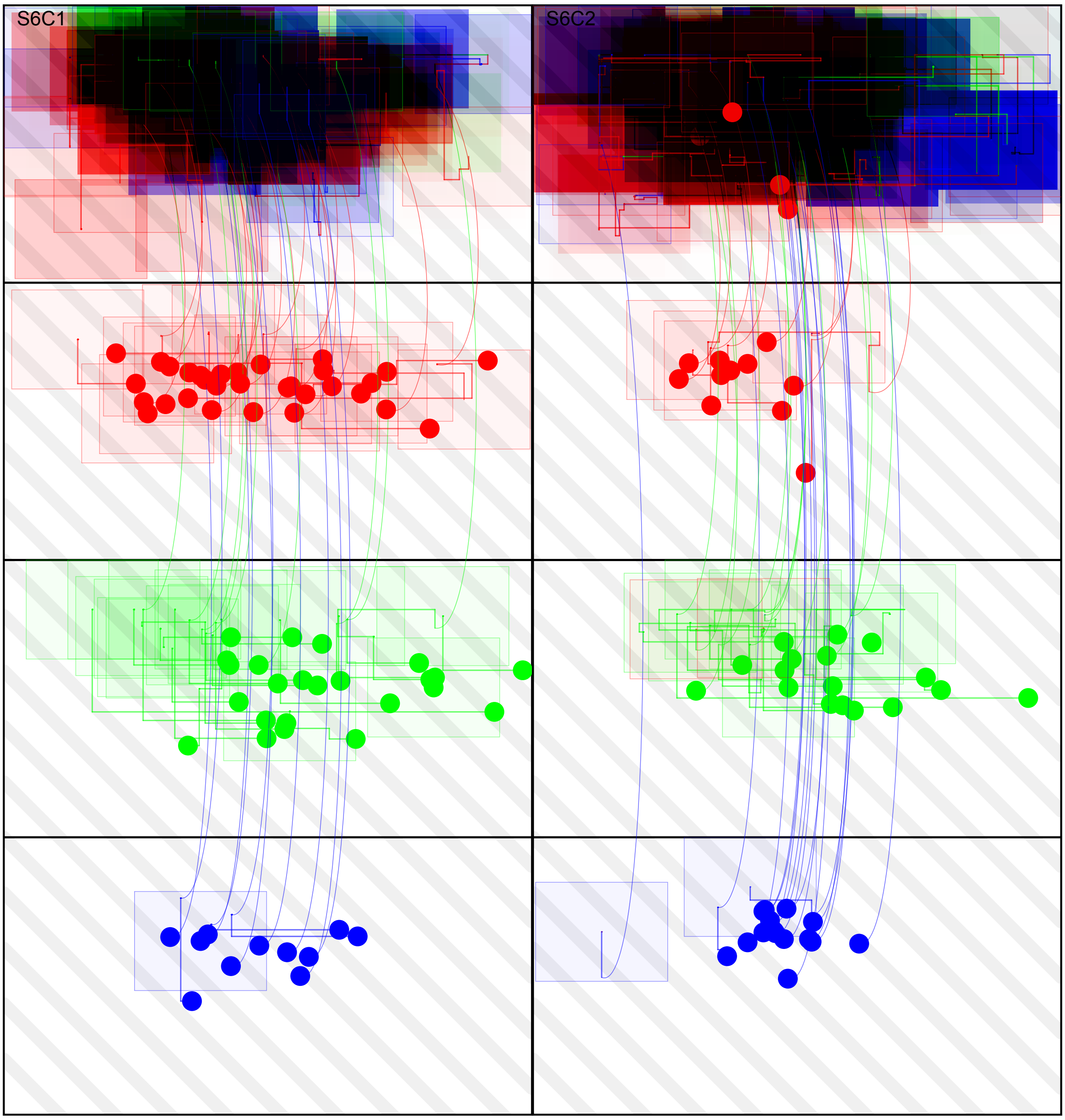

S3C0
S3C1

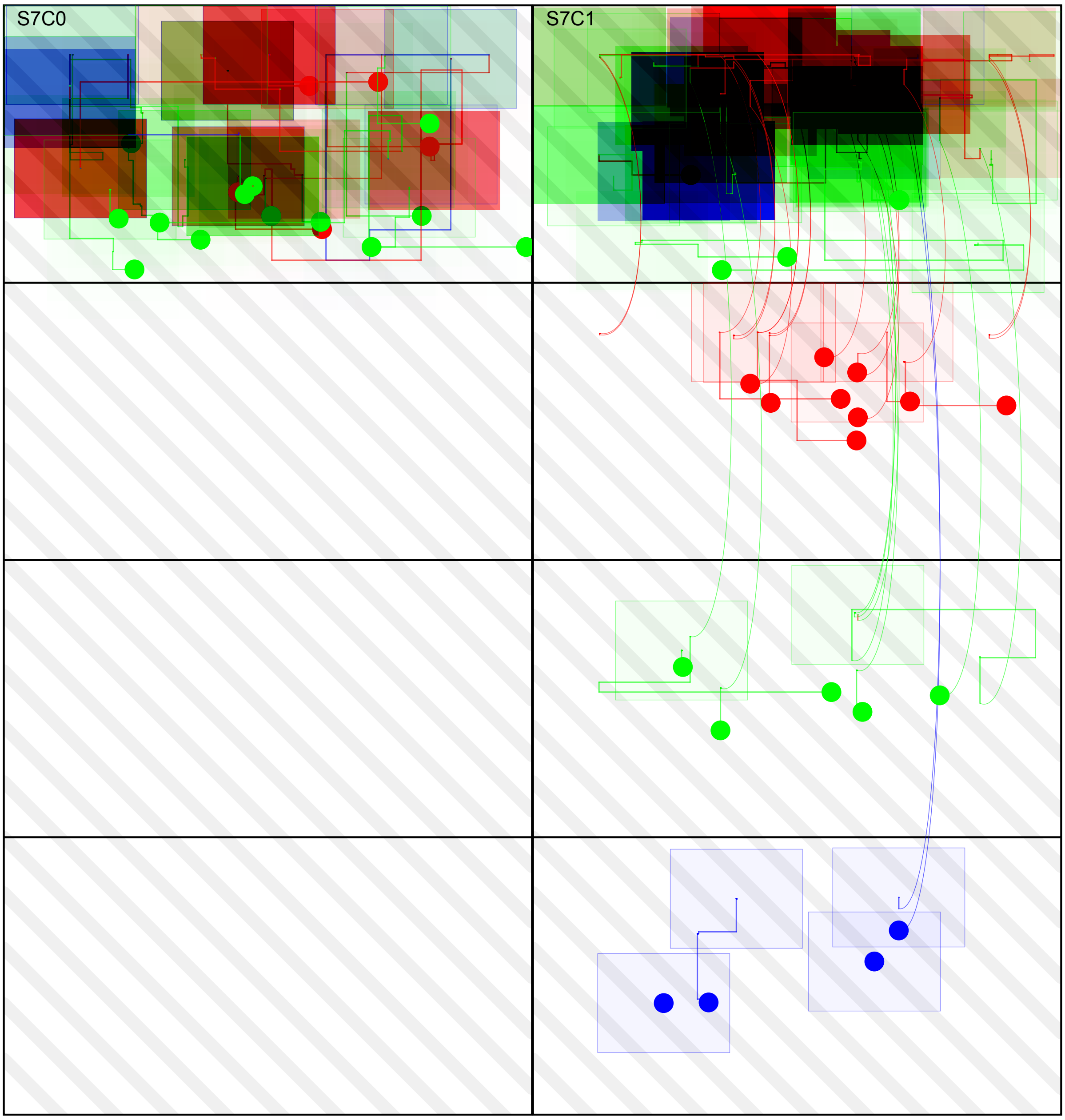

S3C2
S4C0

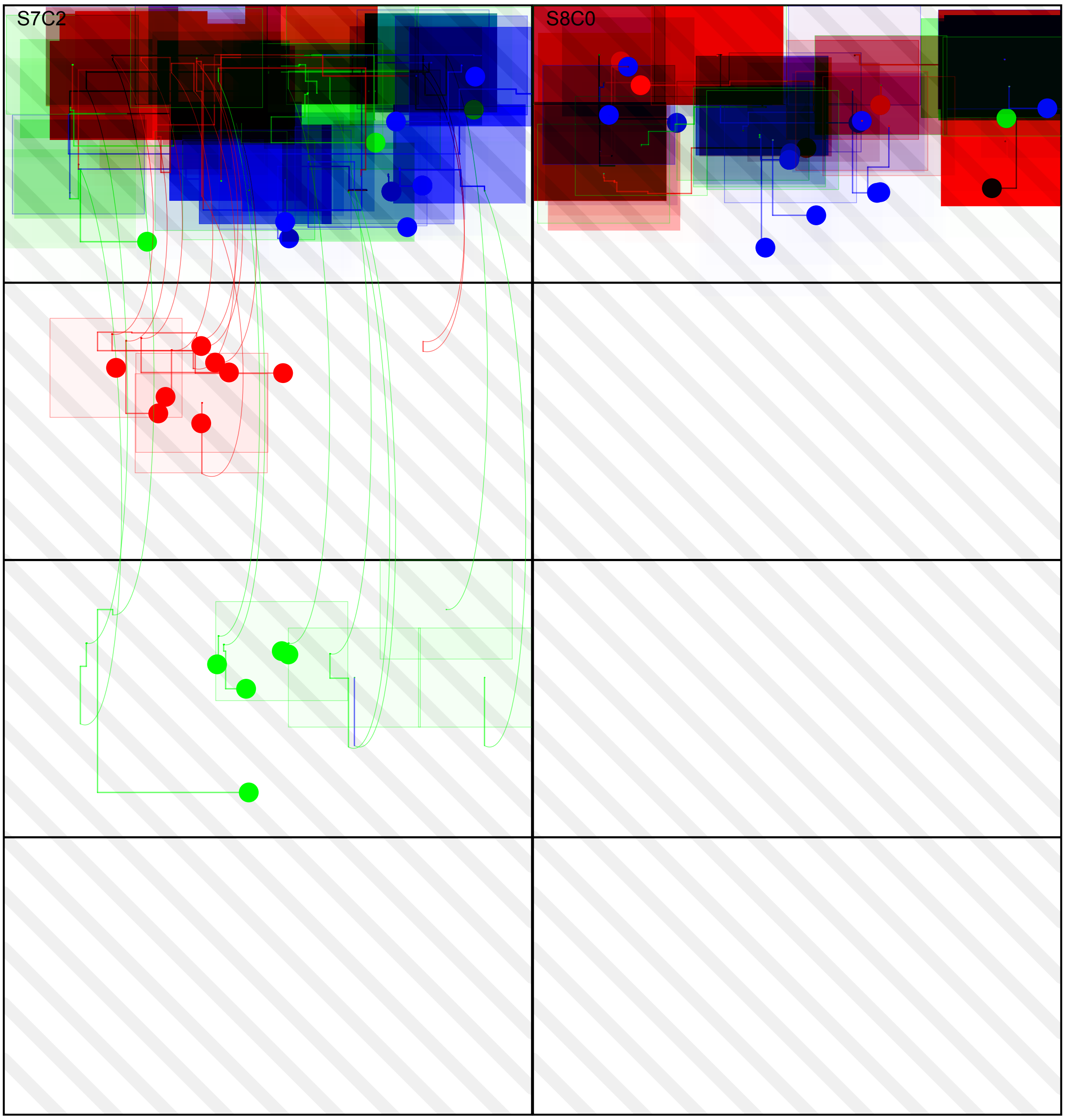

S4C1
S4C2

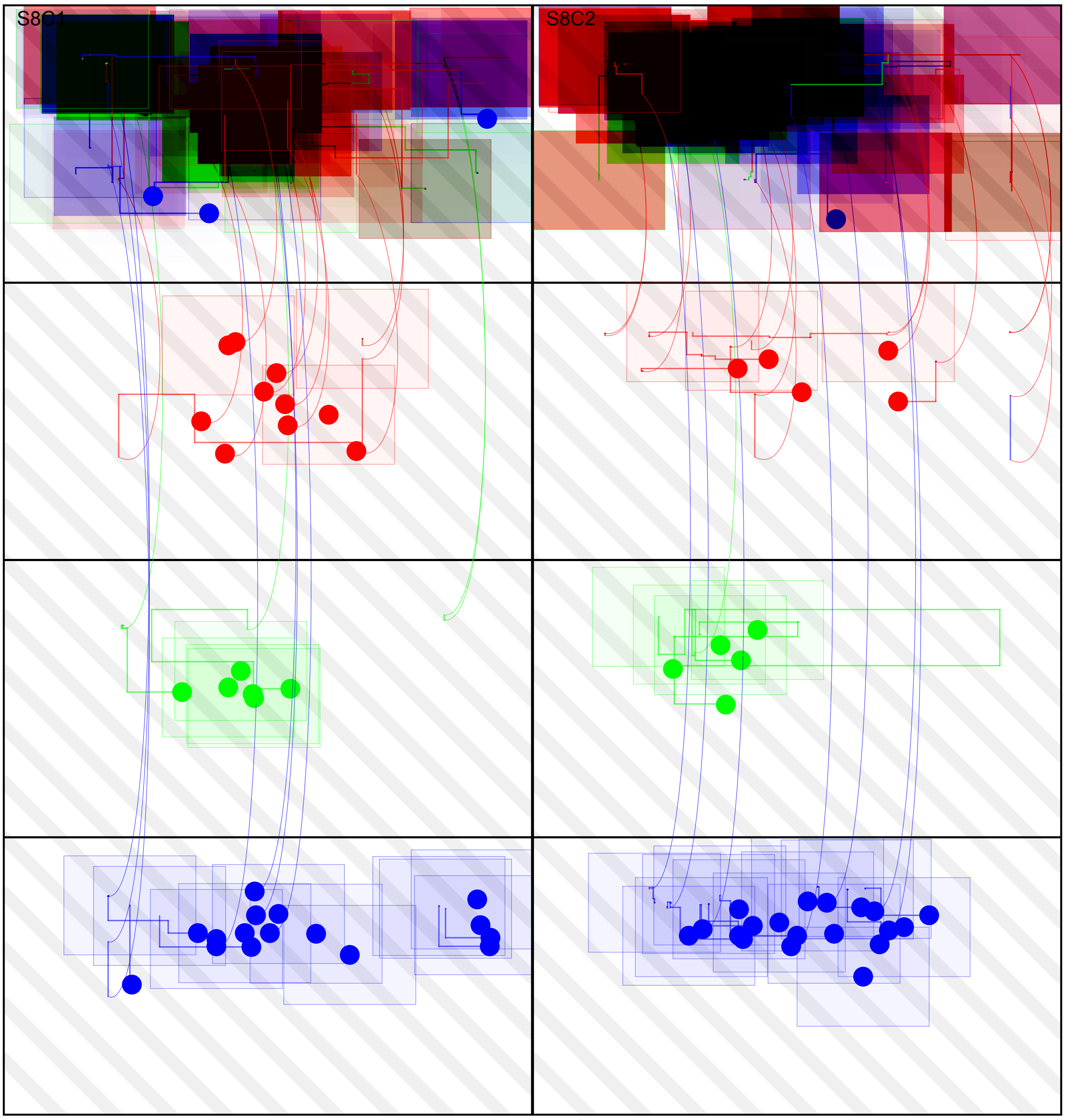

S5C0
S

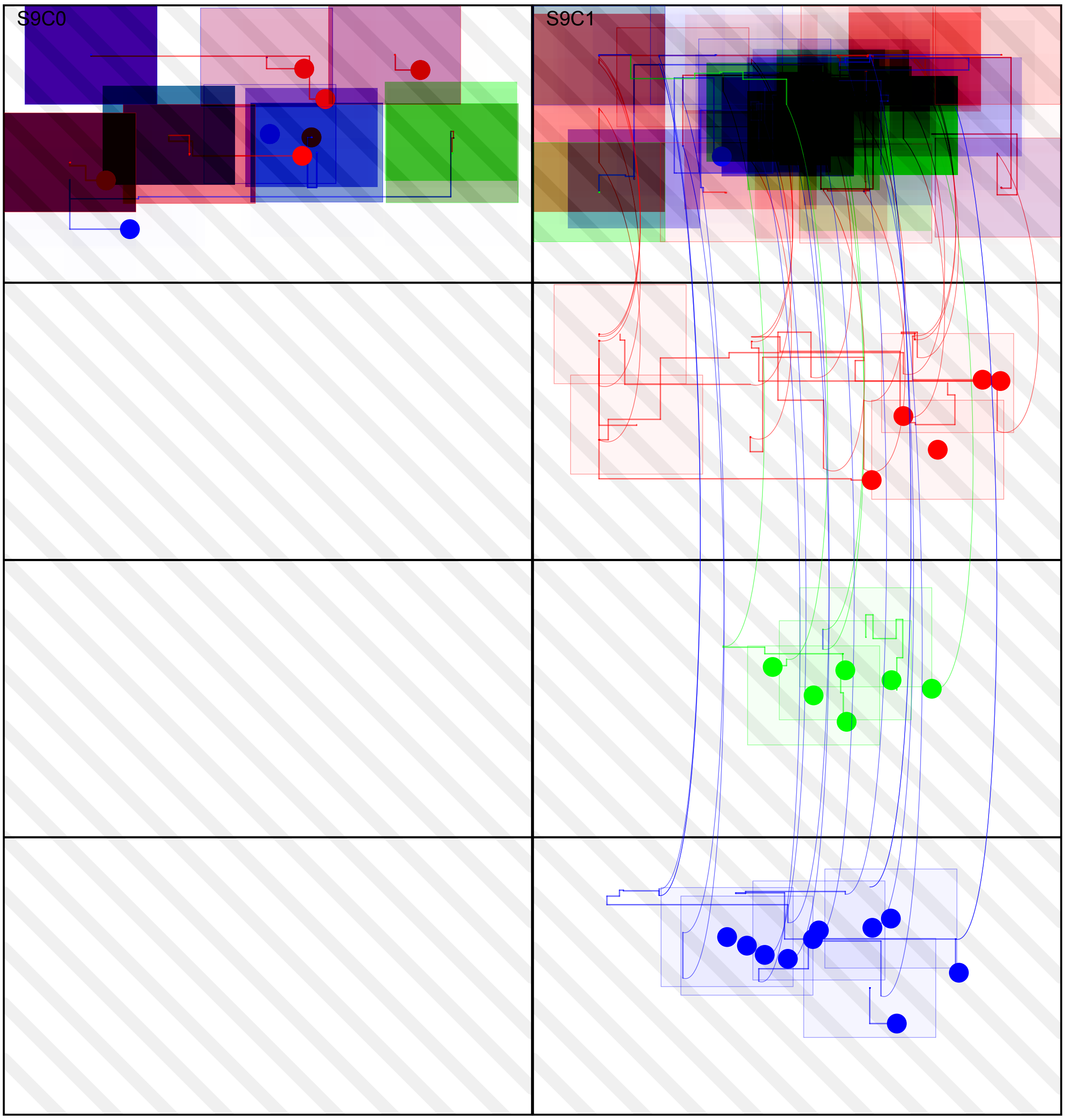

S5C2
S6C0

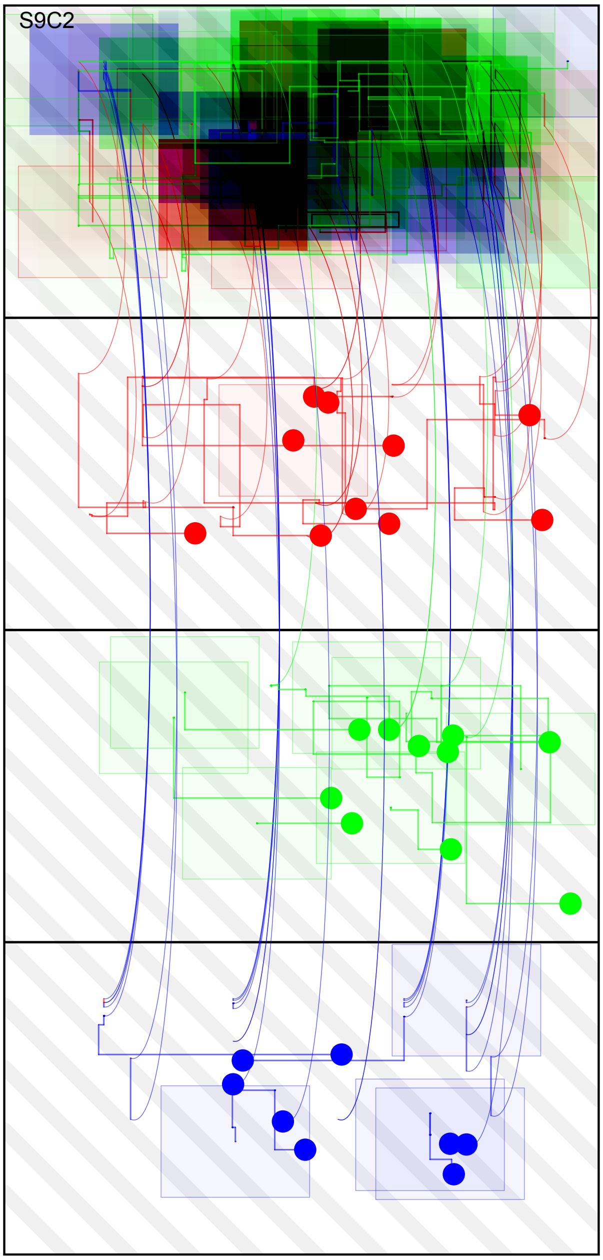

S6C1
S6C2


S7C0
S7C1


S7C2
S8C0


S8C1
S8C2


S9C0
S9C1


S9C2

# Appendix C

# Study Two Materials

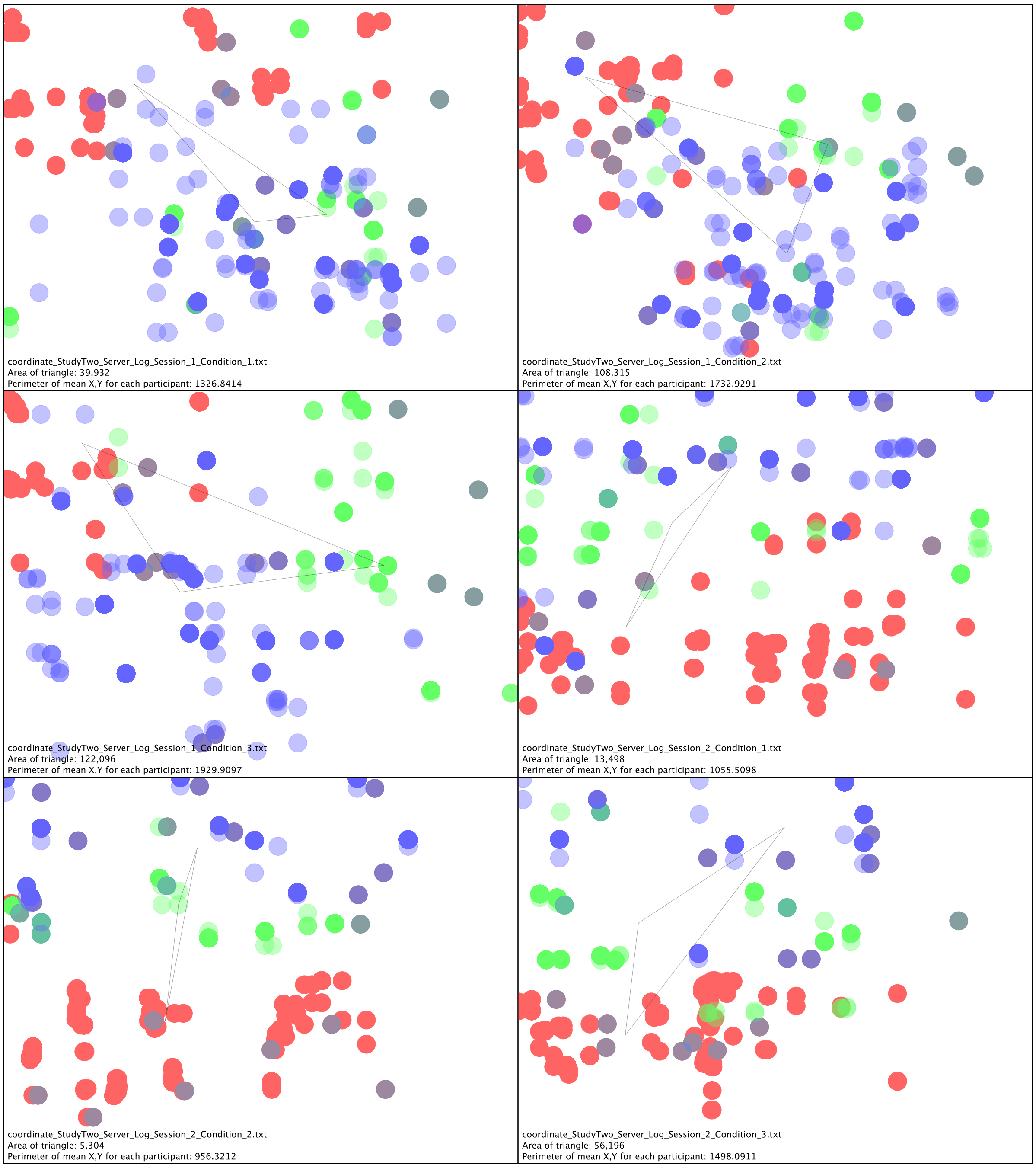
---

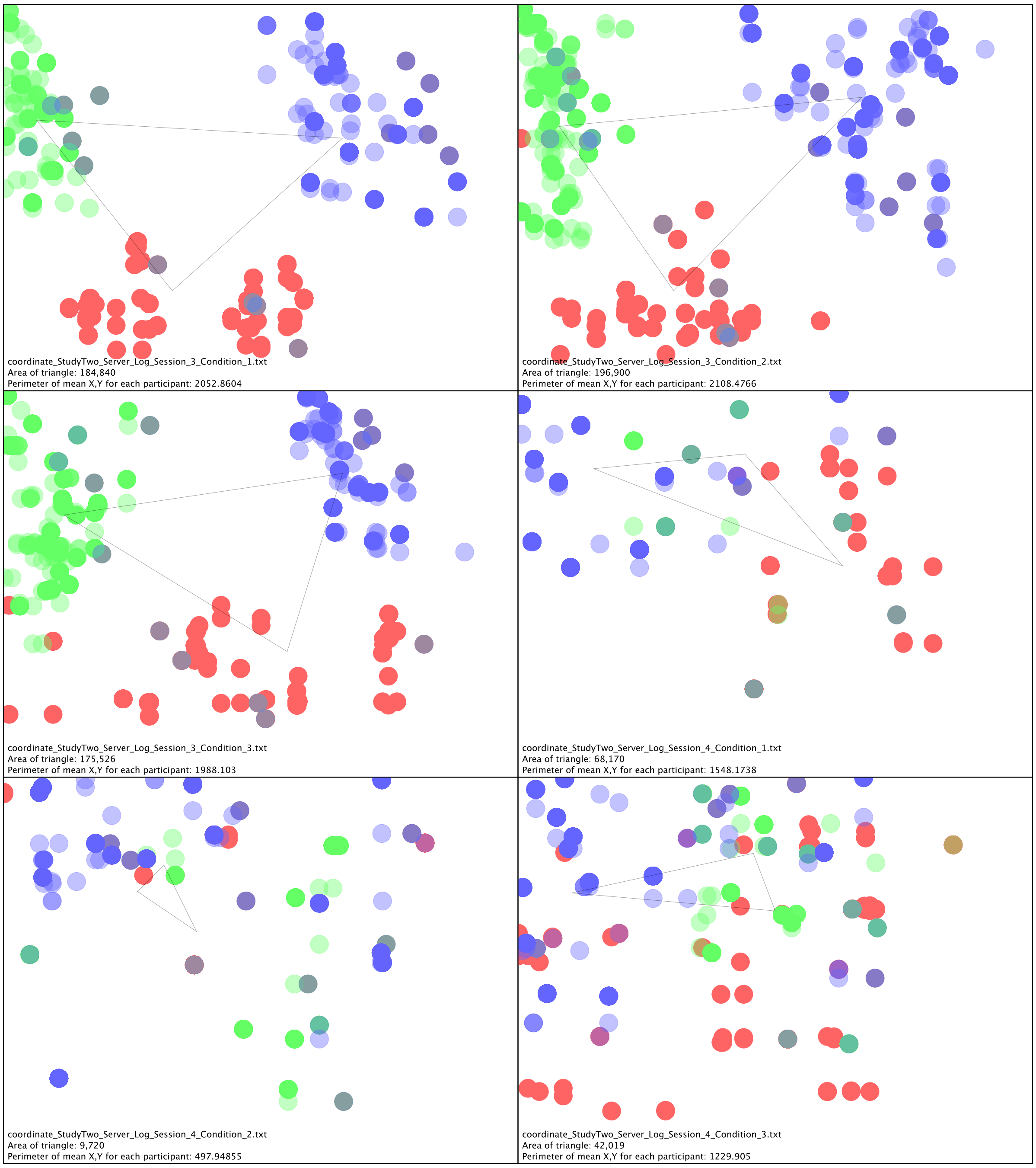
**Study Two Pre-Test Questionnaire**

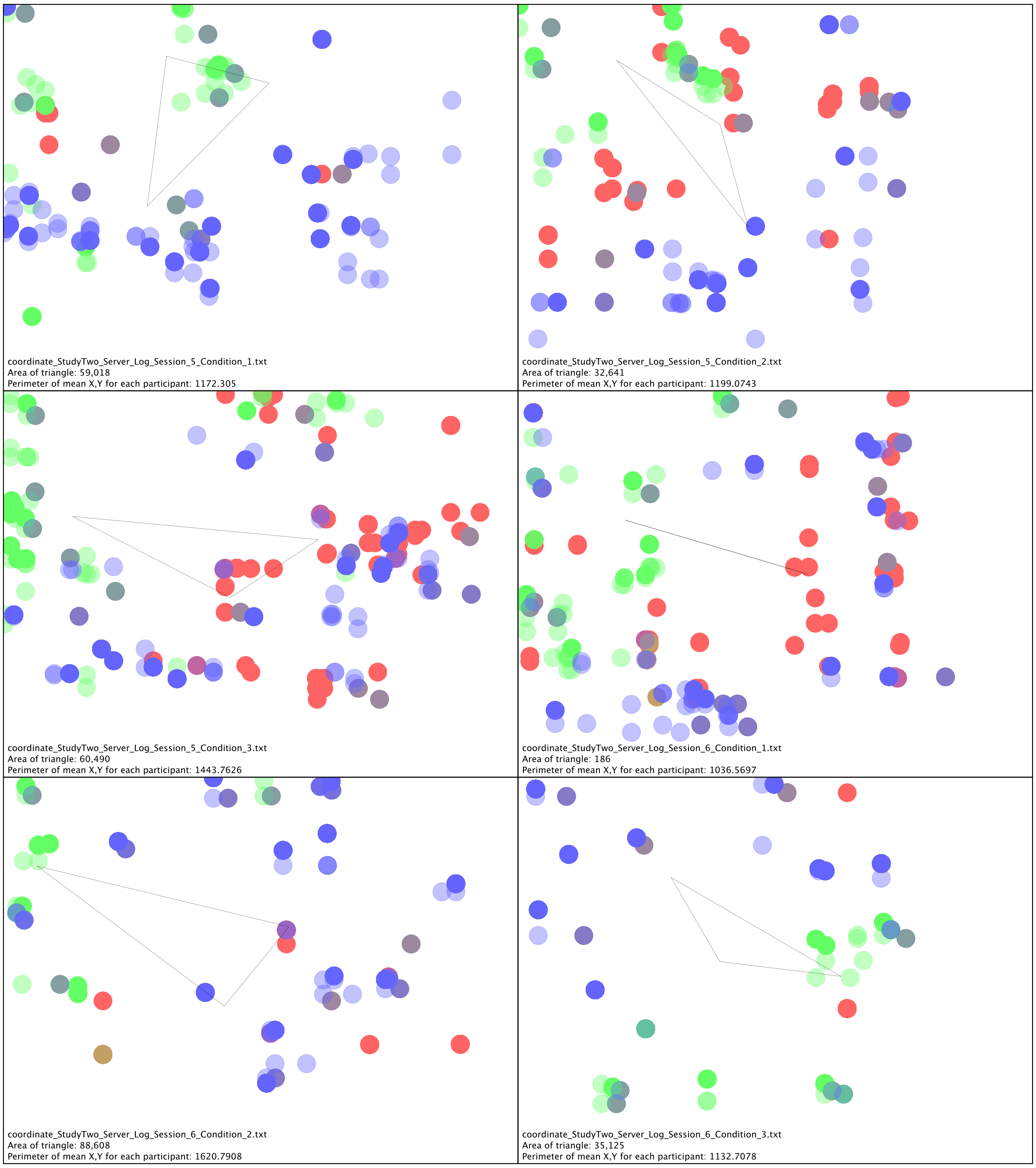
- Gender (Male,Female)
- Occupation
- Age
- Can you play a musical instrument? (Yes,No)
- If yes, how would you rate your musical proficiency? (Beginner, Intermediate, Semi-Professional, Professional)
- Have you ever composed a song on your own? (Yes, No)
- If yes, approximately how many? (1 song, 2-5 songs, 5-10, more than 10 songs)
- Have you ever composed songs with others? (Yes, No)
- If yes, approximately how many? (1 song, 2-5 songs, 5-10 songs, more than 10 songs)
- What types of music do you listen to or like? (Pop, Rock, Indie, Metal, Rap, Hip Hop, R'n'B, Classical, Electronic, Funk, Blues, Jazz, Latin, Country, World, Other - please state)
- How computer literate are you? (Beginner, Intermediate, Expert)
- How frequently are you on-line? (For example, how often do you check your mail, browse the web, make on-line purchases, play online games, listen to music etc.) (Never, Once a week, Once a day, Once an hour)
- Have you ever used an instant messaging or text-chat tool? (e.g., msn messenger, google chat, facebook chat, ICQ) (Yes, No)

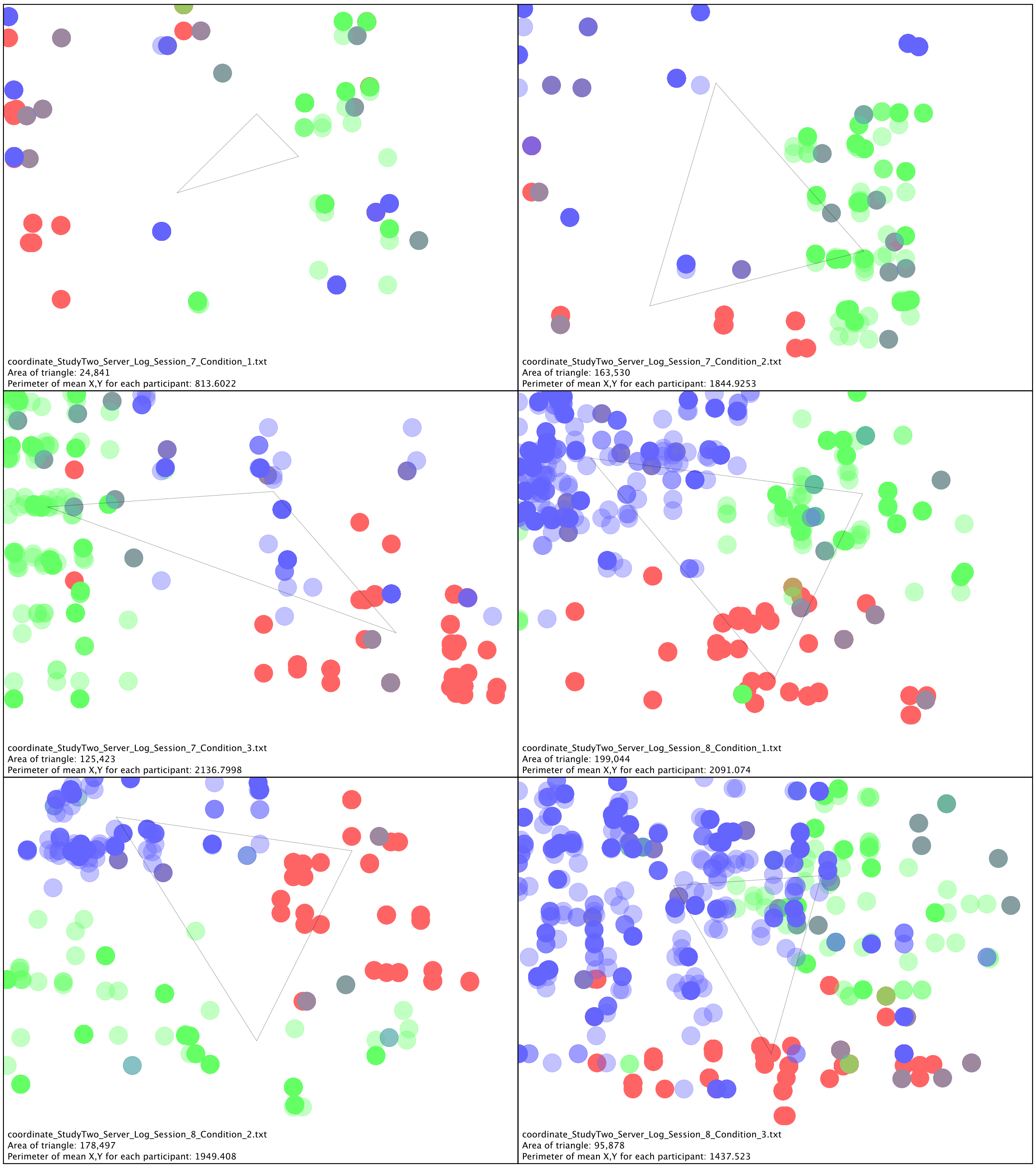
- Have you ever used computer software to collaborate in real time on any tasks? Please use the comments box to give details. (Online gaming, Collaborative document editing, Collaborative writing, Making music, Other - please state).

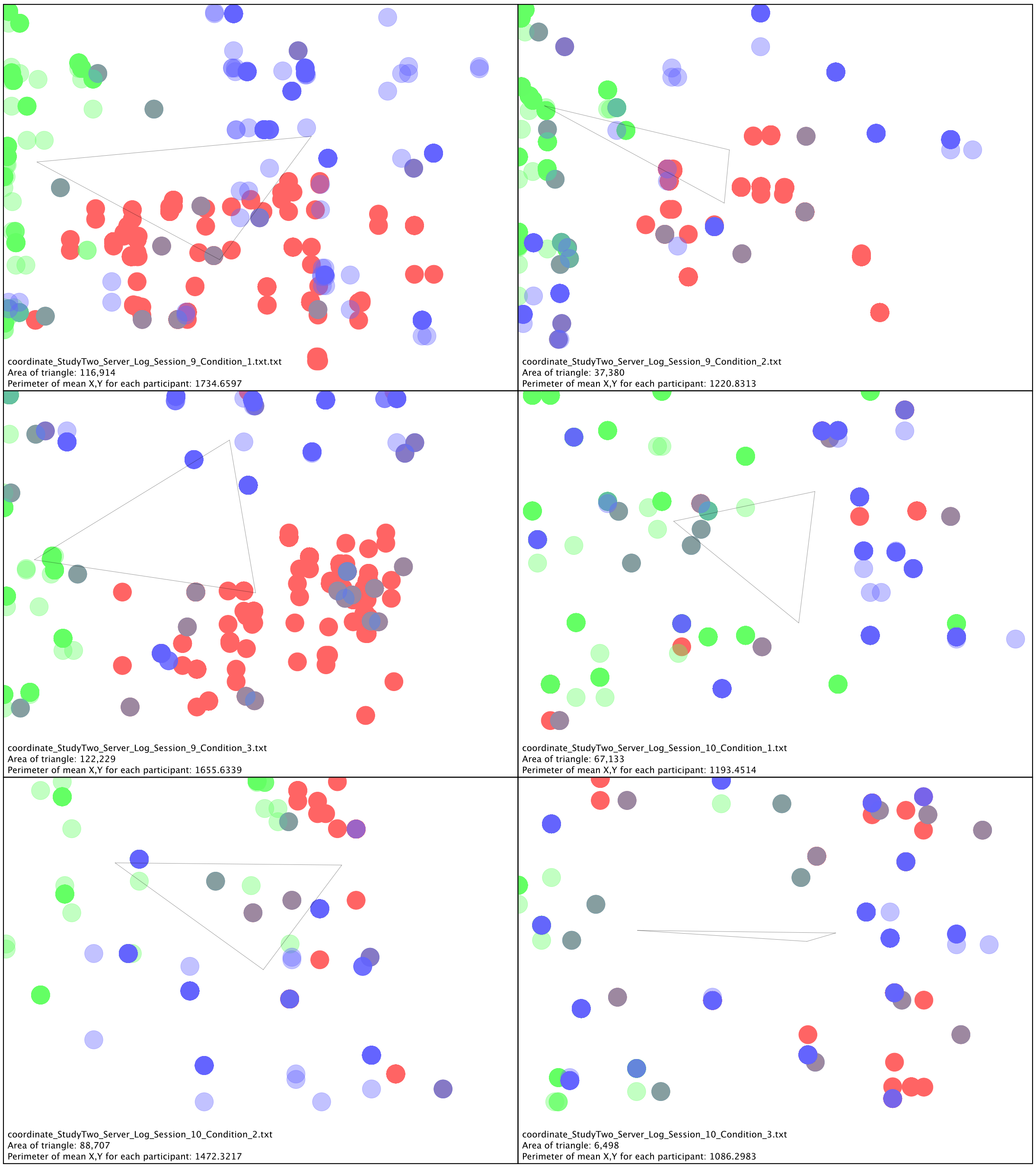
**Study Two Shared Workspace Spatial Visualisations**


coordinate_StudyTwo_Server_Log_Session_1_Condition_1.txt
Area of triangle: 39,932
Perimeter of mean X,Y for each participant: 1326.8414
coordinate_StudyTwo_Server_Log_Session_1_Condition_2.txt
Area of triangle: 108,315
Perimeter of mean X,Y for each participant: 1732.9291
coordinate_StudyTwo_Server_Log_Session_1_Condition_3.txt
Area of triangle: 122,096
Perimeter of mean X,Y for each participant: 1929.9097
coordinate_StudyTwo_Server_Log_Session_2_Condition_1.txt
Area of triangle: 13,498
Perimeter of mean X,Y for each participant: 1055.5098
coordinate_StudyTwo_Server_Log_Session_2_Condition_2.txt
Area of triangle: 5,304
Perimeter of mean X,Y for each participant: 956.3212
coordinate_StudyTwo_Server_Log_Session_2_Condition_3.txt
Area of triangle: 56,196
Perimeter of mean X,Y for each participant: 1498.0911


coordinate_StudyTwo_Server_Log_Session_3_Condition_1.txt
Area of triangle: 184,840
Perimeter of mean X,Y for each participant: 2052.8604
coordinate_StudyTwo_Server_Log_Session_3_Condition_2.txt
Area of triangle: 196,900
Perimeter of mean X,Y for each participant: 2108.4766
coordinate_StudyTwo_Server_Log_Session_3_Condition_3.txt
Area of triangle: 175,526
Perimeter of mean X,Y for each participant: 1988.103
coordinate_StudyTwo_Server_Log_Session_4_Condition_1.txt
Area of triangle: 68,170
Perimeter of mean X,Y for each participant: 1548.1738
coordinate_StudyTwo_Server_Log_Session_4_Condition_2.txt
Area of triangle: 9,720
Perimeter of mean X,Y for each participant: 497.94855
coordinate_StudyTwo_Server_Log_Session_4_Condition_3.txt
Area of triangle: 42,019
Perimeter of mean X,Y for each participant: 1229.905


coordinate_StudyTwo_Server_Log_Session_5_Condition_1.txt
Area of triangle: 59,018
Perimeter of mean X,Y for each participant: 1172.305
coordinate_StudyTwo_Server_Log_Session_5_Condition_2.txt
Area of triangle: 32,641
Perimeter of mean X,Y for each participant: 1199.0743
coordinate_StudyTwo_Server_Log_Session_5_Condition_3.txt
Area of triangle: 60,490
Perimeter of mean X,Y for each participant: 1443.7626
coordinate_StudyTwo_Server_Log_Session_6_Condition_1.txt
Area of triangle: 186
Perimeter of mean X,Y for each participant: 1036.5697
coordinate_StudyTwo_Server_Log_Session_6_Condition_2.txt
Area of triangle: 88,608
Perimeter of mean X,Y for each participant: 1620.7908
coordinate_StudyTwo_Server_Log_Session_6_Condition_3.txt
Area of triangle: 35,125
Perimeter of mean X,Y for each participant: 1132.7078


coordinate_StudyTwo_Server_Log_Session_7_Condition_1.txt
Area of triangle: 24,841
Perimeter of mean X,Y for each participant: 813.6022
coordinate_StudyTwo_Server_Log_Session_7_Condition_2.txt
Area of triangle: 163,530
Perimeter of mean X,Y for each participant: 1844.9253
coordinate_StudyTwo_Server_Log_Session_7_Condition_3.txt
Area of triangle: 125,423
Perimeter of mean X,Y for each participant: 2136.7998
coordinate_StudyTwo_Server_Log_Session_8_Condition_1.txt
Area of triangle: 199,044
Perimeter of mean X,Y for each participant: 2091.074
coordinate_StudyTwo_Server_Log_Session_8_Condition_2.txt
Area of triangle: 178,497
Perimeter of mean X,Y for each participant: 1949.408
coordinate_StudyTwo_Server_Log_Session_8_Condition_3.txt
Area of triangle: 95,878
Perimeter of mean X,Y for each participant: 1437.523


coordinate_StudyTwo_Server_Log_Session_9_Condition_1.txt.txt
Area of triangle: 116,914
Perimeter of mean X,Y for each participant: 1734.6597
coordinate_StudyTwo_Server_Log_Session_9_Condition_2.txt
Area of triangle: 37,380
Perimeter of mean X,Y for each participant: 1220.8313
coordinate_StudyTwo_Server_Log_Session_9_Condition_3.txt
Area of triangle: 122,229
Perimeter of mean X,Y for each participant: 1655.6339
coordinate_StudyTwo_Server_Log_Session_10_Condition_1.txt
Area of triangle: 67,133
Perimeter of mean X,Y for each participant: 1193.4514
coordinate_StudyTwo_Server_Log_Session_10_Condition_2.txt
Area of triangle: 88,707
Perimeter of mean X,Y for each participant: 1472.3217
coordinate_StudyTwo_Server_Log_Session_10_Condition_3.txt
Area of triangle: 6,498
Perimeter of mean X,Y for each participant: 1086.2983

**Post-Test Questionnaire Comments**

Table C.1 presents the comments left by participants in the Study Two post-test questionnaire.

Table C.1: Written comments left by participants during post-test questionnaire

| Statement | Group | Order | **Comment** |
|---|---|---|---|
| The best music | 4 | | The second prbably fitted the film best, and at that point we were probably still concious of other people's sounds, whereas in the third...it went a bit mental. |
| | 5 | | the 3rd condition was the best as we were listening to each other more and the sounds created seemed to flow better. |
| | 5 | | note - think condition 2 and 3 were generally both ranking the same |
| | 10 | | After we've got used with it things became easier |
| | 10 | | got to enjoy the way the program worked towards the end |
| | 10 | | In condition 3 my headphones seemed too loud and there was no way to turn them down without affecting the mix. |
| I felt most involved with the group | 4 | | The first as we were all figuring out how to balance the group. I feel as if one participant become overwhelmed at the noisy third, so I slowed the tempo to see if they could get a hook onto what was going on. |
| | 5 | | note - condition 2 and 3 could have been ranked the same (high) |
| | 5 | | i felt equally involved in all 3 |
| I enjoyed myself the most | 3 | | actually I enjoyed each session equally... |
| | 4 | | The second as we became more familiar with the software and being able to automate the sound manually as it played. The third was fun but difficult to get heard and take control of the musical structure. |
| I felt out of control | 3 | | I didn't feel out of control just finding my way at some points |
| | 5 | | i felt in control most of the time |
| I understood what was going on | 3 | | I wasn't always sure who was making which sound |
| | 5 | | all the time |
| I worked mostly on my own | 1 | | 1 applies to all three |
| | 4 | | People kept thieving my effects... |
| I lost track of time | 1 | | 1 applies to all three |
| | 5 | | in all 3 |

Continued on next page

**Table C.1 – continued from previous page**

| Statement | Group | Order | **Comment** |
|---|---|---|---|
| Other people ignored my contributions | 1 | | Don't know what they were doing/thinking |
| | 3 | | I don't think so.... |
| | 5 | | this didn't happen |
| | 5 | | i didnt feel this in any |
| We worked most effectively | 1 | | It all seemed fairly random |
| The interface was most complex | 4 | | There was quite a nice learning curve in the software and really didn't feel intimidating. |
| I had the most privacy | 4 | | I wasn't sure what others could see, or hear for that matter, so I wasn't entirely sure. |
| I knew what other people were doing | 4 | | It became more confusing as it went on in terms of who was doing what, though I could begin to pick out certain musical forms and ideas/styles that came from the other 2. |
| We edited the music together | 4 | | Very hard to tell in each case |
| | | | I feel that there was little constructive editing taking place. |
| I made my best contributions | | | |
| I was influenced by the other people | 1 | | Can't say |
| | 3 | | in each I think... |
| The condition I preferred the most | 6 | | I perfered being able to monitor my own contribution on the headphones before submitting to group/speakers |

Table C.1: Written (typed) comments left by participants in the post-test questionnaire. Incorrect spelling and grammatical errors have been preserved.

## C.1 Video Codes

### List of codes generated in the first stage of coding dialog during interaction

• headphones shutting you off • not being involved with the group with headphones • good for brainstorming • enjoyable • Preferred headphones • Couldn't use speakers to identify contributions • noticing/appreciating other people's changes • Hearing the music as it is made • fun to play with • difficulty hearing what I was doing • Getting three people's ideas • being excluded • not much difference between personal and public with speakers • Different listening sensation • working together • focusing • density of composition • felt like they were working together • potential of the software • having too many modules • working together • getting busy • sharing • Attending to texture of sound in headphones • Enjoying sonic field with speakers • Could hear better with headphones • concentrating more with headphones • being influenced by others more with speakers only • not remembering what had been made • mixing • asking to change things • Listening closely in headphones • listening to each other • taking things in and out • intensity of working on headphones • hearing alongside other people • turning the volume down • not doing something if someone has done it first • remembering to turn stuff off • hearing changes • muting • combining contributions • rehearsing together in a digital way • putting things in and out • "creating sketches, rough ideas in headphones" • like being in a rehearsal room • hearing your own ideas before • Monitoring your own work • refining ideas before sharing • using personal to make sure things are on time • Controlling what you hear with speakers • personal channel for 'setting things up'. 'trying out stuff' • feeling like being in a chatroom / on the internet • picking out what my instruments were doing • using personal channel to hear what I was doing • getting used to using the software • previewing • understanding how it worked • infringing on others • colour coding • giving a final mix • presenting a final piece • sticking them into the public • getting lost, not knowing which one is yours • stealing or having contributions stolen • telling which was yours • not telling what was yours • knowing who's is who's • losing track of who was doing what • knowing what others are doing • losing track of what was happening • checking what was created by others • not knowing who was doing what • other people hearing your ideas • checking what was yours • Constantly changing • working in a specific area • resolving confusion over which sounds were what by using headphones (ability to single out sounds in a different channel is useful for identifying sounds in a dense mix) • self-organising • working in the top left, top right, bottom right • organic • working in different areas of the screen to know who was doing what • working in top left, top right, bottom • Feeling more involved • working in an area of the screen • getting lost in time (headphones) • sounding better • filling gaps in the sound • working all over the place • putting stuff anywhere there was space • doing the bass • doing the drums • experimenting with weird stuff • using extant musical knowledge • identity with colours • doing the kicks • becoming more familiar as a group • cloning to get double tracks • doing melodies • getting better at making the music gel • doing the snares • tidying up the screen • doing the cymbals • doing the drums, textures, melodies • easier to hear changes in headphones • collaborating more with headphones • changing each other's sounds more in headphones • not asking each other • not talking much • wanting to create roles • focusing on specific elements (beat, sounds) • fiddling • noticing each other's physical movements • noticing people's style and feel

# Appendix D

# Study Three

---

## D.1 Participants

A pre test questionnaire gathered demographic information about the participants. The pre-test questionnaire questions were identical to those used in Study Two (see Section C). Of the participants included in the analysis, 32 were male (67%), 16 participants were female (33%). 46 participants (95%) gave their age. The average age was 27 years (SD = 6.19). The modal age was 23 years. The oldest participant was 45 years old, and the youngest was 18 years. 22 participants (46%) were students. 14 participants (29%) worked as professional musicians or held jobs in creative industries. 7 participants (15%) worked in other forms of employment and 2 participants (4%) were unemployed. 3 participants (6%) did not indicate their occupation.

42 participants could play a musical instrument (88%). 8 participants (17%) rated their musical proficiency as ‘Beginner’ level, 18 (36%) rated themselves as ‘Intermediate’, 16 (33%) rated themselves ‘Semi-Professional’ 3 (6%) rated themselves ‘Professional’ and 3 participants did not respond to this question. 39 participants (81%) had composed on their own, while 9 participants (17%) had not. 2 participants (4%) had written a single piece of music. 7 participants (15%) had written 1-5 pieces, 7 participants (15%) had written 5-10 pieces, and 25 participants (52%) had written more than 10 pieces. 7 participants did not indicate how many pieces they had written. 35 participants (73%) had previously composed with other people, 12 (25%) participants had not, and 1 participant did not indicate either way. 4 participants had jointly composed a single song or piece of music, 16 (33%) participants had composed 1-5 pieces, 5 participants (10%)

had jointly composed 5-10 pieces, and 13 participants (27%) had jointly composed more than 10 pieces of music. Ten participants did not provide a response to this question. Figure D.1 shows the participants reported musical preferences.

Two participants (4%) described their level of computer literacy as 'Beginner' level, 32 (67%) described themselves as Intermediate and 14 participants (29%) described themselves as 'Expert'. All participants used the internet. 17 participants (35%) responded that they went online 'Once a day', and 31 participants (65%) stated being online 'Once an hour'. All 48 participants (100%) had previously used an instant messenger application. 19 participants (40%) had played online computer games. 20 (42%) had used collaborative document editors, 12 (25%) had used collaborative writing editors, 13 (27%) had used collaborative music making software, and 5 participants (10%) had used 'Other' forms of collaborative software, such as 'arts experiments', 'web-cam chat', 'creative theatre making', 'MMORPGs', 'LucidChart' and 'Google Documents'. One participant stated using video editing software in a 'collaborative situation'.

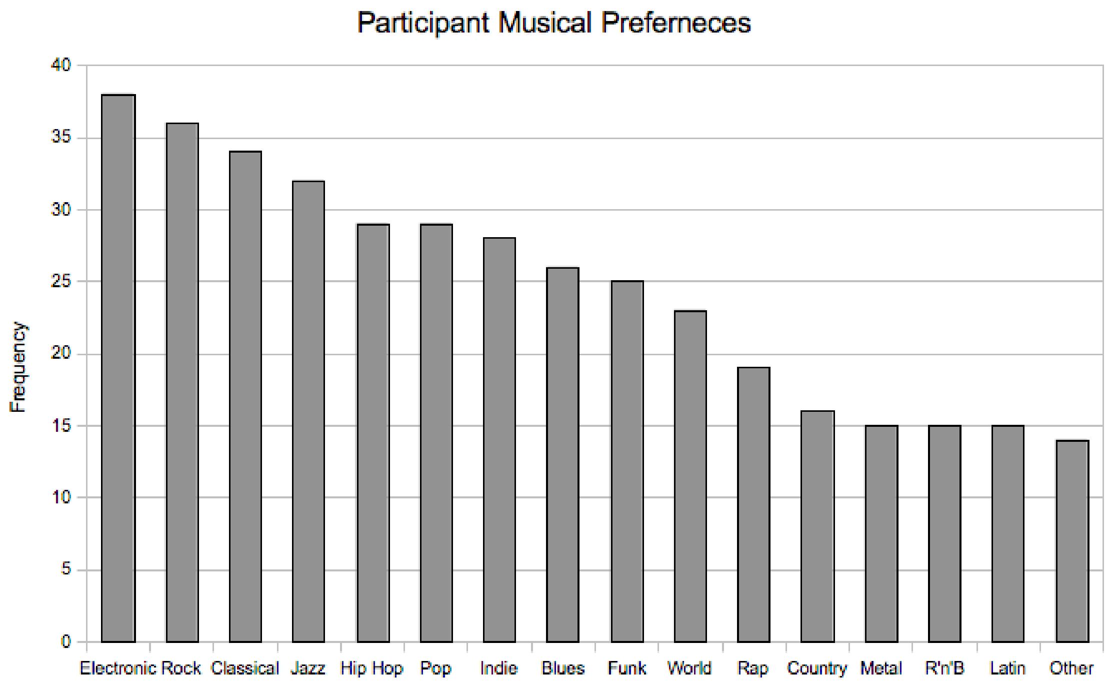



Figure D.1: Participant reported musical preferences


Figure D.2: Screenshot of Study Three server window

## D.2 Server Window

## D.3 Additional Thematic Analysis materials

**List of initial codes generated in the initial stage of coding dialog during interaction**

• vocalising • housecleaning • tidying up • being chaotic/a mass • too much stuff happening • soloing as more/less collaborative • owning - your things • putting things together • discussion of sound • reference to specific module • discussion of musical parts • comment on musical parts • Statement of preference • sounding good • positive statement • discussion of what one did • comparison of software to other programs • discussion of software possibilities/potential • discussion of software • problems with the software • hitting limitations of software • features • discussion of approach to using software • soloing • asking for help with the software • identification of others • providing help with software • revealing/sharing/disclosing knowledge of software • uncertainty regarding consistency between users • Video loop • roles • thinking aloud • invitation to make changes • statement of activity • statement of intent • planning • discussion of style/genre • discussion of tempo • spatial organisation • discussion of co-editing • instruction to others • enquiry about activity of others • asking to co-edit • statement that co-edit was made • ownership • comment on success / comment on failure • comment on quick passage of time • comment on being engrossed • comment on enjoyment • reaction • identifying a sound • confusion - instances of confusion trigger dialog • discussion of experiment • discussion of conditions/changes/differences • discussion of experiment room • comparison of group behaviour between conditions • Comparison to other experiences • discussion of life, skills and knowledge outside of the experiment

**List of initial codes generated from coding the M1 Condition interviews:**

• Private • Fun • Exploration • More Time • Dissatisfied • Software Use • Software • Questions about experiment • Finding a sound source • maybe you can work on your own • nice • exploring • neede more time • not very satisfied • didn't write comments • program was minimalist • experiment • see which bit something is coming from • build lines the bring it back into • it went fine • messing around • needs patience • Song as a whole • thought it would be more fiddley • question about software functionality • experiment • Soloing • having the stems solo • not for you if you have to work alone • fun • not possible to organise in short time • change to make it sound better as a whole • didn't use mute or solo • real-time music • used instrument soloing a lot • soloing instruments helpful for melody • first one did everything ourselves • fun • might have started to organise more with more time • Feature Suggestions • sub groups • normally danuted by loads of dials and things • Wanting the recordng • didn't use soloing • find out which device is doing a particular sound • muteing everybody elese • fun • develop chemistry over time • would be useful to add an instrument that is exactly the same • didn't know how to connect the filter at one point • don't have to do much • send us the audio • used soloing less • used instrument soloing more than person soloing • can listen on your own • interesting • Positive • might have got more interesting songs with another session • would be useful to add an instrument that is exactly the same • wouldn't know how to connect if not told • relativley easy to make music from it • create a track with the recording • client soloing helped get rid of mess • make sure editing right one with solo • listen on own • amusing • drum sequencer is good • wouldn't be so cluttered if more time to discsuss what we wanted • single user version • if on own might layer it a certain way • everybody has same instruments • getting copy of music • easier to focus with client soloing • should have option to preview before adding to the pool • enjoyed it • some parts sounded alright • Problems • maybe include a keyboard • Chaotic • wouldn't have it all clogged up if working on own • huge potential • can I get the music? • didn't use soloing • listening in private • really fun] • went better than expected • Spatial • minor problems • would be useful to see other people's mice • chaotic • two copies going to the same filter • same thing with pro tools or logic • never listened anyone • make everyone hear when you're ready • really fun • went really well • bottom right • problems to do with the instruments • maybe 3d instead • chaotic • two copies an octave apart side by side • problems of connecting through internet • didn't use instrument soloing • function so not everybody has to listen to you messing around • fun • had a grea ttime • the right part • problem of making more synths to make a chord • rotating around a table • first time was chaotic • handy to be able to move things around • unique • didn't understand insturment soloing • add it, kind of • cool • surprised how quickly could create good loops • keeping my stuf in the top • too repetative • different interface ideas • everyone doing things at the same time • features and use • lots of potential • used thing on left (client soloing) more than instrument soloing • really fun • would love to use in my music • bottom • goes haywire • rotating like a cloud • more messy • slowing down to get a longer loop • introductory, user firiend collaborative music software • Research • Questions to researcher • difficult because always on • Planning • great experinece • experience was nice • right hand side • melodic sequencer was difficult • displayed as different perspectives • chaging pattern behind beat to create something • was explained to be really simple • does researcher use the software? • easier to tell who was doing what in second on because of soloing • made plans in second one • was good • nice to have things set up properly • stayed at the bottom • difficult to pin-point pitches with sliders • interest in expeirment • hear your things thing and compare with • talked to each other helped in terms of planning • nice experience • top left side • not sounding very pleasant • authorship might be nice • house cleaning • used in schools • interest in research • seeing how something contributes • deligated in second one • music could be ussed somehow • bottom left • no problems working together • colour code system • can be used as an instrument • used on Wii, with cute characters and tweaked graphics • what are you desigining the software for? • Politness • used solo if I wanted to adjust something • no plan • was interesting • stuck to the right middle • not soloing properly • maybe colour changes after people have been editing • not just used for editing • releasing software? • how long did it take to make • no body is daring to do something • understood more in first (with soloing) • like how it works as a drum machine • put it in one of the corners which was useful • soloing worked for some not others • is it for commercial purposes? • polite • comparing sounds • sounds very good • ok, this is the chord, it's coming from there • soling chains was confusing • option to change size of windows • interest in experiment • strangers • listen to what other people are doing • top left • software problem • choose to see your stuff in different ways • Easy to learn • Listening • influence of experiment • would be more critical if wasn't strangers • mostly soloed on synth • Negative • started bottom right • software problem • feature suggestion • software was easy to pick up • stop adding, and listen • planning to market software? • being shy • soloing to find a melody • some parts didn't sound very good • someone took my position • high tempo was glitchy • seeing scale on sequencer • didn't need knowledge of signal flow • listeing to the whole without myself • interest in experiment • didn't want to spoil something • soloing different ingredients • top right • couldn't see connections when it was crouded • should make other boxes disapear when soloing • software was simple • taking time out to listen to other people • interest in experiment • soloing provided relief • Difference Between Conditions • bottom part • difficult to find your way around • difficult finding pitches on the step sequencer • anyone can use it • listening to things individually • interest in experiment • muted or took the volume out of things without Client Soloing • had more practice in second one • bottom • took time to find what I wanted to edit • should highlight connections as well • listening to individual elements • questionnaire was confusing • soloing groups of modules was more important than soloing people • were better at it second time • playing with the space • couldnt use delay tools • scrolling • like idea of forcing people to listen to others • comment about questionnaire • Making Changes • solo just for us • might have used soloing more if it was added second time • stuck to same place • trouble making connections • scrolling could get confusing • questionnaire was repetative • I'm going to change this now • taking out layers and groups made mixing easier • Soloing Self • made plans in second one • Looking • top left • routing didn't always work • make window you are working on bigger • People within the group • interest in experiment • let everyone know what's happening • taking out layers to get basslines and rhythms to mix • listen to my own and figure out what I was doing • first one did everything ourselves • looked at the lights to see what they were doing • noticed working in separate places on the screen • it was always busy • could be more user friendly • individual styles • was interesting • make sure editing right one with solo • soloed a lot • used instrument soloing to know what I was doing • doing own thing in first one • looking carefully for slide movement before co-editing • definitely working in particular areas of the screen • takes time to find the notes • spaces for different instruments • thought people had different things and sounds they liked • would like to see summary of PhD • used soloing for instrument or effects • listen to my own and figure out what I was doing • deligated in second one • looking at things • doing bottom • trouble finding the note • more physical interface • individual styles • soloing to separate • started listening to myself • all using the same instruments in the second one • see your synth next to other peoples • doing top • jumped when changing tempo • walking around • Introducing Contributions • Onine • solo to hear only instruments • see if my sound went with other people • had one box each in second one • noticed tweaking at same time as others • same arrangement for most of the tunes • out of modules • should have option to preview before adding to the pool • putting volume down when adding something • if you could use it online • soloed when the option was available • see if I am making loads of noise • had many in first one • arrangement just happened • terminology issues • personal view before adding would be good • waiting for one person to do one ting • fine tune with soloing • solo myself so I know what I'm doing • working together in second one • bring all your stuff to one side • crashing problems • could be interesting to have two lines to same effect • Comparison • Making Music • adding something • Creative Process • needed to solo and put volume up to listen • good to know what I was contributing to the whole • first one we were kind of independent • Chance • started on right • software problems • common to have sounds going through one effect • reactable • evolves • taking out the layers is very important for music • soloing is a good option • can't solo in one version • turned out better by chance • no one else noticed bottom corner • problems with interface • crossfader between synths • like instant messaging, but with synthesizers • inspired by what you hear • need to relax ears every few bars to welcome new material • didn't find any difference apart from soloing • making something at random • grouped things in certain areas • editing tool tool for after • a live jamming program • inspiration from outside • had a beat in my head • knew more what was going on in the second condition (without soloing) • grouped step sequencers • like playing in a band • might effect the way other people do things • Reflection on music • just expeirmenting • understood more in first (with soloing) • Roles • grouping modules so you could see what they meant • is there toher stuff like this? • leads to unexpected things • like a gian breathing • I wasn't thinking too much • missed relief of client soloing • no roles • doing right hand side • Discussion of Software • coming from playing in bands • dance track • satan band • using intuition • soloed when the option was available • someone start with the synth • knowing where people were on the screen • share riffs with other people • compared to working alone • hit record on our hands • a few monsters • got really opportunistic • Screen real-estate • people more focusing on one in second time • someone do drums • changed position on screen • mostly for a live set • digital jamming • building on each other • like a giant sleeping • Longer Sequence • confusing because of screen real-estate • first time was chaotic • someone do something else • bottom right • interfaces for people who are not musicians • simlar to Max/MSP • responding to each other before finished • evolving loop throughout • wished there were more beats/longer sequence • screen real-estate was an issue • second time decided to be more minimal • they'll do that, I'll do something else • top left • would take longer to explain to someone who wasn't a musician • interactive sequencer • changed many times for beginning to end • longer loops • ran out of space • changed in second one • if someone was doing the bass I'd do something else • quite at the bottom • too hard to work in studio with this software • compared to working together on one computer • interesting tempo changes • Uncertaintly • wanted longer sequence • stacking things on top of each other • second was more consistent • orgnaised the chords • top right • creative tool • reacting to other people's beats • uncertaintly about muting for everybody • longer sequence might detract from simplicity • space would always be limited unless a big screen • first take was more melodic • not remembering who it was • assumed from direction of screen from time before • flowing • didn't know purpose of effects • would be nice to have a longer sequence • filled up the screen • second was jst an experiment • organised master delay • turn taking with Ableton Live • really easy • not sure if it soloed for everybody • just sixteen beats • things were overlapping • better melodies in first one • mostly doing drums • Awareness of others • Working Together • it was quite interactive • amazing how far you can go with basic things • forgot what Q means • desperatley want to find an empty space for a box • second one was just playing with software • put together the three machines • didn't pay attention to who was doing what • working on something that fits with someone • not like other programs • starts by playing] • Ownership • Constraints • started to fill up • didn't know what people we doing in second one (no client soloing) • doing the bassline • wasn't bothered who was doing what • sometimes you can't work with other people • reminds me of a thing at glastonbury • reflects on method of working • unsure which was mine • working within limitaitons • no one else noticed bottom corner • had no idea what was whos' in second one (no client soloing) • didn't assign roles • didn't have a clue who was doing what without soloing • teaches you you have to cooperate sometimes • is there similar software? • just playing • matters little whos each one was • being creative within limitations • Soloing Individuals • created less boxes in last one • focused more on melody • hard to guess who is doing what • takes less time when you're working with other people • drum machines can be confusing for people who have not used them before • building on each other's ideas • within parameters of software • used soloing to see who was creating which stuff • kept fewer but made just as many • take a pick • wasn't important to solo other people • might take more time with people • different to drum machines on market • starts by playing • just sixteen beats • I could say, ah, this is what Marie is doing • in last one weren't into opening modules, more into making music • messing with the kicks • not bothered if this part was his or hers • throw ideas at each other • important to start with playing when making music • A Mess • limited by one bar of music • see what others were doing • started to listen in last one • experimenting with crazy things • didn't know what people were working on • having other people might distract you • Co-editing • can forget to play • a bit of a mess • can't really experiment with one bar • isolate what each person is doing • different approach in last one • picked some drums • had no idea what was whos' in second one (no client soloing) • were good friends • comfortable editing each other • I couldn't write a traditional song • too many things going on • see what everybody is doing • second had better beats • tried to follow the drums with the bass synth • didn't know who was doing what so much in the second one (no soloing) • would be more difficult with strangers • grabbing same thing as someone else • started to have more fun with it • difficult to build lines with lots of stuff going on • clicked everyone to see what it sounded like • didn't know who was doing what so much in the second one (no soloing) • focused on meldoy • colours to help know who was doing what • comfortable editing each other • took time to start editing other people's stuff • produced something which sounded like a bit of music the second time • not possible if you hear all other stuff together • Preferrence • listened to everyone's seperatley • grouped drums • assumed from direction of screen from time before • thought we worked together well • didn't edit other people's stuff • lots of stuff going on • preferred second one • listen to whatever things everyone was doing • taking over melody • play of what somebody else is doing • wasn't appropriate to edit other peoples • start from scratch and pile it up • listen to what other people are doing • screwing around with drums • awkward not knowing each other • editing other people's because it sounded good • gets messy • preferred second one • loved soloing each and every one of us • Seeing parts fit together • get additional input so you get inspired • editing other people's because it's fun • cluttered • created something more coherant second time • check what the other people are doing • see how parts fitted together • house cleaning • created something more coherant second time • see if my sound went with other people • deleted modules that were not connected • preferred condition with Client Soloing • first time sounded a bit of a mess • preferred first one because of group function to see who did what • preferred music in first one

**List of initial codes generated from coding the M0 Condition interviews**

• fun • A Mess • Video • Seeing • Listening • Organisation • mute • Screen Space • Live • Roles • enjoyed it • fun • just a mess at some points • not looking at the video • needing to see things • everyone listening • organised ourselves • mute • working in any space available • live creation • listening to each other encourages roles • fun • enjoyed myself • can be chaotic • having video as backgroud • seeing things moving • listen and see • finding seat at the dinner table • see it more as a live performance tool • people started adding different things they thought would work • fun • good fun • gets cloggy • forgot about video • singling out modules visually • everyone else listening • divided the screen • live performance • started doing mixing • fun • time went quickly • too many things going on at once • quetsion about video • seeing things moving on the screen • resting back and listening • generally on the right • live distributed concerts • taking a backseta • fun • chaos was funny • didn't look at video once • listen to • maybe on the left • things develop in real time, • doing cymbals, snares • enjoyed it • screen becoming a mess • started listening to each other later • in left • editing in real time • donate on synthesizers • fun • crowded • listening to overall thing • top right • on the fly • harmonising • was good • getting confused • simplistic • listening to yourself • probably on the left • jamming • choose an instrument • good fun • confusing what is what because of limited number of sounds • simplistic • separating it in your ears • screen position was not set • people want human elelent, performance element • changing roles • enjoyed it • screen was messy • Learning • simple • forcing others to listen • control • expectations • stock to bottom • not so much for editing music • changing roles • enjoyed myself • messy if more people • more experience • quite simple • control • expectations • same positions as training • composing in real time • doing the drums • good time • getting hectic • got better • liked simplicity • having lots of control • having a space to return to • not recording • did the first melodic thing • it was cool • intruging • getting more used to it • more simple than other software • wanting more control • assumed where people were • felt like performance tool • doing more of the synths • it was pretty interesting • got complicated • initially figuring it out for ourselves • accessible • wanting more control • longer sequence • working at bottom • improvising together, not composing • using kick • enjoyable • things got familiar and forgotten • just sort of worked • limits of control • longer sequence • working at left • evolving • doing drums • enjoyied • so much stuff going on • figuring out each bit • easy • Individual Style • series of sequences within one box • working on right • make it a live performance • being more confident with drums • enjoyable • getting better over time • easy to use • learning each others' strengths • more steps in sequencer • tended to shift • live performance • melodic and abstract • good fun • didn't know what things would soujnd like • easy for people with experience in music software • got a feel for each person's style • longer sequence • moved around screen • performative element to the experience • taking charge • enjoyed it • get comfortable • it's instant • starting to feel more comfortable with the group • longer sequence • knowing where people were anyway • Co-Editing • did beats • fun • not experience • things are going instantly • didn't want to limit anyone's opportunity to experiment • longer sequence • working in different places • didn't want to limit anyone's opportunity to experiment • did bass • fun • learning curve • sixteen steps is limiting • stayed in same position • happy to let people work on things anyway • getting basis of tracks down • fun experience • not having control or experience • becomes boring • forcing arrangement of windows • might be different if online • one would do bass • enjoyed • take a while to get used to the step sequencer • longer pharase • minimising modules • being shy • one person would do melody • Soloing • Collaboration • easy to learn quickly • Chat tool • Soloing People • Introducing ideas • changing resolution of sequence • on the right • not wanting to change other people's • doing beats • didn't use Client soloing • collaboration felt natural • some software is easy for beginners • chat messages go unnoticed • live music • hearing individual people was handy • put it live • making longer sequences by alternating • whoever got there first • using other people's stuff • doing a bit of everything • tried the solo • kind of always works • chat messages become unsynchronos • who did what • throwing in a sound • only so much you can do to the loop • screen sharing is useful • editing other people • stopped feeling like roles • not sure if soloing was shared or individual • did OK even though didn't know each other • linking messages with sounds • figuring out who was who • put it in • need to go to the next sequence • right hand side • better to work together • working on the drums • not sure if soloing was shared or individual • melding more • messages lose their meaning • hearing sounds from each person • not interferring • getting worse instead of better • working along bottom • going straight to editing someone's stuff • working on details • didn't use soloing • collabroation • chat wasn't useful • identifying others • things become established without discussion • putting things on screen in relation to seating • taking too long to ask to edit someone else's • gravitate more towards melody • soloing was handy • better second time around • disconnection between chat messages and music • awareness of others • difficulty making melody • using screen to distingusih who was what • changing, rather than asking for permission • inclinded towards melody • would have appreciated soloing after familiarity • can't do this with other people • chat notification • didn't see the Solo as being helpful • difficulty making melody • neurolinguistic programming and screen position • editing other people later • tweaking melody • didn't use soloing • stamping authority • chat tool doesn't connect to music • solo wasn't representative of what people were doing • difficult to use synth • running out of space • working for the best outcome • I was the drums • Soloing is useful • getting hits as to where it was going • chat tool didn't scroll properly • more representative of what they wanted to be part of the piece • synth is just luck • who was where? • attributed contributions to people • got into rhythms • using solo to take things out, listen and put back • looks between people • Identifying Sound Source • manually scrolling chat • hearing things for yourself • difficulty making melody • bottom right • at the end was more common (shared) • I was drawn to the synths • used soloing a bit • would be more organised with more time • finding out where a sound was coming from • things disappearing • see what something sounded like • problem finding harmony • bottom left • someone else using my box • making whatever • soloing didn't make a difference • wanted to play • using Person Soloing to identify sounds • not being able to type quickly enough • isolate each other's stuff • hard to find the right notes • having too many boxes at the same time • being anonymous • dividing instruments to get an idea of what we were working towards • soloing didn't make the group more collaborative • can't beat being around people • using Person Soloing to identify sounds • get into the intricacies of what each other was doing • Element of Chance • hard to find notes • decided to divide the screen and make things easier • using clone to make a modification to someone else's work • getting confused about roles • solo self • being with people within the technology • using soloing to restrict field of possible sound sources • element of chance • hard to do melodic stuff • section each, plus an effects section • modifying someone else's work • supposed to do drums • soloed at end • collaboration is better • identifying a sound/source • using keyboard keys • running out of space • moved someone else's • doing drum fills • did use soloing • organic feel • problem finding a sound • sometimes couldnt' see notes • divided instuments second time • moved someone elses' by accident • doing the chords • work around not being able to solo • doing a jam • wanting to change a specific sound • hard to find the right key • didn't mute or delete each other • listening to others • didn't contribute as much second time, as liked what other people were doing • using soloing to pin stuff out • hard to find notes • deleting other people's stuff • listening to combinations • more of a flow • using client soloing to find out where a sound was coming from] • Soloing Specific Modules • repetative • trouble selecting notes • maybe more likely to co-edit if more annonymous • viewed composition as a whole • things changing when you weren't listening • ebb and flow in music • using solo to cut shit out when hectic • would be interesting to solo one thing at a time • repetative • notes seem random • accidentally changing someone else's • viewed composition as a whole • using solo to create more fiddley parts • not knowing what others were doing • using solo regardless of who • would be interesting to solo one thing at a time • repetition • wanting to harmonise • editing each others • soloing to work on a melody • eye contact • isolating sounds • a solo button for each • hard to make interestingly melodic stuff • not wanting to mess with other people's work • soloing so as not to clash • knowing what sounds good etc • comparing sounds • soloing a single object • more used to thinking of notes than slider positions • might mess with other people's stuff in the future • soloing to find the right notes • wanted more discussion • having a way to isolate something so no one eles could change it • soloing specific instruments to make a complex part • can't see the notes • tingering with other people;s • was hard to hear things without the solo button • talking while doing it would have been better • didn't always know where sounds were coming from • solo just for yourself • hard to make harmonies • his bassline • confusing software error with actions of other users • would have liked to solo individual instruments • Screen Space • solo more than one tihing • Making Bad Contributions • hard to go between note increments • not messing with his basline • Comparison • making a big change might have been weird • not enough screen space • solo more than one instrument • disrupting • using increment number to find notes • question of editing stuff together • similar to max/MSP • keep screen neater • soloing certain instruments • making it bad • key next to the slider • don't want to invate private space • more interactive than logic • Discussing who did what • soloing a group • people responding responding to mistakes • more fun than looking at logic • talking about who did what • avoiding bum notes • like max/msp • changing tempo • putting things in at a quiet volume • not sharing the mouse • talking about tempo • Getting to know people • similar both times • add sounds that are not alient • software problems • annonymity • compard to pedal rig in guitar • talking about what happened • interesting with people you don't know • did similar to first time • removing things if they were intrusive • overloaded software • Suggested Features • annonymity • like a game • discussion could have changed direction of music • getting to know people thorugh making music • didn't notice a difference between conditions • making bad contributions • problem of it always running • rather play a melody first • maybe more likely to co-edit if more annonymous • facebook app • feeling in the middle of something • more used to playing with people you know • not much difference between conditions • didn't want to mess it up • routing was difficult with lines • play a melody before programming it • laptop garage jam • meeting other musicians online • laggy • play a meldoy without running the sequence • Online • like three DJs mixing together • laggy with too many samples • additional features • online system/community • software is interacting producing, arranging • colour coding • learning about people • software died • quick auditioning of sounds • internet based would be good • routing through effects • colour coding • learning about people • poor sound qualty • additional features • online bands • middle thing is your mixer • visual cues • nice way to get to know people • poor sound qualty • play something into the loop • online might be chaos • plug into the mixer • individual coloured boxes • corporate tool for bonding • good to have other people's ideas • Being Shy • poor sound qualty • additional features • online interaction with specific people • simultaneously see the same thing • visually easy to remember with colours • ice breaker at a confrence • good to have other people's ideas • starting to feel more comfortable • poor sound qualty • panning • remote interaction with friends • signal path sort of thing • boxes are coloured individually • getting to know people better • inspired by others • waiting before changing someone else's • poor quality headphones • additional features • online interaction • felt like a game • coloujr was helpful • no idea about the technology or people's backgrounds • not offending • poor quality headphones • realtime sampling • online • new garage band • colour is helpful to make connection • not wanting to mess with other people's work • poor sound qualty • microphone • editing the same thing could get annoying • connecting computers via midi • different colours • being polite • difficulty with slider • additional features • musical paint • turning on to see colours the whole time • Experiment • Good features • don't want to invate private space • problem with slider • wanting more instruments • live, real-time music paint • different colours • questionnaire isn't a good way to evaluate • liked chat function • if it was friends it would be differnet • couldn't get right levels then put to public • external controller • game • questions at end were difficult • good setup • you would mangle and mash your friends stuff • not much feedback from interface • additional features • collaborative music program • interested in the report • clone feature was cool • can't make musical phrases • additional features • similar to Max/MSP • interested in the research • couldn't drop in a beat • additional features • joing together objects that do things • what is the purpose of the software • Strangers • Communication • couldn't follow traditional dance music style • more options on things to put in • box with instruments, plugging them together • three is a good number of people • no idea about the technology or people's backgrounds • forced to communicate without Soloing • software problems • keyboard shortcuts • reacatble • identifying others • personalities are a big variable • not knowing what other people liked • discussing what to make • problem with effects routing • touch screen • limitations of other electronic music making • trying to find out who was doing what • purpose of experiment • if it was friends it would be differnet • Music Made • missing some features • Privacy • screen sharing is uncommon • not knowing who did what • what is the idea of the software? • sounded dubsteppy • Problems • Software • automation • soloing things out • screen sharing in other software • who did what • interest in experiment • liked the beats at the beginning • inconsistency • didn't use the clone • different sections of music • listeining to sounds before putting them into the pool • never used collaborative software for making music • just knew where people were • interest in experiment itself • preferred tempo • want something to go back to • using clone to make a modification to someone else's work • separators for different musical parts • hearing what you were doing in private • online gaming • not using soloing to hear specific people • interest in experiment • got a bit hard work • creation of object does not always mean the person who was editing • what is the idea of the software? • touch screens • audition button • plugging things into the centre • hidden agenda of expeirment • preferred the middle • creation of module isn't a good indicator of ownership • getting a copy of the software • adding structure • preview before sharing • similar software? • Positive Comments • interest in research • constraint • don't know where it's going • can't get certain sounds you want • muting centre • moving between different musical parts • see something before hand • compared to max/MSP • enjoyed simplicity of drums • interest in experiment • constraint • good rave riffs • hard to make a connection with what is happening • complicated machines • simplified synthesizer to compliment simple drum machine • personal channel • Visi/Max/MSP • surprised there are not more similar programs • hidden agenda of expeirment • limited musical style • some cohearant bits • will software be available? • synth could be more 'about sounds' rather than notes • hearing what it sounds like with the rest • could use three computers and midi • nice way of doing music • interest in research • not allowed • more like sounds • clone function • different screen layouts or views • needing privacy • plugging things into the centre • perfect for club • constrained • sounds industrial • depends who software is designed for • more global parameters • previewing by yourself • not much collaboration in a DAW • first program like this • limited • sketch • interest in code • seeing other people's cursors • soloing to try things out • difficulty working together with a DAW • congratulations • constrained by the four four • good for making grime • faster lfo • using soloing as personal channel • lots of possibilities • constrained by four four bar • driven by rhythm • talking about who did what • Musical Creativity • chat notification • auditioning stuff • software is impressive • limited of software • droping something out • making master effects • wanting to create what's in your head • different sequeces running at different rates • spoiling the composition by adding something • new experience • limitations • unusual musical style • talking about who did what • everyone using the same software, no rebellion • second bar • working on own initially • it's instant • synthesizer was limiting • just chaning things • reflecting on what others were doing • previous experience in bands • horns, vocals • soloing to be more on your own • Know what you did • things are going instantly • limitations • good at bringing in beats • talking about what happened • worked by self with technology • external control • using solo to isolate own stuff • what you did • with friends who don't use this sort of stuff, one person could keep it together • electronica style • building up • talking about what happened • works on own a lot • visual representation of track • instead of putting it straight through • making sure what you are hearing is your sound • you can do a lot with it • can't get certain sounds you want • more of a flow • misses collaboraiton • feature suggestions • soloing meant trying to make own stuff • forgetting what you've made • easy to collaborate • limitations of software • ebb and flow in music • what's inside counts • software functions • using soloing to monitor • soloing to listen to exactly what I was doing • can't make musical phrases • rhythm was important • challenge • mini arrangement view • personal channel • single pointer/finger interaction • accidents • modern technology allows people to get ideas out of their head • flashing a box to draw attention • managing in ear first • Advantages of Collaborative Software • drums are dominant • working within a system] • flash function on box • feature suggestions • start working on your own • sharing a single computer • Listening back to music • Comments about experience • Comparison of Conditions • can come up with something really good if creative • external control • being perfectionist • not sharing the mouse • would like to listen back to the music • feeling like coming out of a dark room • first condition was harder (without soloing) • reflection on abilitiies • mixer • testing • sharing a single computer • feeling spacey • soloing encouraged more individual work • writing music is theraputic • solo a track • listening to your own stuff before sharing • not sharing the mouse • passage of time • didn't contribute as much second time, as liked what other people were doing • hard to collaborate instantly • iPad app would be cool • hearing things in your head, but distracted by everything else • good to have individual screen and mouse • first time not agreeing what to do • microphone for communication • soloing encouraged more individual work • would be fun to do with friends • was hard to hear things without the solo button • wii controllers for music on youtube • mute at the start • ok for having fun • intuitive controller like wiimote • application in recording studio • different effects to fuck with the sound • not much collaboration in a DAW • see it as an iPad or iPhone app