Exploring Collaborative Music Making Experience in Shared Virtual Environments

Liang Men

Media and Arts Technology
School of Electronic Engineering and Computer Science
Queen Mary University of London

2020

Abstract

Virtual Environments (VEs), as media providing high-level immersion, offer people an opportunity to mimic natural interpersonal interactions digitally. As a multi-player version of VEs, Shared Virtual Environments (SVEs) inherit VEs' advantages in enabling natural interactions and generating a high level of immersion, and will possibly play an increasingly important role in supporting digitally-mediated collaboration. Though SVEs have been extensively explored for education, entertainment, work, and training, as yet, few SVEs exist in the field of supporting creative collaboration and as a result, research on the creative aspect of collaboration in SVEs remains very poor. This raises questions about how to design the user experience to support creative collaboration in SVEs. This thesis starts with an introduction and related work. An SVE called Let's Move (LeMo) will then be briefed. LeMo allows two people to interact with each other and create music collaboratively in its virtual environment. Three studies based on LeMo will then be presented: Study I explores how free-form visual 3D annotations and work identity influence the collaboration, Study II and Study III explore how working space configurations affect the collaboration. Results indicate that: (1) 3D annotations can support people's collaborative music making (CMM) in SVEs through five classes of use; (2) group territory, personal territory, and territorial behaviour emerge during collaborative music making in SVEs; (3) manipulating characteristics of personal space affected collaborative behaviour, formation of territory, work efficiency, sense of contribution, preference, and so on. Then an overall discussion between studies is made and further implications for SVEs supporting collaborative music making (and other types of collaboration) in SVEs are given. The findings of this thesis contribute towards the design of Human-Computer Interaction of Shared Virtual Environments focusing on supporting collaborative music making.

Acknowledgements

I would like to take the opportunity to thank them to express my appreciation to all those who helped and supported me during this journey.

First, I would like to thank my primary supervisor Prof. Nick Bryan-Kinns, who provided seamless help and support throughout the four years. Without him, the last four years would be much more challenging and the PhD might be an impossible mission for me. It was my great fortune to know him and to be one of his students. I shall expand my thanks to my second supervisor Dr. Laurissa Tokarchuk, and the independent assessor Dr. Miles Hansard, both provided valuable comments and advice during the PhD. Thanks should also go to my examiners, Dr. Paul Marshall and Dr Matthew Purver for their valuable and insightful feedback, which certainly helped to improve this thesis.

I feel so fortunate to do the PhD in Media and Arts Technology Centre for Doctoral Training (MAT CDT). Feeling of togetherness and sense of belonging provided by all my contemporaries in MAT greatly reduced the loneliness and helplessness during the PhD process. Specifically, I would like to thank Shivani, Raphael, Leshao, Thomas, Louise, Tom, Jacob, Betül, Beici and An, talking and working with all of whom are so great and beneficial to this PhD. I would also like to thank MAT CDT staffs Jonathan, Geetha and Karen for their valuable support for the PhD research.

Thanks should also go to my friends Haoyuan, Bhusan, Zixiang, Jiajie and other friends at Queen Mary, all of whom are always my solid backing supporters. I would also like to sincerely thank my grandma Jun, my parents Baoyi and Xiqin, and my parents in law Xuanfeng and Hong for their unconditional love and support.

Last but not least, I would love to thank my beloved wife Danqi, for her unwavering love and continuing support throughout these years. It was her who brought me light during the darkest times.

This work is supported by EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology (EP/L01632X/1).

Statement of Originality

I, Liang Men, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below. I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material. I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis. I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university. The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author. Signature:

Submitted: 10 September 2019

Details of collaboration and publications: all research and contributions in this thesis and the associated publications are my own work. The research was supported by Dr Nick Bryan-Kinns within the scope of his role as my primary supervisor, and he is acknowledged as a second author in all related publications. Previous publications related to this thesis are described in Section 1.3.

Contents

List of Figures

List of Tables

List of Abbreviations

aME Classification	Annotations for Mutual Engagement Classification Scheme
API	Application Programming Interface
AR	Augmented Reality
CMM	Collaborative Music Making
CMMs	Collaborative Music Making systems
CQ	Comparison Questionnaire
CSCW	Computer-Supported Cooperative Work
CVE	Collaborative Virtual Environment
DMI	Digital Music Interface
DOF	Degrees of Freedom
FOV	Filed of View
FPS	frames per second
HMD	Head Mounted Display
LeMo	Let's Move
VE	Virtual Environment
VMI	Virtual Musical Instruments
VR	Virtual Reality
VRMI	Virteual Reality Musical Instruments
PSQ	Post-Session Questionnaire
SFQ	ShortFeedback Questionnaire
SVE	Shared Virtual Environment
SMI	Smart Musical Instruments
UI	User Interface

Chapter 1

Introduction

Virtual reality (VR) may provide a greater sense of community and more intuitive interactions than traditional media (Wallace & Maryott, 2009). E.g. people have been found to interact with virtual objects and people in virtual environments (VEs) similar to the ways they would interact with real objects and people in the real world (Jackson & Fagan, 2000; Hoyt & Blascovich, 2003). Indeed, whilst many screen-based interactive systems view users as people on the outside looking in (Benford et al., 1995), VR offers an opportunity to immerse people into the interaction and collaboration. Compared to traditional media, VR may provide a novel space for multi-sensory experience (Turchet et al., 2018), a greater sense of community and more intuitive interactions (Wallace & Maryott, 2009), and may offer opportunities for new forms of human-to-computer interaction (Men et al., 2017) and human-to-human interaction.

The potential of multi-user immersive virtual reality to facilitate social activities is well established (e.g. AltspaceVR[1], VRChat[2], Convrge[3]), and VR is argued to be a prospective media for collaboration and community (Wallace & Maryott, 2009). However, there is a paucity of literature on collaboration in VR and only a few VR applications currently exist in the field of creative collaboration, which includes but not limited to collaborative music making (CMM, Brown & Dillon 2007), collaborative sketching (Purcell, 1998), collaborative idea generation (Bødker et al., 2000), and collaborative drawing (Shah et al., 2001). As a result, there are many open research questions on how to design the virtual space and virtual tools to better support these creative collaborations in VR. Understanding CMM in VR would be a good starting point, because music making, as a collaborative activity that relies on common goals, under-

[1] AltspaceVR `https://altvr.com`
[2] VRChat `https://www.vrchat.net`
[3] Convrge `http://www.convrge.co`

standing and good interpersonal communication, has long been a key form of collaborative activity (cf. Titon & Slobin 1996; Bryan-Kinns & Hamilton 2012; Serafin et al. 2016). Its ability to create shared social experiences (Serafin et al., 2016) makes it an excellent activity for research on understanding collaboration in VR. By answering how to better support CMM in VR, more insights on how to support creative collaboration might be obtained.

Collaborative Music Making (CMM) is inherently multimodal involving not only the produced sound itself, but also other presentations such as bodily gestures (Healey et al., 2005), physical activation of instruments (Bin et al., 2017), and written notations and sketches (Thiebaut et al., 2008; Nabavian & Bryan-Kinns, 2006) to manage the joint creation and production of music. Many of these modalities such as body position are facilitated through the physical proximity of musicians. When engaged in music composition and improvisation, composers who are working together may want to communicate in modalities other than sound since that is the primary medium of the creative activity. This raises questions on how to support CMM in virtual environments, specifically in terms of how to design user experiences which support collaboration without being detrimental to the product actually being created. The CMM systems Daisyphone and Daisyfield in Bryan-Kinns (2011) provide people with a shared annotation mechanism and work identity, the former feature enables collaborators to draw lines that are publicly visible and the latter enables them to know the contributor of each music note. It has been argued that music making may benefit from this. With the inspiration from this, part of the aim of this thesis is to explore how similar visual cues (e.g. 3D annotations and work identity) might impact the creative collaboration when it comes to VR context.

Aside from that, part of the aim of the thesis is to obtain a better understanding of the virtual space, which seems to be the foundation of collaboration in VEs because the virtual space is the basic material for all kinds of activities happening in VEs (cf. Raffestin 2012). The real world provides us with a shared, encompassing space, in which humans perceive rich information about the events and objects that can be manipulated and explored. As such, space can be seen as a given material offered to human activity (Raffestin, 2012). The use or recognition of space to communicate the domain of area or possession occupancy ownership is referred as territoriality (Beebe et al., 2000). In collaboration, territoriality, as a result of performing activities in space and a spatial strategy to affect, influence or control resources and access (Sack, 1983), plays an important role in managing interaction and resources (Scott et al., 2004). In contrast to non-immersive media which have very different properties to real-world interaction (Gaver, 1992), Virtual Reality (VR) offers the potential for people to perceive and interact with others and objects in a way much more

similar to the real-world experience. Therefore, an effective arrangement and utilisation of a working space can possibly be a crucial factor to a successful collaboration in VR. Despite the existing research investigating spaces and collaboration (e.g. collaboration via tabletop interface in Xambó et al. 2013; Scott et al. 2004), the role of space and territoriality in VR is still little researched. Hence, exploring this field becomes part of the research aims.

In summary, the research agenda of this thesis is informed by the prospective opportunity VR provides in supporting collaboration, the lack of VR systems supporting creative collaboration (particularly CMM), and the need of understanding virtual space's role in creative collaboration in VR. An overarching research question has been raised: *How to better support collaborative music making in shared virtual environments (SVEs).* To answer this question, this thesis will outline the design and build of a shared virtual environment named Let's Move (LeMo), which allows two people to create a piece of music collaboratively in VR by manipulating virtual music interfaces. It should be noted that instead of making a professional music production tool and supporting a professional level of music composition, the focus of this research is to obtain an understanding of how to support CMM in VR. Though it is for exploratory purpose in the field of supporting of CMM in VR, the findings could possibly be inspirational for other creative collaborations beyond music making in VR, e.g. collaborative drawing. Three empirical user studies based on LeMo will then be presented to explore the design of multimodal support and space settings for CMM in VR, looking specifically at how visual cues might affect collaboration, and space is used by people in a SVE context.

1.1 Research Question

The overarching research question of this thesis is: *How to better support collaborative music making in shared virtual environments?* Next, the meaning of the terminologies used in the above question will be specified. The verb *support* is to offer tools that assist user's awareness, communication and performing during collaboration, see (Nguyen & Duval, 2014). The term *Collaborative Music Making* is a creative social activity, it happens when more than one person compose together in a collaborative way. During the process, people involved in this activity should be enabled to communicate, exchange and share sources and thoughts and be free to do so. More work related to CMM will be discussed in Section 2.3.1. The term *shared virtual environments* is a multi-player version of the Virtual Environment, which defined as a simulated environment in which a perceiver experiences telepresence (cf. Steuer 1992). Different from single-player VEs, in shared virtual environments, players can also experience

presence of other players in the same VE and perform inter-personal interactions (cf. Schroeder 2012).

Based on the challenges and gaps revealed in the relevant work presented above, this thesis will address the following research questions: i) How to aid CMM in SVEs through multiple modalities (e.g. visual cues); ii) How collaborators behave during collaboration in the virtual space, and how to design the virtual space to support CMM in SVEs by using spatial configurations; iii) How to measure and evaluate CMM in VEs.

1.2 Contributions

The main contributions of this thesis are:

- Chapter 3 presents the system LeMo, which fills the gap between CMMs and SVEs, and can function as a basis for exploring CMM in SVEs.
- Chapter 4 identifies five classes of use of 3D visual annotation in supporting CMM in SVEs, three of which are particularly relevant to the future design of Sonic Interactions in virtual environments.
- Chapter 5 identifies: (i) two types of territory and working configurations emerged during collaborative composing; (ii) additional personal space supported individual creativity and increased efficiency with the some side-effects; (iii) a publicly visible personal working space was preferable to a publicly invisible one. Based on these findings, three corresponding design implications for SVEs focusing on supporting CMM are given.
- Chapter 6 identifies: (i) providing personal space is an effective way to support collaborative creativity in SVEs; (ii) personal spaces with a fluid light-weight boundary provided enough support, worked better and was preferable to ones with rigid boundaries ; (iii) a configuration that provides a movable personal space was preferred to one that provided no mobility. Following these findings, four corresponding design implications for shared virtual environments focusing on supporting CMM are given.
- Chapter 7 does further discussion based on the three studies in Chapter 4, 5, 6, identifies the important tole the embodiment plays in CMM in SVEs, proposes ways to construct privacy for SVEs, provides suggestions for building tools in visual and auditory approaches, and concludes implications for CMM in SVEs. Chapter 7 also includes a discussion of the possible value of the implications and suggestions beyond CMM.

1.3 Associated Publications

This thesis covers three empirical studies, all were carried out between 2015 and 2019 in campus of Queen Mary University of London. The majority of the work presented in this thesis has been presented in international peer-reviewed publications:

1.3.1 Published

Men, L., & Bryan-Kinns, N. (2018, March). LeMo: Supporting collaborative music making in virtual reality. In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)* (p. 1-6). doi: 10.1109/SIVE.2018.8577094 [This paper covers part of Chapter 4.]

Men, L., & Bryan-Kinns, N. (2019). LeMo: Exploring virtual space for collaborative creativity. In *Proceedings of the 2019 on Creativity and Cognition* (pp. 71-82). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3325480.3325495` [This paper received Honorable Mention for Best Paper Award. Chapter 5 is mostly covered by this paper except the interview result part.]

Men, L., Bryan-Kinns, N., & Bryce, L. (2019). Designing virtual spaces to support collaborative creativity. *PeerJ Computer Science*, *5*, e229 Retrieved from `https://doi.org/10.7717/peerj-cs.229` doi: 10.7717/peerj-cs.229 [This paper covers Chapter 6.]

1.3.2 Under Review

Men, L., & Bryan-Kinns, N. (2019). Multi-modal approaches to supporting collaborative music making in shared virtual environments. Book chapter submitted. [This book chapter covers part of Chapter 4, part of Chapter 6, and part of Chapter 7.]

1.3.3 Supplementary Published Publication

Men, L., Bryan-Kinns, N., Hassard, A. S., & Ma, Z. (2017, March). The impact of transitions on user experience in virtual reality. In 2017 IEEE Virtual Reality (VR) (p. 285-286). doi: 10.1109/VR.2017.7892288

1.4 Thesis Structure

The following chapters will be structured as:

Chapter 2 presents an overview of related work, including: (i) virtual environments, shared virtual environments and their key features, (ii) creativity, music making and how they are related to VEs, and iii) space, territory and territoriality, and privacy in collaboration.

Chapter 3 explains the methodology for the research and the technical implementation of the prototype LeMo.

Chapter 4 presents Study I, in which two visual features (work identity and 3D annotation) are systematically manipulated to see their impact on collaboration.

Chapter 5 presents Study II, in which 42 users composed music together using three different virtual working spatial configurations. The aim of the study is to observe how users use the virtual space, and the additional personal space with different features.

Chapter 6 presents Study III, in which 52 users composed music together using four different virtual working spatial configurations. The aim of the study is to explore how to provide personal space with a light-weight form to minimise the negative side effects identified in Study II.

Chapter 7 reflects LeMos and findings of Study I, II and III, discusses the embodiment usage, ways to balance privacy and openness, visual and auditory approaches of the three studies and proposes implications for SVEs and VRMIs.

Chapter 8 concludes the thesis, summarises the contributions of the thesis and provides future perspectives on exploring how to better understand and support collaborative music making in SVEs, and potentially a wider range.

Chapter 2

Background

Recall that the overarching research question of this thesis is how to support creative collaboration in SVEs, music making in particular. This question can relate to four research fields: (i) the medium being used - VE, (ii) VEs' key features, (iii) creativity, music making and VEs, (iv) space, territory, and territoriality and privacy in collaboration. Next, this chapter will follow these four topics to explain the relevant works in detail. These backgrounds together clarify the basic principles of research questions and the research design of three empirical studies conducted in this thesis.

2.1 Virtual Environments

This thesis is concerned with understanding how to support creative collaboration in immersive Virtual Environments (VEs). To do so, understanding VEs is important. Next, This section reviews related work on defining Virtual Environments, example systems of VEs, multiplayer version of VEs – Shared Virtual Environments (SVEs), and VEs' key features including presence and its measurement, embodiment.

2.1.1 Defining VEs

The terms virtual worlds, virtual cockpits, and virtual workstations were used to describe specific projects. In 1989, the term virtual reality was coined by Lanier, aiming to bring all of the virtual projects under a single rubric. The term Virtual Environment (VE) can be traced back to the early 1990s (Bishop & Fuchs, 1992), it emerged as a competing term to Virtual Reality (VR), however, both are usually equally used to refer to a world created totally by computer simulation (Luciani, 2007). In this thesis, in most instances after this section, the

term VE will be used to refer to such an artificial environment to comply with Shared Virtual Environments (SVEs), related work of which will be reviewed in Section 2.1.3. Many definitions of virtual reality involved the technological equipment set. Below are three examples of such definitions:

> *The term virtual reality...typically refers to three-dimensional realities implemented with stereo viewing goggles and reality gloves.*
> (Krueger, 1991, p. xiii)

> *Virtual Reality is electronic simulations of environments experienced via head-mounted eye goggles and wired clothing enabling the end user to interact in realistic three-dimensional situations.*
> (Coates, 1992)

> *Virtual Reality is an alternate world filled with computer-generated images that respond to human movements. These simulated environments are usually visited with the aid of an expensive data suit which features stereophonic video goggles and fibre-optic data gloves.*
> (Greenbaum, 1992)

These devices-driven definitions include the notion of "electronically simulated environments" and accessing systems for users to immerse and interact in the VEs. Defining VR by materialising the system makes it simpler for novices to understand VEs, but at the same time, this kind of object-centred view inevitably loses a lot of non-material aspects as well and could not uniquely define VE experience. Specifically, according to Steuer (1992), this kind of definition shows its weakness in three ways. First, a given system is recognised as a "VR system" or "non-VR system" simply by checking whether it contains the essential hardware components or not. Second, this sort of definition failed to provide sufficient meaning that exists beneath the hardware. Finally, it provides small variation across which VR can vary. To deal with these inadequacies, defining it from another way is essential. Defining VR by concentrating on the experience instead of a hardware collection might be an alternative way. Believing that the key point of defining virtual reality is the concept of presence, Steuer (1992) defines of Virtual Reality in terms of telepresence:

> *A "virtual reality" is defined as a real or simulated environment in which a "perceiver" experiences telepresence.*
> (Steuer, 1992)

Steuer believed this definition shifted the criteria of VR from the hardware set to individual experience, provided an experience-based measurement for VR and allowed variation across technologies along with several dimensions. Similarly,

with the immersion as a key component, Sherman & Craig (2018) describe VR by decomposing it into three different components: immersion, sensory feedback and interactivity.

Virtual reality possibly has the greatest potential in terms of facilitating experiences that cannot be encountered in the real world (Serafin et al., 2016). A successful synthesis of an environment requires an understanding of what the environment is made of, according to Ellis (1995), an environment has three parts: content, objects and actors: Content is the objects existed in the environment; Geometry is a field for action, including dimensionality metrics and extent and dynamics are the rules of interaction among its contents (Ellis, 1995).

2.1.2 Systems of VEs

Though VR technology was invented several decades ago, not until recently with the development of technology, and affordable commercial products being released (e.g. HTC Vive[1] and Oculus Rift[2]) had it become accessible for larger audiences and researchers, cf.(Summers et al., 2015). To enable a plausible immersive virtual reality, simulating virtual 3D world and allowing users to see, hear and touch the virtual environments are the key points. As people are visually-oriented creatures, most VR systems primarily address the visual sense. To date, there are two usual types of implementations for immersive VEs, Cave Automatic Virtual Environment (CAVE) and Head Mounted Displays (HMDs), see (Loomis et al., 1999).

Cave Automatic Virtual Environment - CAVE. A Cave Automatic Virtual Environment (CAVE) is an immersive virtual reality environment that contains at least three to six walls, surrounding a viewer with projected images. It was firstly invented by Cruz-Neira et al. in 2014. A CAVE usually gives a much wider field of view (FOV) than most HMDs do. The viewer can also move in a CAVE, and their movements are tracked by the sensors typically attached to the 3D glasses worn by the viewer and the projected images or video continually adjusts to retain the viewer-centred perspective.

Head Mounted Displays - HMDs. Compared with CAVEs, HMDs are more common. HMDs have advantages including smaller volume, lighter weight and lower cost. These features have made it the mainstream in both research and commercial field at the present. An HMD basically contains a pair of lenses, a screen and inertial sensors (accelerometer, gyroscopes, and magnetometer). The lenses and screen together provide a slightly different image for each eye of the user, the larger disparity between the images, the more far away the users

[1] `https://www.vive.com`
[2] `https://www.oculus.com`

feel the virtual objects shown in the image. The inertial sensor provides real-time orientation data, enabling 3-degrees-of-freedom (3DoF, i.e. freedom in 3 rotational axes). Note people all live in a 3D world and interact with all 6 DoFs (3 directional axes and 3 rotational axes), to enable 6-DoF in VR, some HMDs also come with room-scale tracking devices. There are two major approaches in doing this: (i) "Outside-in" tracking, which applies extra tracking devices (e.g. tracking camera) to track the HMD and controllers, thus by far it's the best and most accurate way, an exemplary HMD that uses this technique is HTC Vive[3]; (ii) "Inside-out" tracking, the headset contains sensors tracking the room, by doing so, it can calculate its position. This solution is less accurate, however, there is no demand for external sensors. An example HMD using this technique is Oculus Quest[4].

Technically speaking, for both CAVEs and HMDs, the precision of the sensor, the frame-rate and the resolution of the screen and the FOV are the key parameters that determine the quality of the VR visual experience, cf. (Summers et al., 2015).

2.1.3 Shared Virtual Environments

In the mid-1990s, the development of network technology had made it feasible to link many users simultaneously in the same Virtual Environment (VE), prompting the Shared Virtual Environments (SVEs, see Schroeder 2012). Besides "SVEs", other terms being used include multi-user virtual environments, multi-user virtual reality (Carlsson & Hagsand, 1993), Collaborative Virtual Environments (CVEs) (Zhang & Furnas, 2003) and Social Virtual Reality (SVR). To align with mainstream usage, herein the term SVEs will be used to refer to VE systems in which users experience other participants as being mutually present in the same environment and in which they can interact inter-personally (cf. Schroeder 2012). Whilst single-person VEs may focus on creating a detailed (visual) simulation, the design of SVEs typically prioritises enabling collaboration between users (Nassiri et al., 2010). The emergence of SVEs bring many interesting questions: Whether the interpersonal interaction in VEs is similar to face-to-face interaction in the real world. How the features of avatars might affect the way of interaction.

2.1.4 CSCW and SVEs

To answer the questions above, a better understanding of SVEs is needed. We start by reviewing Computer Supported Cooperative Work (CSCW), as SVEs

[3]HTC Vive: `https://www.vive.com/us/`
[4]Oculus Quest: `https://www.oculus.com/quest/features/`

can be seen as particular applications of CSCW (Sarmiento & Collazos, 2012). The term CSCW was firstly coined in a workshop by Grudin (1994), since then CSCW has remained a hot research field. CSCW research, as a newly evolving area, focuses on providing computational tools to support and facilitate group work (Sarmiento & Collazos, 2012). Three key processes of CSCW have been identified by previous research (Menon 1997; Ray 2002), as re-declared by Sarmiento & Collazos (2012), they are (i) communication between team members; (ii) coordination under a set of clear rules; and (iii) information sharing in a democratic mode. Two principles have been argued to be particularly important for CSCW: (i) cooperation is not a separable activity (like compiling a program or writing a letter that has definite start and end times); (ii) successful groupware system should allow people to cooperate by overcoming barriers of space and time that are imposed on people (Wexelblat, 2014). Wexelblat further used the time-space distinction to categorise CSCW systems according to types of interaction it supports (Wexelblat, 2014). As shown in Figure 2.1, the space axis reflects the degree of spatial distribution (co-located or distributed) whilst the time axis reflects the degree of synchronousness (co-temporal or sequential). Some examples were also given, see Figure 2.1.

The development of immersive Virtual Environments technology has drawn CSCW researchers' attention and interests. SVEs' emergence is a convergence of research interests in VEs and Computing Supported Cooperative Work (Benford et al., 2001). Researchers have proposed the concept "Shared Virtual Environments (SVE) as an alternative for improving those aspects in CSCW systems, cf. (Sarmiento & Collazos 2012).

By providing a natural medium for three-dimensional CSCW, in which people can communicate and interact with each other (Billinghurst et al., 2000), SVEs provide much more affordances for natural interaction, and might improve in several aspects that CSCW systems have shortage in, e.g. the immersive environments and (full body) avatar provided in SVEs make it possible for team members to use tacit and non-verbal information (like visual appearance, facial expressions) to communicate more naturally, to coordinate naturally by having the sense of proximity and sharing information/objects more naturally through the avatars (Benford et al., 1994).

2.1.5 Design Process of SVEs

SVEs' advantages have been verified and exposed by Churchill & Snowdon (1998), however, these advantages also imply that SVEs are unique from CSCWs and could have specific restrictions on both collaborative work and technological requirements of SVEs (Sarmiento & Collazos, 2012). As such, guidelines

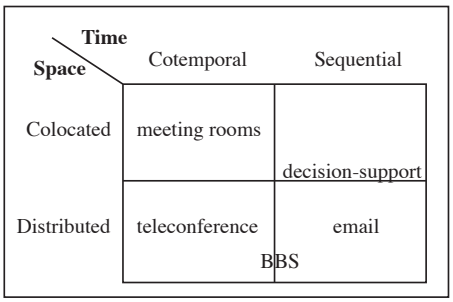


Figure 2.1: Types of groupware (from Wexelblat 2014).

are needed to inform the understanding and application of SVEs. Sarmiento & Collazos (2012) proposed an iterative process to develop CSCW systems in VEs, i.e. SVEs, the process is composed of 7 steps: (i) Software requirements - Eliciting, specifying and analysing the users' needs in a formal and standardised way (Noguera et al., 2010; Rabin, 2005); (ii) 3D design - The design 3D VEs should generate empathy with the users. Although this step is conceptual, it should include scenarios, avatars, shared artefacts and interaction modes and to be results of a formal user observation process (Sarmiento & Collazos, 2012). (iii) collaboration model - a formal schema that abstracts the key process including communication, coordination and sharing (Sarmiento & Collazos, 2012); (iv) visualisation and interaction model - specifying how the users will interact in that VE; (v) software development - implementation of the application; (vi) collaboration monitoring - observing and documenting users behaviour while they are interacting with the system; (vii) awareness monitoring - monitoring self-awareness, awareness of other people and the work group. Note the process should start and finalise with "user observation" process, which is the core for building any systems that involve user interaction.

2.1.6 Applications of SVEs

SVEs are considered emerging tools for a variety of purposes, including system, community activities (Lea et al., 1997), online education (Roussos et al., 1997), teleconferencing distributed work and training (Nedel et al., 2016), gaming and entertainment (check Toybox[5], and PlayStation VR social demo[6]).

Despite this, relatively sparse research exists in the field of supporting collaborative creativity, leaving many open questions.

[5]Toybox: `https://www.youtube.com/watch?v=iFEMiyGMa58` (Accessed: 2019-07-26)

[6]PlayStation VR social demo: `https://www.theverge.com/2016/3/16/11246334/playstation-virtual-reality-social-vr-demo` (Accessed: 2019-07-26)

2.2 Key Features of VEs

Some key features are essential to the success of VEs. The VE technology, according to how the term was defined, is all about simulating the sense of "being" there. Therefore the sense of presence is a key point for VEs. Presence is important to SVEs as well, as SVEs need to immerse people in the VE to make the illusion that they are inside the same world. Measuring presence is also essential to maintain a proper level of that sense. As mentioned above, VEs and SVEs provide the many possibility for new forms of interactions and communication, the embodiment functions as a substitute of users' real body, demonstrating their visual appearance, facial expressions and physic interactions not only to the users themselves, but more importantly, to other users in the same VE (cf. Benford et al. 1994). Therefore supporting embodiment is crucial not only to systems of VEs but also to SVEs. Another important feature is identification, which is especially crucial in SVEs. This is because collaborators would need to know who is performing the action. VEs and SVEs also have other important features, such as 3D design, collaboration model, visualisation and interaction model (check Sarmiento & Collazos 2012), most of these features are very dependent on the types of collaborative task and the two fundamental features: presence and embodiment. Therefore, this subsection will only focus on reporting related work on presence and its measurement, and embodiment in VEs. Next, related work of these two key features of VEs, which are also applicable to SVEs, will be reviewed in detail.

2.2.1 Presence

In 1980, Minsky coined the term "telepresence" to describe the feeling of being at a distant place that a user may have while interacting via a teleoperator system. Presence is the subjective experience of being in an environment or a place, regardless of one's physical environment (Witmer & Singer, 1998). A closely related phenomenon is "distal attribution" or "externalisation", which refer to the referencing of our perceptual experience to "external space beyond the limits of the sensory organs themselves" (Loomis, 1992). Presence is the people's experience in the physical environment surrounding them. It does not mean the surrounding environment but means the people's perception of those surroundings as mediated by both controlled and automatic psychological processes (Gibson, 2014). Many perceptual components help to produce this sense, including input from some or all sensory channels, as well as more mindful attentional, perceptual, and other mental processes that assimilate incoming sensory data with current concerns and past experiences (Gibson, 1966).

Presence was argued to be a result of virtualisation, which was defined by

Ellis (1991) as "the process by which a human viewer interprets a patterned sensory impression to be an extended object in an environment other than that in which it physically exists" (Ellis, 1991). Ellis further pointed out that virtualisation includes three levels: (i) Virtual space - "the process by which a viewer perceives a three-dimensional layout of objects in space when viewing a flat surface presenting the pictorial cues to space"; (ii) virtual image - "the perception of an object in depth in which accommodative, vergence andoption-allystereoscopic disparity cues are present"; (iii) virtual environment - adding "observer-slaved motion parallax, depth-of-focus variation and wide field-of-view without a prominent frame". Although currently, most VE systems focus on simulating visual and auditory senses, an ideal VE should be able to supply the totality of sensory information to the user continually (Usoh et al., 1996). The information displayed by these sensory modalities continually update according to users' movements and interactions, giving the users that they are present in the environment simulated by the computer, i.e. the sense of presence.

Presence is defined as the sense of being in an environment (Gibson, 2014). When applied to virtual environments, presence refers to the experience of the virtual environment rather than the real physical surroundings (Witmer & Singer, 1998). The importance of immersion and presence's role in VEs can be seen from the aforementioned fact that many researchers defined VEs with a focus on the sense of presence or immersion (e.g. Steuer 1992; Sherman & Craig 2018). Previous research has suggested that presence is important to a virtual environment (Sheridan, 1992; Barfield et al., 1995; Slater et al., 2009). For example, a highly present individual is more likely to behave in the VE in a manner similar to their behaviour in similar circumstances in everyday reality (Usoh et al., 1996).

2.2.2 Measuring Presence in VEs

Since presence is a variable reflecting user's feelings, a sense of being in a real or virtual world rather than the ability of VR technology to immerse a user, this makes it difficult to describe presence objectively. To date, there are two common ways of measuring presence. One is subjective measures, requiring the subjects' introspection, the other way is objective measure (Schuemie et al., 2001). Objective measures include behavioural and physiological ways (Schuemie et al., 2001) and subjective measures include continuous measure (IJsselsteijn et al., 2000), presence counter (Slater & Steed, 2000), focus group (Slater & Steed, 2000; Heeter, 1992) and questionnaire (Schuemie et al., 2001). There seem to be a number of ways, however, in presence research, most measures are based on subjective ratings through questionnaires (Schuemie et al., 2001). Question-

naires are usually used to validate the objective measures, a reason for this is the theory of presence is still being developed and thus the rich feedback from questionnaires are essential to promote the understanding of the phenomenon being measured (Schuemie et al., 2001).

In the field of questionnaire-based measures, Witmer & Singer (1998) thought any measure of presence should be both reliable (i.e. only dependent on the characteristics under consideration) and valid (i.e. measuring what it is designed to measure and measuring it well). During an experiment in a VE, questionnaires can not only measure subjective sensations but also ask the subjects to describe other things (e.g. the VE and their behavioural and physiological responses), in which the investigators may be interested (Schuemie et al., 2001). The observations through questionnaires are of course less reliable as a result of its subjective nature (Schuemie et al., 2001). Some questionnaires consist of a single question to measure presence, for instance, "I feel a sense of actually being in the same room with others when I am connected to a MOO" (Towell & Towell, 1997). To improve the reliance of the measurement, several questions are needed (Dinh et al., 1999; Hendrix & Barfield, 1996a). Recently, several questionnaires are made in a more or less systematic way (Schuemie et al., 2001). Next, we introduce several predominant ones. Based on previous studies (Usoh et al., 2000), the questionnaire proposed by Slater et al. (1995) consists of several questions. All are variations on three themes: the subjects' sense of "being there"; the extent to which the VE becomes more "real or present" than everyday reality; the "locality", the extent to which the VE is thought of as a "place" that was visited rather than just a set of images. Witmer & Singer (1998), in their Presence Questionnaire (PQ), determined four factors contribute to presence, the presence score is the sum of the users' ratings. These four factors are: (i) *control factors* - the amount of control the user had on events in the VE; (ii) *sensory factors* - the quality, number and consistency of displays; (iii) *distraction factors* - the degree of distraction by objects and events in the real world; (iv) *realism factors* - the degree of realism of the portrayed VE. Witmer & Singer (1998) further reduced questions believed to reduce the reliability, and three factors were found: (i) *involved/control* - the control and responsiveness of a VE, and how involving a VE is; (ii) *natural* - the naturalness of interactions and control of locomotion, and the consistency of a VE; (iii) *interface quality* - the amount of interference or distraction from task performance. Igroup Presence Questionnaire (IPQ) is also a popular questionnaire in this field. By combining earlier published questionnaires (these include parts of Witmer and Singer and Slater and colleagues' work). Schubert et al. (1999) built their IPQ. With testing the resulting questionnaire to 246 volunteers, three presence factors were found: spatial presence (SP), having a sense of physical presence in VE; involvement

(INV), measuring the involvement experienced and the attention paid to the VE; realness (REAL), measuring the subjective experience of realism in the VE.

2.2.3 Embodiment in VEs

In the real world, human beings are using face-to-face interaction in their everyday lives. For example, by observing facial expressions, we could recognise others' emotions and feelings. By viewing the appearance and dressing, we could even discover more details (e.g. the origin and job) of a man or woman. We might then respond differently based on the data we got and the conclusion we made. Benford et al. (1995) have indicated that it is our bodies that provide immediate and continuous information about our presence, activity, attention, availability, mood, status, location, identity, capabilities and many other factors to ourselves and others, and suggested that people are using their bodies (so-called body language) explicitly in communication (Benford et al., 1995). Social interactions in virtual worlds and the real world are regulated by the same social norms (Yee et al., 2007), so user embodiment is an important feature to technology-mediated collaboration as well. Researchers have raised questions ranging from how the embodiment could affect users' social communication and behaviour (cf. Wallace & Maryott 2009), the impact of appearances and behaviours on people's sense of presence or immersion (Draper et al., 1998; Minsky, 1980; Sheridan, 1992; Steuer, 1992) and social presence (i.e. co-presence, check Rice 1993; Short et al. 1976). Findings indicate that the embodiment conveys presence, location and identity (Benford et al., 1995, 2001), all of which are important to a successful collaboration (cf. Ellemers & Rink 2005; Bryan-Kinns & Hamilton 2012). As such, choices about embodiment are crucial and can affect the quality and extent of collaboration in VR (Wallace & Maryott, 2009).

An appropriate use of embodiment can increase the sense of telepresence (Nowak & Biocca, 2003), the sense of social presence–the feeling that others are present with the user in the mediated environment (Benford et al., 1995; Nowak & Biocca, 2003), and promote the sense of community (Rovai, 2002). Gutwin & Greenberg (2004) coined the term "workspace awareness" as *"the up-to-the-moment understanding of another person's interaction with the shared workspace"* and addressed that it is essential for groupware systems to give team members a sense of workspace awareness. They listed all the elements of workspace awareness, see Table 2.1. In SVEs, the embodiment might have a great potential in providing support for these elements, e.g. a proper and sufficient embodiment could help people see not only the changes made by each other's but also the actions of how they made the change, which is extremely

Table 2.1: "Elements of workspace awareness" from (Gutwin & Greenberg, 2004).

Category	Elements	Specific Questions
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?
	Artefact	What object are they working on?
Where	Location	Where are they working?
	Gaze	Where are they looking?
	View	How much can they see?
	Reach	How far can they reach?

useful for our understanding of the co-work (cf. Benford et al. 1995), which cover all the three categories (who, what, and where) of the worskspace awareness proposed by (Gutwin & Greenberg, 2004), see Table 2.1.

The embodiment can also mould a strong sense of identification, which is important in collaboration because it is a key element for constructing workspace awareness (Gutwin & Greenberg 2004, see Table 2.1) and it can affect collaboration positively as well as negatively in group works (Ellemers & Rink, 2005). Awareness of the identity of others can significantly increase mutual engaging interaction (Bryan-Kinns & Hamilton, 2012). And in VEs, to a large extent, the identification is moulded by the embodiment.

In many traditional CSCW (computer-supported cooperative work) systems, users become known to one another through their (disembodied) actions. This is due to insufficient embodiment (Benford et al., 1995), i.e. we know someone is working with us by seeing the changes they made. For example, in Google Doc, the co-worker adds a word, which is seen by other coworkers immediately. High-level sufficient embodiment could help coworkers see not only the result of each other's behaviour but also the actions of how they made the changes, which is extremely useful for the understanding of the co-work because now coworkers know where the changes come from. A better sense of coworker's actions is also helpful to the construction of workspace awareness (Gutwin & Greenberg 2004, see Table 2.1). Technological advances have increasingly enabled computer-generated entities to mimic both the appearance and behaviours of humans (Brent & Thompson, 1999; Dryer, 1999). For example, within the prototypes developed for the experiment discussed in this thesis, the Light-house tracking cameras of HTC Vives[7] and hand tracker Leap Motion[8] are used to

[7] `https://www.vive.com/eu/accessory/base-station/`
[8] `https://www.leapmotion.com`

provide a real-time capture of the movements of heads and hands and gestures. Another example is Veeso, which is a headset with built-in face tracking technology. Besides tracking, contemporary technology also enables us to rebuilt 3D embodiment vividly. The avatar can be as simple as a T-shape with eyes to show orientation and direction of view, or as complex as a full 3D body scan of the user (Sherman & Craig, 2018). In Section 3.5.2, the design and development of the avatars applied in the three studies of the thesis will be specified.

2.3 Creativity, Music Making and VEs

From primitive to civilisation, from every small step forward on earth to the "giant leap" Armstrong did on moon, every progress throughout human history, no matter great or tiny, will never be made without creativity. At both the individual and societal levels, creativity is critical to invention, innovation, and social progress (Sternberg & Lubart, 1999; Candy & Hori, 2003). Creativity defines human and makes human human, yet it stays mysterious (Sawyer, 2011).

Given its importance, creativity has traditionally been explored across a vast range of activities and domains for a long period of time (Kaufman & Sternberg, 2010; Shiu, 2014; Sternberg, 1988). Lubart (2005) defines creativity as the ability to produce work that is novel and appropriate. Boden (2003) defines creativity as the the ability to form new, surprising and valuable ideas (including concepts, poems, musical compositions, scientific theories, cookery recipes, choreography, jokes etc.) and artefacts (e.g. drawings, sculptures, steam engines, vacuum cleaners, pottery, origami, penny whistles). The report published by UK National Advisory Committee on Creative and Cultural Education in National Advisory Committee on Creative and Cultural Education (NACCCE) (1999) tries to define creativity by depicting its features: (i) Using imagination - during the process of generating originality; (ii) pursuing purposes - having the imagination applied; (iii) being original - originality can be in relation to a person's own previous work, to their peer group or anyone's previous output in a particular field; **iv**) judging value - valuation in respect to the objective applied, creativity involves both idea generation and evaluation. Defining creativity can be "illusive" (Ford & Harris, 1992), C. W. Taylor (1988) even traced "some 50 or 60 definitions" of creativity.

To ease the process of understating and researching creativity, researchers try to classify creativity into different sub-types. Based on the level of novelty, these ideas or artefacts may be novel to the individual who created it, or to the whole of previous human history. The abilities to create novelties of the former kind is defined as Psychological-creativity (P-creativity for short), the latter is defined as Historical-creativity (H-creativity for short), and H-creativity can be seen

as a special case of P-creativity, which is a more fundamental notion (Boden, 1998). A similar way is based on the scope of the creativity and the experience of the creators. "Big-c" creativity was used to refer to clear-cut, genius level of creativity, where creators usually have rich experience, and "little-C" creativity refers to everyday creativity that may make a contribution and "mini-c" is more about interpersonal creativity during the learning process, both require much less experience of the creator (Plucker & Beghetto, 2003; Beghetto & Kaufman, 2007; Kaufman & Beghetto, 2009). Creativity can also be grouped based on the ways of generating novelty, three types were proposed by Boden (1998): (i) combination creativity, which is combinations of familiar ideas; (ii) exploratory creativity, which is exploring structure conceptual spaces; and (iii) transformational creativity-which is generating new structures by transforming some dimension of the space.

VEs, with great potential to support or even enhance creativity, are perhaps an underused tool (Thornhill-Miller & Dupont, 2016). Thornhill-Miller & Dupont (2016) argued VE technologies can perhaps be applied to enhance creativity and problem solving (i) by changing aspects of self and self-perception; (ii) by optimising interactions and collaboration; (iii) by optimising environmental conditions and influences; (iv) facilitating guidance or gamification of the problem-solving process; and (v) by offering an arena for the integration of other technologies of creativity enhancement such as pharmacological enhancement, brain stimulation, and neurofeedback. A good example is Google Tilt Brush[9], which enables people to draw in 3D space with variety of virtual brushes, see artful works Artist in Residence[10]. Despite VE's great potential in supporting creativity, VE is still barely researched in this field. How to apply VE technology in supporting creativity still remains an open question.

2.3.1 Collaborative Music Making Systems

Music making, as a collaborative activity that relies on common goals, understanding and good interpersonal communication, has long been a key form of collaborative creativity (cf. Titon & Slobin 1996; Bryan-Kinns & Hamilton 2012). The ability to create shared social experience is one of the fundamental aspects of music making (Serafin et al., 2016). These unique features of music making make it an excellent activity for research on exploration and understanding of collaborative creative interactions in SVEs.

In 2003, Blaine & Fels explored the design criteria of Collaborative Music Making systems (CMMs), pointing out key features including the media

[9]Google Tilt Brush: `https://www.tiltbrush.com`

[10]Artworks created by various artists, painters, cartoonists, dancers, designers at: `https://www.tiltbrush.com/air/`

used, player interaction, the systems' learning curves, physical interfaces and so on. To classify CMM systems as they emerged in the ensuing years, Barbosa (2003) developed the Networked Music Systems Classification Space inspired by the Classification Space for groupware proposed by Rodden (1991), to distinguish systems in terms of the time dimension (synchronous/asynchronous) and space dimension (remote/co-located). For example, Daisyphone in Bryan-Kinns (2004), which provides a shared loop that up to 10 people can edit remotely and any changes are updated in real time, falls into the remote synchronous network music systems in Barbosa's Classification Space. Other examples include reacTable (Xambó et al., 2013) and BilliArT (Bressan et al., 2017), both of which provide co-located shared musical experiences with tangible interfaces, and Ocarina (Wang, 2009), which provides a distributed music-making experience. Though many CMMs have been developed, according to Wozniewski et al. (2008), most of them rely on users to be in a relatively fixed position (e.g. in front of a computer). The head tracking and spatialised audio offered by VE technology can be good tools to break this chain and to free the users. However there have been relatively few SVEs which support CMM, making this research area barely explored, especially in terms of the collaborative aspect.

2.3.2 Annotations in Groupware

Although challenging, supporting annotation of artefacts is widely considered necessary in the collaborative process (Phalip et al., 2009). Several solutions for annotating text-based documents are already in widespread use (e.g. Microsoft Office[11], Google Docs[12] & Adobe Acrobat[13]). Beside texts, annotating other type of documents or data is also necessary. For example, Scribblr is a tool to annotate drawings and image sketches online (Weakley et al., 2007), and Viddler[14] is an online tool to enable users to post textual or video-recorded comments at particular points in the timeline of a video. Annotation compelling systems have also been built to annotate multimedia documents and video material (Bouvin et al., 2002; Ramos & Balakrishnan, 2003). Likewise, for collaborative interactive systems, writing and sketching can possibly be used to exchange ideas, act as a memory aid, convey approvals, ideas, doubts and so on. For example, the CMM system Daisyphone and Daisyfield in Bryan-Kinns (2011) provide people with a shared annotation mechanism, and it has been suggested that music making may benefit from this.

[11] `http://www.microsoft.com/office/`
[12] `http://docs.google.com/`
[13] `http://www.adobe.com/products/acrobat/`
[14] `http://www.viddler.com/`

2.3.3 Music Making in VEs

To date, most virtual reality research focuses on exploring the application of visual feedback, with other senses (e.g. auditory feedback, tactile feedback) playing a secondary role (Serafin et al., 2016). Though this sole focus might be capable enough to provide a wonderful and convincing experience in some cases, it brings a ceiling effect, which can only be broken by bringing the other senses together. This section reports related work of Virtual Reality Musical Instruments (VRMIs), which apply not only the visual channel, but also the audio channel, and possibly other sensational channels as well. The integration provides rich feedback for user and enhancing the user experience.

In the areas of new interfaces for musical expression (NIME) and sound and music computing, virtual musical instruments (VMIs) have been developed and refined in recent decades (see Cook 2002; Smith 2006). VMIs are software simulations or extensions of existing musical instruments with a focus on sonic emulation, for example, by physical modelling synthesis (Välimäki & Takala, 1996). Considering the visual feedback is closely related to the sound production mechanism in the natural world, e.g. vibrating strings, providing visual feedback becomes essential for VMIs (cf. Serafin et al. 2016). Whilst VMIs focus on providing auditory and even tactile feedback, A similar competing term "Virtual Reality Musical Instrument" (VRMIs) was coined, with immersive head mounted displays (or other forms of immersive visualisation systems) applied, providing rich the visual feedback, cf. (Serafin et al., 2016). VRMI was refined to indicate representing sound processes and their parameters as 3D entities of a virtual reality so that they can be perceived not only through auditory feedback but also visually (Bowman et al., 2010).

Different from VMIs, which have been researched for several decades, VRMIs have not drawn much attention. Cadoz et al. (1993) developed CORDIS-ANIMA, a system that does not use immersive technology but includes all the elements of immersive multimodal musical instruments, allowing user to create physically plausible and rich computer-generated sound images and movements. The additional visual channel of VRMI might provide stimulated intelligent feedback, aiding music players in performing, and improving the interaction between players and the instruments. Mäki-Patola et al. (2005) explored interaction metaphors based on existing musical instruments as seen in his Virtual Xylophone, Virtual Membrane, and Virtual Air Guitar implementations. Most of the instruments were created in a CAVE-like virtual room. Take the Virtual Air Guitar for example, it uses data gloves and motion tracker to detect plucks, vibrato, slide and mute, enabling users to play electric guitar sounds by performing gestures in the air. The authors further suggested that exist-

ing technologies could be used to produce more responsive instruments, and expand design and performance possibilities. Gelineck et al. (2005) developed several physics-based VRMIs, users were provided with drum-like or flute-like controllers, rotation and position of which were tracked and mapped the rotation and position of virtual 3D instrument models. The authors further discussed the possibilities not achievable in the physical counterpart, such as the ability to change dimensions of the simulated instruments whilst playing them, or providing additional visual feedback to immerse the users and augment the experience, e.g. visualising changes of sounds in amplitude and frequency altering lights and particles. In 2010, Berthaut et al. created Drile, a multiprocess immersive instrument built for musical performance. The technique "hierarchical live-looping" technique applied in Drile allows musicians to create and manipulate complex musical structures. Different from common CAVE systems which have 3 to 6 screens, Drile only includes one large screen, users need to wear 6-DOF-tracked stereoscopic glasses and hold tracked controllers in front of the large screen to interact with the system.

With the commercialisation and resurgence of VR technology, more Virtual Reality Musical Instruments have been developed as commercial products. Next, we list a few of them. *The Music Room*[15] is a collection of instruments that allow players to strum, slide, bend and drum in virtual reality. It is also a MIDI controller, meaning users can loop and record music, select presets and launch clips in VR. *Soundstage VR*[16,17] is a virtual reality music sandbox built specifically for room-scale VR (e.g. HTC Vives), it provides a diverse musical tool-sets (e.g. sequencer, keyboard, horn and drum-kit) for both musical professionals and musical hobbyists. Apart from the interactive VRMIs, Soundstage VR also includes a modular mix chain with a library of effects and processing. It also provides a looping and recording stage, which can be used for post-production or other media productions. *Music Inside*[18] is a multiplayer music rhythm game, in which players play with virtual drum-kits to match the songs. Different from most music rhythm games, beats of which have been previously determined, *Music Inside* features a self-developed algorithm system, which analyses the sound source in real-time, allowing users to play any songs. Similarly, *DrumBeats VR*[19] is also a VR music rhythm game, players can import custom songs into the game and drum along with the songs that are custom made for this game. *Electronauts VR*[20] is a VR multiplayer music

[15] `http://www.musicroomvr.com`
[16] `http://www.soundstagevr.com`
[17] `https://xinreality.com/wiki/SoundStage`
[18] `https://www.realityreflection.com/musicinside.html`
[19] `https://store.steampowered.com/app/1015480/DrumBeats_VR/`
[20] `https://survios.com/electronauts/`

creation software, which sits at an interesting intersection somewhere between game, experience, and tool. Players can make music solo or with friends in VR. The Music Reality Engine used in *Electronauts* makes it easier to make good tune. Different from the VRMIs briefed above, *Electronauts* provides full-body avatars wearing spacesuit, and players have the choice to customise the avatars. *EXA the infinite instrument VR*[21] is an immersive musical studio that allows players to compose, record, and perform music using expressive instruments. It supports multiplayer, so players can create virtual bands. Players can also jam with sound library embed in the system or their sound. Musical outputs can be exported to WAV/MIDI for further usage or editions.

2.3.4 Design Principles for VMIs

As VRMIs were developed from NIMEs and inherited many features from NIMEs, we will review research on design principles for NIMEs first and then go into detail about design principles for VRMIs. In 2001 based on experience in designing and constructing musical interfaces, Cook proposed a loose philosophy and a set of design principles that covered both human and technological factors. These principles were reinforced by observations on the design, artistic and human factors of digital music controllers. In 2009, Cook revised and expanded the principles, pointing out the principles changed due to technological progress. Though a decade has passed, a few of these principles are still closely relevant to VRMIs, e.g. "copying an instrument is dumb, leveraging expert technique is smart". I.e. instead of replacing or copying existing instruments, it's better to consider controllers inspired by virtuosity (Serafin et al., 2016). In 2008, to guide the development of VMIs built for research purposes, Johnston et al. summarised a set of design criteria that aligned with the criteria proposed in Wessel & Wright (2002) and Fels et al. (2002). With inspiration from the principles of (Cook, 2001, 2009), in 2014, Wang presented a total of 13 principles for visual design of computer music. These principles categorised into "user-oriented", "aesthetic", and "other", with a specific focus on real-time integration of graphics and audio. Many of Wang's principles are also relevant for VRMIs if extended as explained appropriately (Serafin et al., 2016). In 2016, Serafin et al. proposed 9 design principles for virtual reality musical instruments (VRMIs), with a focus on the experience of the performer while interacting with the virtual world. In 2017, Morreale & McPherson proposed a set of design considerations accounting for longevity of Digital Musical Instruments (DMIs). Such considerations were based on a survey conducted on designers of DMIs asked to reflect on design issues that limited the uptake of

[21] `https://store.steampowered.com/app/606920/EXA_The_Infinite_Instrument/`

the crafted instruments or the aspects that facilitated their establishment and adoption. In 2019, whilst defining Smart Music Instruments (SMIs), Turchet meticulously reviewed the research on design principles of NIMEs, pointing out these principles overlooked the possibilities that enabling co-located and remote musical interactions of the player with other players, audience members, and machines through interconnecting of the instrument with other devices.

2.3.5 Designing Principles for VRMIs

As the principles proposed by Serafin et al. (2016) are closely related to the software built for this research, next they will be reviewed in detail.

Principle 1: Design for Feedback and Mapping. Serafin et al. (2016) argued all modalities (sound, visual, touch, and proprioception) should be designed in tandem, thus mapping between these modalities should be carefully considered. 3D sound is an especially relevant issue for VRMIs, the location and motion of the auditory virtual objects should match the location and motion of the visual virtual objects (Begault & Trejo, 2000).

Principle 2: Reduce Latency. Similar to Wang's focus on real-time (Wang, 2014), Serafin et al. (2016) argued that ideally, all interactions should be smooth and minimised latency was preferred, and it was important for VRMIs to provide timely audiovisual feedback according to users' actions. The perceptual binding that occurs in response to an event producing multimodal stimulation is influenced by the synchronisation between the arrival of stimuli in different modalities (Kohlrausch & van de Par, 1999). According to LaViola Jr (2000), the system's latency can lead to cybersickness.

Principle 3: Prevent Cybersickness. Cybersickness, or VR sickness (Fernandes & Feiner, 2016), may involve several different symptoms, including but not limited to, disorientation, headaches, sweating, eye strain, and nausea (Davis et al., 2014). The oldest, most accepted and most predominant theory relating to motion sickness and cybersickness is sensory conflict theory (Reason & Brand, 1975), i.e. the discontinuity between either visual, somatosensory, and proprioceptive input results in sickness symptoms. Serafin et al. (2016) argued that apart from optimising technological factors (e.g. efficient tracking, high update and frame rates), creators of VRMIs should also be mindful of the way users navigate in VEs, using a one-to-one mapping between virtual-world and real-world translations and rotations was strongly suggested.

Principle 4: Make Use of Existing Skills. Serafin et al. (2016) advised using metaphors derived from interactions existing in the real world to help users grasp the concept at the beginning. Following Cook's notion (Cook, 2001), a simple copy of existing musical instruments in the real world is less interesting

and should be deprecated. Instead, new kinds of interfaces that are bested suited for VR medium should be discovered to extend existing experience with musical instruments (Serafin et al., 2016).

Principle 5: Consider Both Natural and "Magical" Interaction. As defined by Serafin et al. (2016), interactions and instruments qualify as magical if they are not limited by real-world constraints (e.g. the laws of physics, human anatomy, or the current state of technological development). Otherwise, they qualify as natural if they conform these constraints. Serafin et al. (2016) argued designers and developers of VRMIs should consider using both types of interactions and instruments, this is also in line with recommendations by Bowman et al. (2010) for 3D user interfaces.

Principle 6: Consider Display Ergonomics. Hiding technology allows users to focus on the experience itself. However, to the present, the HMDs are still far from perfect. Though wireless technology and mobile HMDs have shaken off the wires, the weight of the HMDs is still non-negligible, not to mention the filed of view (FOV) provided by most HMDs is still far narrower than the perceptive FOV by eyes in the real world. And note that the performance for search tasks, comparison task, and walking tasks is positively related with increased FOV (Arthur & Brooks Jr, 2000; Ball & North, 2005). Therefore, when creating VRMIs, designers should keep those strain and potential discomforts in mind (Serafin et al., 2016).

Principle 7: Create a Sense of Presence. Presence is one of the major factors that affects VR experience. Whilst creating VRMIs, the limitations of the system should be borne in mind and users should be discouraged from relying on sensorimotor contingencies not fully supported by the system (Serafin et al., 2016).

Principle 8: Represent the Player's Body. Though ownership for one's body and its parts as well as their localisation in space play an important role in constructing bodily self-consciousness (De Vignemont, 2007; Gallagher, 2000; Tsakiris et al., 2007), users cannot see their real body anymore whilst immersed in VR. However, a virtual representation can be provided by tracking and mapping the real body. The virtual body results a sensation called virtual body ownership (Serafin et al., 2016). Seeing the virtual body from the first-person perspective (i.e. the virtual body substitutes the real body) can generate a perceptual illusion of ownership over the virtual body (Gallagher, 2000).

Principle 9: Make the Experience Social. To present, VR has merely been an individual activity due to the occlusive HMDs, which block communication with the outside (Serafin et al., 2016). However, VR might have the potential in facilitating social interaction between individuals virtually (Nowak & Biocca, 2003). Additionally, music is a good activity to create shared social experience

(Serafin et al. 2016; Bryan-Kinns & Hamilton 2012; Titon & Slobin 1996), so shared and social VRMI experiences should be encouraged.

2.4 Space, Territory and Privacy

VEs constitute illusive but meaningful virtual spaces (Steuer, 1992). This virtual space provides richness and affordance for people's interactivity in VEs. Therefore gaining a better understanding of the space is an effective way to understand the collaboration in VEs. "Space" is a material given prior to the happening of actions, and territory emerges as a result of the actions and a production of the actors (Raffestin, 2012). Territory helps people to mediate their social interaction (Altman, 1975), which is argued to be a key element to collaboration (Kreijns et al., 2003). Additionally, people were found to perform creative collaboration in a similar way with the real world, they divided the working space and formed territory (Men et al., 2017). Hence, potentially, with more knowledge of the virtual space, we can even manipulate the virtual space to influence the collaboration in SVEs. Note that the term "space" here specifically refers to the dimensional physical/virtual space rather than the space in psychology or social science, which falls out of the scope of this thesis.

2.4.1 Personal and Group Space in Collaboration

A "personal space" herein refers to a specific space assigned to a specific person and "group space" refers to a specific space assigned to a specific group prior to the start of activities (e.g. an experiment). For CSCWs that focus on supporting collaborative creativity, providing personal space is argued to be useful (Scott et al., 2004). Integrating personal and group spaces, allowing users to work individually in their personal spaces at their pace, cooperatively work together in the shared space, and smoothly transition between both of the spaces are important (Greenberg et al., 1999; Sugimoto et al., 2004). As a starting point of this exploration, Greenberg et al. (1999) developed a PDA-based prototype, observed how users shifted between the two spaces and recommended against a rigid notion of "personal". Instead, they suggested the boundary between personal and public should be provided with gradations in subtle and lightweight ways, supporting a fluid transition between personal and public. Following that, Shen et al. (2003) addressed this concern in their project UbiTable by providing a flexible gradient of sharing semantics. Specifically, rather than the binary notion of public and private space, UbiTable provides an additional semi-private space, in which data is visible but not electronically accessible for others. However, both sets of research were carried out based on 2D media (PDA and

projector), which made their findings less applicable for workspace design in VEs.

2.4.2 Territory in SVEs and Tabletop Collaboration

Human territoriality is a powerful and pervasive element in human being's lives (Sack, 1983). R. B. Taylor (1988) argues that it is an *"interlocking system of attitudes, sentiments, and behaviours that are specific to a particular, usually delimited, site or location"*, which reflect and reinforce, for individuals or a small group *"some degree of excludability of use, responsibility for, and control over activities in these specific sites"*. Similarly, Sack (1983) sees it as a basis of power, a spatial strategy to affect, influence, or control resources and people. By claiming a space, territory helps people mediate their social interaction (Altman, 1975), which is a key element to collaboration (Kreijns et al., 2003).

Because there is limited research on territoriality in VR, and rich research on this in tabletop-based collaboration, we review territoriality in tabletop research as a supplement, which might be informative as it is also a computer-mediated collaboration with territory involved. The term "tabletop" here refers to interactive *table top displays*, which usually include high-quality projectors, flat panel and plasma displays, and touch-sensitive surfaces (Kruger et al., 2004). These electronic tabletops inherit the collaborative benefits of tables, which greatly compensate computers' disadvantages in this regard (Scott et al., 2004). Similar to the crucial role territoriality plays in human being's lives, it also constitutes crucial aspects of the tabletop collaboration. Collaborators were found to use different types of territory to serve different needs, including sharing, exchanging or storing working tools and resources (Scott et al., 2004), though some researchers found that removing territorial constraints can promote exploratory group activity (see Xambó et al. 2013). Two main types of territory are identified from research on screen and tabletop mediated collaboration:

(1) *Personal territory* for performing independent activities. This type of territory serves as a safe place to try and develop alternate ideas before publishing the ideas (Tang, 1991). When provided with a personal territory, users prefer to test their contribution before introducing it to the group work (Fencott & Bryan-Kinns, 2010). Users have been found to prefer to rotate items toward themselves in personal territory (Tang, 1991) and perform very few actions in their collaborators' personal territories (Scott et al., 2004).

(2) *Group territory* for performing the main task. In group territory, people create and develop new solutions, transfer resources and provide help (Scott et al., 2004). It is interesting to note that the orientation properties of objects in the group territory can be used to convey support, to separate ideas or to group

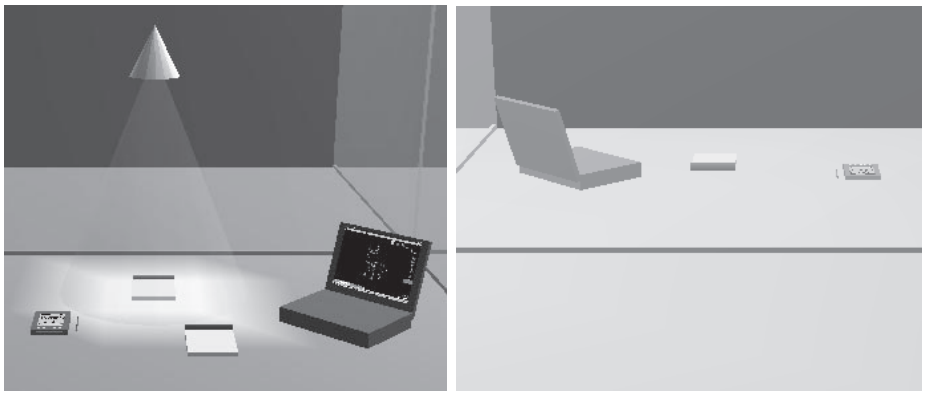


Figure 2.2: (left) A private privacy lamp shining on an object and (right) the view from another cubicle (from Butz et al. 1998).

products (Tang, 1991).

In terms of designing for territoriality, Scott et al. (2004) proposed four guidelines for designing digital tabletop workspaces: (i) visibility of action; (ii) an appropriate size of workspace; (iii) providing functionality in the appropriate locality; **iv**) allowing for the grouping of items to facilitate storage. Furthermore, the visibility and transparency of actions have been found to be important in designing group workspaces, as they help collaborators to monitor each others' actions, maintaining workspace awareness during collaboration (Pinelle et al., 2003; Fencott & Bryan-Kinns, 2010). However, this can result in overloaded cognitive information, which some people found to be difficult to handle (Fencott & Bryan-Kinns, 2010). To date, little research has explored how such features of territoriality might be designed for and used in SVEs.

2.4.3 Privacy in Collaboration

Privacy gives people the ability to seclude themselves and schedule their activities independently. It has long been an important issue in the design of any multi-user system, e.g. the way of visually representing objects' privacy states in SVEs (Butz et al., 1998). One traditional way to differentiate public and private things is simple and direct, which is "visibility = public = accessibility", and "no visibility = private = restricted accessibility". It is true that most desktop CSCW systems assumes that only explicitly shared things are public, whilst all other things on a user's computer screen are private by default, because most items on the screen inherently have nothing to do with the collaboration (Baecker, 1993). However this approach does not fit SVE, because the aim of SVE is to simulate an illusion that coworkers are working in the same

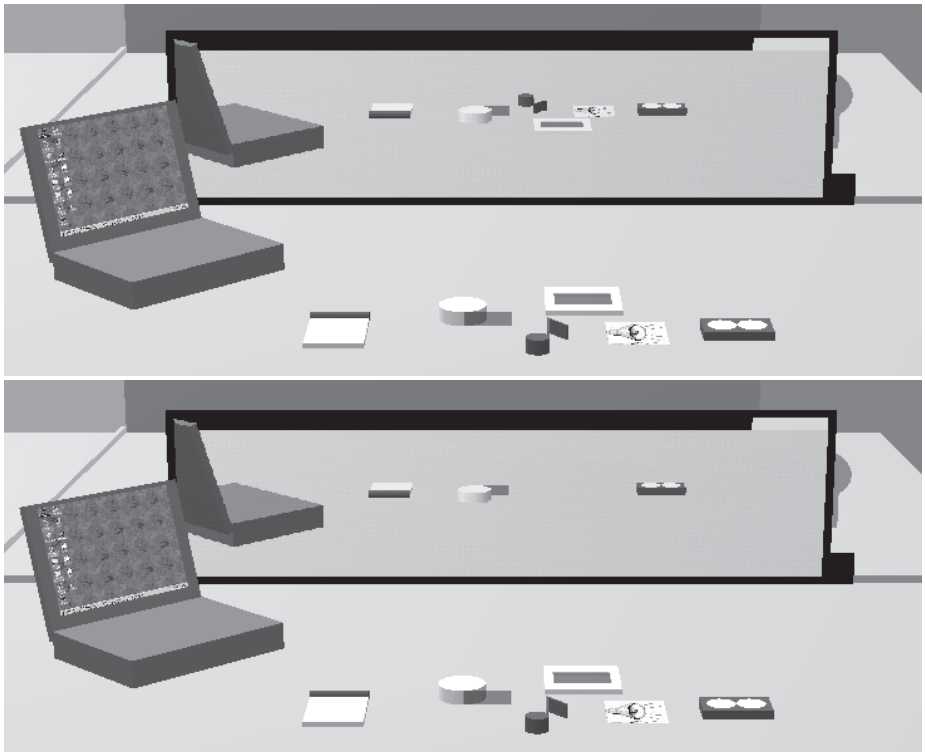


Figure 2.3: A vampire mirror (upper figure) with all objects public and (bottom figure) with selected objects made private (from Butz et al. 1998).

VE. As such, the same set of virtual objects should exist in the VE and the visibility should not be selective. Butz et al. (1998) suggested two methods to manipulate the privacy condition of objects in augmented reality, the first one is *privacy lamps*. If all objects are public by default, a *privacy lamp* can be naturally picked up and positioned over the objects that users want to make it private, see Figure 2.2. Similarly, if all objects are private by default, users can use public lamps to make things public. Another tool is called *vampire mirror*. as shown in Figure 2.3, it is a tool that only selectively reflects what other users can see (i.e. it only reflects the publicly visible objects), so users can review objects' privacy state. However, such solutions might not suit SVEs, as it breaks the illusion that all the collaborators are present in the same VE and manipulating the same set of virtual objects.

2.4.4 Chapter Summary

To summarise this chapter, the first section has reviewed virtual environments, including its definition, example Virtual Environment systems, and Shared Virtual Environments (SVEs). SVEs are believed emerging tools for many purposes, however, the topic is still insufficiently researched, especially in supporting creative activities. The second section discusses the related works on key factors of VEs and SVEs, including presence (and its measurement) and embodiment, providing sense of presence is the core of VEs and embodiment is a very powerful element to enhance the immersion and other aspects of VEs. The third section talks about the relationship between creativity, music making and VEs, pointing out VEs provide a great opportunity to enhance creativity and music making is a nice activity to create shared virtual experiences. Finally, related work about space, territory and territoriality, and privacy and collaboration are discussed. Little research has explored these topics in a SVE context, therefore many research issues still remain unsolved.

Chapter 3

Methodology

This chapter delineates the methodology approach for the research and the technical implementation of the prototype. It starts with a reflective review of the methods of evaluating collaboration in SVEs, followed by a description of the methods applied in the studies of this thesis, and ended with a delineation of the development of the collaborative music system named LeMo, which was developed for the three studies.

3.1 Measuring Collaboration

Measuring collaboration is by no ways easy, evaluating collaboration is at least an order of magnitude more complex than typical human-computer interaction (HCI) evaluations because it involves multiple humans interacting with networked systems (Damianos et al., 1999). Collaboration is argued to be multidimensional (Burkhardt et al., 2009). For example, as pointed out by Burkhardt et al., in numerous application domains (e.g. software design, architectural design), empirical studies on collaboration in design teams (for a state of the art, see Détienne 2006) have highlighted the importance of distinctive collaborative processes for successful design: communication processes such as grounding (Clark & Brennan, 1991), task-related processes (e.g. exchanges of knowledge relevant for the task at hand; argumentation processes), coordination processes, and motivational processes. Seven dimensions of creative collaboration have been given and each of them contains several indicators (Burkhardt et al., 2009). E.g. *Fluidity of collaboration* has three indicators including fluidity of verbal turns, fluidity of tools use and coherency of attention orientation. Similarly, Nguyen & Duval (2014) argued, collaboration in SVEs include three processes: awareness, communication and performing. Each also includes several indicators. E.g. the awareness includes self awareness, awareness of others (others,

and others' location, actions and activities), and awareness of the environment (both the virtual and the real environment).

To evaluate these dimensions of collaboration, there are two types of evaluation methods: (i) *usability inspection methods* which do not rely on the participation of real users (e.g. Pinelle et al. 2003); (ii) *usability evaluation methods* based on users' studies. *Usability inspection methods* are less costly, however, their focus remains on individual-centred task models (e.g. Tromp et al. 2003). I.e. instead of exploring effective collaboration processes in a real context, these methods elicit goals and actions required for users to interact together and not on the collaboration processes and their quality per se (Burkhardt et al., 2009). *Usability evaluation methods* collect both qualitative and quantitative data based on empirical studies, and seem to be a more popular way to study collaboration in CSCW. Various methods and techniques have been developed and applied for data collection and analyses in experimental settings, such as computers logs, interactions between participants (coding methods or ethnomethodological methods), and interviews (Burkhardt et al., 2009). For example, various means – such as videotaping each group, audio recording their discussions and logging chat conversations and player activities during the game – were applied to gather data to evaluate collaboration in a virtual game environment for vocational learning, findings of different types of data were compared and then be cross-verified or refuted accordingly (Hamalainen, 2008). In experimental studies, fine-grained interactions are usually quantified as indicators to assess usability regarding collaboration processes. Two examples in Hornbæk (2006) are: (i) *"use frequency"* - number of keystrokes, number of mouse clicks, number of functions used, number of interface actions, amount of mouse activity, number of times help is consulted; (ii) "*communication effort*" - number of speakers' turns, number of words spoken, number of interruptions, amount of grounding questions. According to Burkhardt et al. (2009), the empirical-based methods' drawbacks include: (i) higher costs and time-consuming; (ii) sometimes even impossible to apply with prototypes; (iii) the extent to which existing empirical-based methods cover all the dimensions of collaboration and iv) their generality or ad hoc nature.

3.2 Measuring Collaboration from Three Aspects

To explore the research question of this thesis – how to support CMM in SVEs, measuring the collaboration is essential. Therefore measures of collaboration need to be developed. Specifically, the measures access collaboration from three aspects: awareness, communication and performance, echoing the argument in Nguyen & Duval (2014) – collaboration in SVEs includes three processes:

awareness, communication and performing.

Awareness - The awareness in SVE was defined as obtaining a complete knowledge of the environment within which an individual is working and of other people they are working with, it provides a context for one's own activity(Nguyen & Duval, 2014). Here, we see presence as self-awareness and extend the scope of awareness to cover this. As such, awareness includes both self-awareness, awareness of others and awareness of other's activities)(Holmquist et al., 1999). The former is the fundamental concept and a key contributor to VR (Steuer, 1992), whilst the latter two are commonly, collectively called group awareness (Cruz et al., 2014). Group awareness concerns the knowledge of those they is working with, in other words, the sense of connection between people in VE, A reinforced group awareness was found to be closely related to a successful collaboration outcomes (Weisband, 2002).

Communication - Communication plays a vital role in collaboration, it is an important factor to complete the collaborative work Nguyen & Duval (2014). Here, communication refers to exchange of information inter-personally, usually in verbal, visual or physical forms and can be measures by user ratings (Cugini et al., 1997). Clark and his colleagues (Clark & Schaefer, 1989; Clark & Brennan, 1991) have described communication was described as the process of maintaining and mutual understanding (Watts et al., 1996), which is essential to the success of collaboration (Baker et al., 1999).

Performance - Task performance is widely used to assess collaborative virtual environment (Bailenson et al., 2002; Bailenson & Yee, 2006). Specifically, participants will be asked self-rated questions to rate the quality of the musical outcome, reflecting the quality of their performance during the collaboration. Moreover, system log will be used to access the interaction related to performance, e.g. number of notes additions and deletions.

Some of these measures (e.g. awareness measures) are very subjective, whilst the others are more objective (e.g. times and length of use of tools). To make full use of participants' reports advantage in collecting subjective feedback and system-logging's strength in collecting objective data, both these two methods were applied in the three studies to cover these three themes.

3.3 Data Collection Methods Used in This Thesis

According to Damianos et al. (1999), four common ways to collect data from the evaluations include: i) *logging tools* for collecting a time-stamped record of participant actions; ii) *direct observation*; iii) *questionnaires/interviews/rating*

scales (open-ended or closed/fixed alternatives); iv) *video and audio recordings*. Most of these methods have been practised in the three studies of this thesis. The following sections of this chapter will first discuss related work on data collection and then brief the methods applied in the three studies.

3.3.1 Interview

As a research method, interviews could be used individually to gather data or in conjunction with other methods as aids to follow up unexpected results, validate other methods and so on (Graham, 1995). Interviews are best suited for unfolding people's perceptions and experiences (Blandford, 2013). Post-task interviews are often used to obtain users' subjective feedback on interactive system participation (Haywood & Cairns, 2006). According to Cohen et al. (2013), there are four main types of interviews: the structured, the unstructured, the non-directive and the focused. A completely structured interview is akin to a questionnaire, in that all questions are predetermined, although a variety of answers may be expected. A completely unstructured interview is more like a conversation, albeit one with a particular focus and purpose. Semi-structured interviews fall between these two poles (Blandford, 2013). Different from text-based questionnaires, interviews take the form of a conversation where the investigator asks questions and the participant replies orally (Blandford, 2013).

3.3.2 Questionnaire

Since this research depends strongly on users' subjective experience (e.g. whether participants feel there is any significant difference between collaboration in different conditions), applying self-rated questionnaire becomes essential.

Numerous questionnaires have been developed for measuring the usability of VR applications. VRUSE developed by Kalawsky (1999) is a well-known usability questionnaire for measuring usability in terms of users' attitude and perception whilst using a VR application. It includes 100 questionnaire items organised under 10 concepts: functionality, input, output, user guidance, consistency, flexibility, simulation fidelity, error correction, presence and overall system usability. The reliability of VRUSE has been tested (Cronbach's alpha > 0.9), however, the main drawback of this test is the large number of questions that the patients are required to answer (Gil-Gómez et al., 2017). Short Feedback Questionnaire (SFQ) developed by Kizony, Katz, & Weiss (2003) is a simplified usability questionnaire related to Witmer and Singer's Presence Questionnaire (Witmer & Singer, 1998). SFQ is composed of eight questions with a five-point Likert attitude scale, measuring the sense of presence, perceived

difficulty of the task, and any discomfort during the experience. SFQ has been applied in Kizony, Raz, et al. (2003); Kizony et al. (2006); Imam et al. (2013).

Both interviews and questionnaire allow gathering subjective data, which is quite important to evaluate visual appeal, preferences, aesthetics, missing functionalities, and very useful as means to compare or cross-reference performance data (Bach & Scapin, 2004). Therefore, interviews and questionnaire are certainly interesting candidates for being applied in the three studies of this thesis. Specific lead questions for interviews and questionnaire items will be tailored accordingly and applied to pursue answers and more insights for the research questions.

3.3.3 Data Log

As briefly mentioned in Section 3.1, fined-grained interactions can be quantified as indicators of collaboration, these fragments can then be logged with time-stamps and reviewed (see Damianos et al. 1999). Experimenters should be very careful and be aware that such a logging tool might slow system's response, which might adversely affect the performance of user tasks (Damianos et al., 1999). An example of using this method is (Damianos et al., 1999), in which a logger was used to record time-stamped speech, typed text, whiteboard activity, and other user events. Another example is (Hamalainen, 2008), in which activity log was applied to cross-verify findings shown by other types of data.

In the three studies of this thesis, a data logger will also be applied to collect objective data, which will then be used to reflect the quality of collaboration. Such a system will be described in detail in Section 3.6.

3.4 Required Features of LeMo

A system supporting the exploration of the research questions of the thesis should allow multiple users to make music collaboratively in the VE, specifically, such a system should be able to: (i) support multiplayer, which will require techniques of networking; (ii) support direct interpersonal interaction, which would require techniques of networking, avatar. These will require knowledge of hand tracking and gesture recognition to enable bare hand interaction; (iii) support music making, this will require knowledge of music instruments; (iv) a virtual environment to immerse the users.

3.5 Technical Development of LeMo

As mentioned in the section above, designing and programming such a system require knowledge in many domains. To make sure all the knowledge had been ready, before building the final system, several prototypes were made. Specifically, Prototype I was built to get the very basic network of Unity working, Prototype II was made to get the data synchronised through the network in real time, Prototype III was a 3D non-immersive step sequencer based on computer screen, it was built to test the step sequencer system, ensuring the music making system was ready. More details about these prototypes are available in Appendix D. The game engine Unity[1] was used to develop the system. Unity is a cross-platform game engine developed by Unity Technologies. Next, three key technical features of LeMo will be detailed: networking, avatars and the music making system.

3.5.1 Networking

The Unity High Level API (HLAPI[2]) is a built-in system for building multiplayer capabilities for Unity games. Below are two basic concepts to understand the networking system used in LeMo:

Host server and remote client - A server is an instance of the game which all other players connect to when they want to play together. The server typically manages various aspects of the game, such as maintaining game conditions and transmitting the data back to the clients. A server can be either a *dedicated server* or a *host server*. The *dedicated server* only acts as as a server, whilst a *host server* acts as both a server and a client. Given that a *dedicated server* usually requires an independent PC, LeMo used *host server* to reduce the need on number of PCs. Clients are instances of the game that usually connect from different computers to the server. Figure 3.1 shows the remote client contacts to a local client, which also plays the role of a server. In that context, the local client on the left is also a *host server*.

Players, local players, and remote actions - In Unity's multiplayer HLAPI system, there are two types of GameObjects[3]: player GameObjects and non-player GameObjects. As shown in Figure 3.2. Each player joined in the game owns a local player GameObject, which is only controlled by the player, all players can only modify non-player objects that are on their client. However players can send command to require the server to send Remote Procedure

[1] Unity: https://unity.com/

[2] Unity HLAPI: https://docs.unity3d.com/Manual/UNetUsingHLAPI.html

[3] GameObjects in Unity: GameObjects are the fundamental objects in Unity that represent characters, props and scenery. Different components (e.g. light, audio, collider) can be added to a GameObject according to demand.

Figure 3.1: The diagram shows the framework of networking of LeMo, i.e. a remote client running on PC2 connects to a host running on PC1.

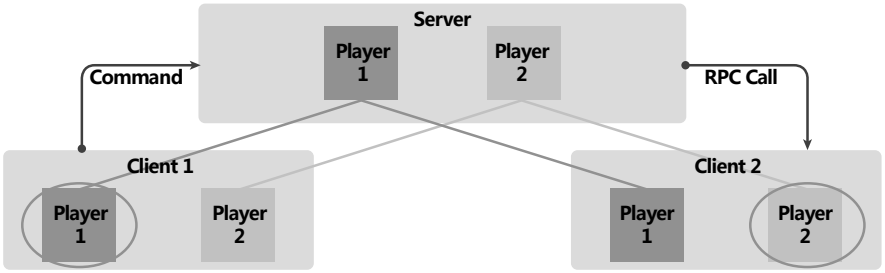


Figure 3.2: The diagram shows how remote actions works in LeMo. Circled players are local players, others are remote players.

Calls (RPC) to perform remote action on all or specific clients. The local player GameObjects (circled player in Figure 3.2) have their instances (non-circled player) on all the clients and the server. All the player objects are running the same scripts. Through Commands or remote procedure calls (RPC), we can send a request to send data and run specific codes on their instances on either the server or on a specific remote client.

3.5.2 Modelling and Animating Avatars

Avatars, according to related work in Chapter 2, are important in maintaining a high level of presence and team cognition (Benford et al., 1995; Nowak & Biocca, 2003; Rovai, 2002), the former of which constitutes crucial aspects of VR experience and the latter is key to effective team work and team performance (DeChurch & Mesmer-Magnus, 2010).

Given the important role avatar plays in SVEs, in LeMo, avatar heads and hands were specifically chosen to be supported as they contribute to the people's face-to-face communication (Bavelas & Chovil, 2006). Leap Motion, a hand tracking device is also frequently applied with HMDs to enable bare hands interaction in VR, see examples at Leap Motion Gallery[4]. To mimic a natural interaction experience, and enable users to use natural gestures to interact

[4]Leap Motion Gallery: `https://gallery.leapmotion.com/`

Figure 3.3: Physical sequencer: Chess Sequencer[6]; Non-immersive digital sequencer: BeatWave[7](upper middle); Daisyphone (upper right, Bryan-Kinns 2004); Immersive VR sequencer: SoundScape Gear VR[8](bottom left); SoundStage VR[9](bottom right).

with music interfaces, and with each other, LeMo applies Leap Motion for hand gestures tracking and Vive headset tracking for head position tracking. A prototype (see Prototype IV in Appendix D.4) was built before implementing LeMo to ensure the tracking system and avatar synchronisation over the network were working properly.

3.5.3 Instrument for Making Music

Given that the studies of this thesis target music novices, step sequencer becomes a great choice because it is relatively simple and does not require much music knowledge. As shown in Figure 3.3, plenty of music step sequencers have been made, either physical (e.g. Chess Sequencer[5]) or digital ones. The latter could be further divided into non-immersive (e.g. BeatWave[6], Daisyphone by Bryan-Kinns 2004) and immersive (e.g. SoundScape Gear VR[7], SoundStage VR[8]), see all examples in Figure 3.3. Technically, all these sequencers contain two parts: (i) A grid (usually has two axes) made of child-element, in which each element stands for a note. Notes usually change along the normal direction of the play head's moving direction, offering the same set of choices for each segment of the loop. The matrix of child-element is also where the player(s) operate the sequencer directly, e.g. adding/removing notes simply by tapping; (ii) A play-

[5]Making music with a Chess Board: `https://www.youtube.com/watch?v=0PWE2QHIXdk`
[6]Beatwave`https://itunes.apple.com/gb/app/beatwave/id363718254?mt=8`
[7]Soundscape Gear VR Launch Trailer: `https://www.youtube.com/watch?v=jLEDgGpnhl8`
[8]SoundStage VR: `https://xinreality.com/wiki/SoundStage`

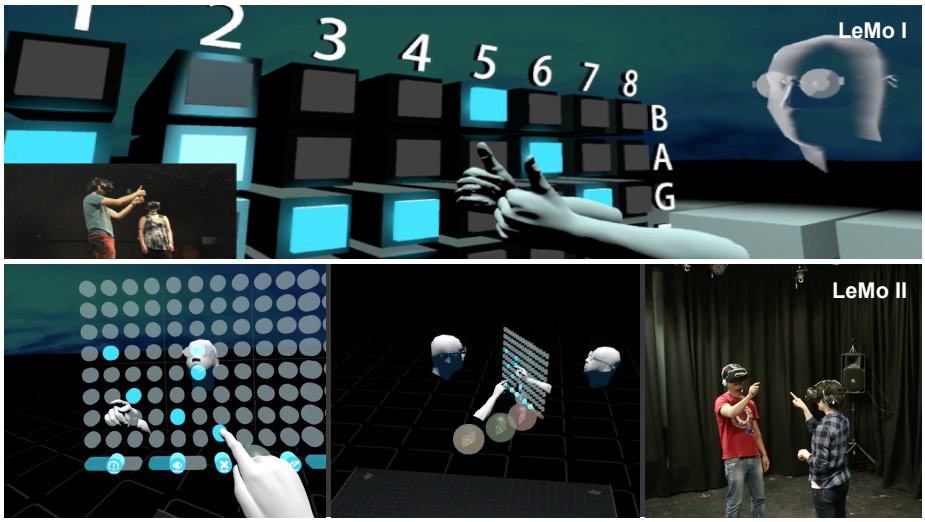


Figure 3.4: LeMos enable two players to work together on a music loop in VR.

head, which keeps moving in a loop, plays the notes it passes over. Prototype III (more details in Appendix D.3), a mouse-click-based non-immersive step sequencer was developed before programming the final LeMo, ensuring the music making system was working properly.

3.6 Final Form of LeMo

As aforementioned, LeMo[9] was built for researching answers for the aforementioned questions. Hence LeMo enables two users to manipulate virtual music interfaces together in an SVE to create a music loop, see Figure 3.4. LeMo was programmed in Unity, models and textures were made in Cinema 4D and Adobe Photoshop respectively. LeMo currently has two major versions: LeMo I and LeMo II[10] (collectively referred to as LeMos). Study I was carried out based on LeMo I whilst Study II and Study III were based on LeMo II.

Both LeMos have three key elements: **i) Music interface** - for producing music. As shown in Figure 3.5, the *matrix* interface contains a grid of grids/dots. Each row represents the same pitch, forming an octave from bottom to top, see Figure 3.5. Users can edit notes by tapping the grids/dots. A vertical playhead repeatedly moves from left to right, playing corresponding activated notes it passes. In this way, each interface can generate a music loop. **ii) Avatars** - each user has an avatar, including a head and both hands (Figure 3.4). Avatars are synchronised with users' real movements in real time, including position and

[9]Full source available at: https://sites.google.com/view/liangmen/projects/LeMo
[10]Check a short video clip of LeMo II at `https://goo.gl/n9ZhPf`

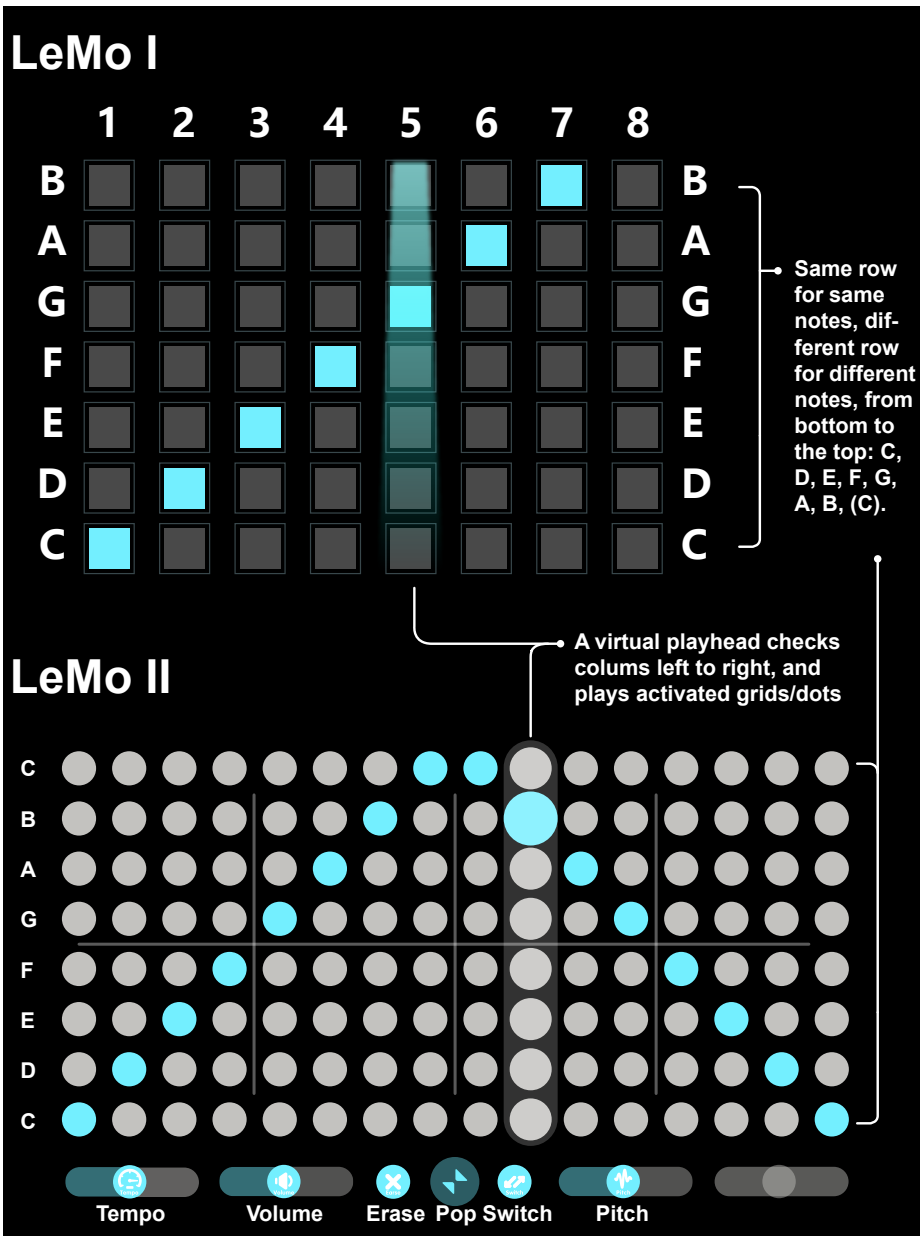


Figure 3.5: The interfaces of LeMo I and LeMo. A C major scale, starting from C4 and finishing at B4, and going back to C4 II (Upper). A C major scale, starting from C4 and finishing at C5, and going back to C4 (bottom).

rotation of heads, as well as gestures. LeMos provide visual aids for collaboration by synchronising the virtual environment (virtual space and music interfaces) and avatars across a network, providing participants with the sense of being in the same virtual environment and manipulating the same set of interfaces. **iii) A virtual space** in which users interact with each other and virtual objects.

Both LeMo I and LeMo II have a **data logging system**. The system of LeMo I is relatively simple, it only logs four types of interactivity: number of note additions, number of deleting self-owned notes, number of deleting other's notes, the average distance between coworkers. LeMo II has a more capable data log system which can also log time-stamped information of user's head and hands' position, musical interfaces' position and pattern.

LeMo I and II have 3 major differences, which are mainly because LeMo II was built later on the basis on LeMo I, and thus provides more and possibly better functionalities. These differences are: **i) Size of interface matrix** of LeMo I is 8*7 whilst that for LeMo II is 16*8. So participants can create an 8-beat loop in LeMo I, and can create a 16-beat loop in LeMo II, see Figure 3.5; **ii) Freedom of manipulating interfaces -** Whilst LeMo I only provides one stationary music interface, LeMo II allows users to generate, remove, position and edit virtual music interfaces (details of the control gestures of study II will be briefed in Study II). **iii) Freedom of manipulating music features -** Compared with LeMo I, LeMo II allows users to control more music features, including instruments, tempo, volume and pitch, see Figure 3.5.

Apparatus - The experimental set-up of using LeMo in all the three studies in this thesis are shown in Figure 3.6. Each participant wears a HTC Vive headset, which has a Leap Motion (hand tracker) mounted, see photos of real participants in Figure 3.4. The devices are connected with PCs and supported by these PCs. During the experiment, two PCs are running LeMo, and are communicating via a LAN cable. When being played, the virtual worlds and virtual objects of LeMo are synchronised in real time, giving participants the illusion that they are in the same viral world, interacting with the same set of objects.

3.7 LeMos Fitting VRMI Design Principles

This section discusses how LeMos fit the 9 design principles proposed by Serafin et al. (2016).

Principle 1: Design for Feedback and Mapping. All modalities (sound, visual, touch, and proprioception) should be designed in tandem (Serafin et al., 2016), e.g. the location and motion of the auditory virtual objects should match the location and motion of the visual virtual objects (Begault & Trejo, 2000).

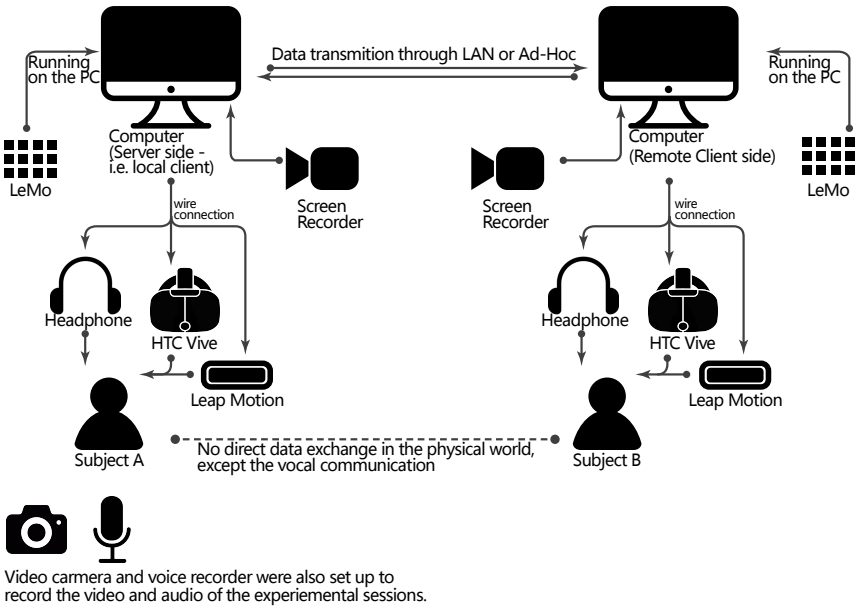


Figure 3.6: Apparatus of LeMos with data recording (video camera and voice recorder) devices included.

As LeMos only involve visual and sound modality, this issue could only be addressed by better mapping visual and audio cues, which had been carefully designed and programmed. Specifically, this was done through: **i)** A mapping between auditory virtual objects and visual virtual object. Spatialised audio has been applied in both LeMos, and each sound of a note is generated exactly from the virtual position of the visual object (grid/dot) that stands for the note. **ii)** A mapping between real movements of hands/heads and virtual movements of hands/heads. By tracking users' hands and heads via Leap Motion and HTC Vive Tracking system, users' real head movements (position and rotation) and hand gestures are mapped to their embodiment's movements and gestures, enabling them to see their own hand gestures, and other's hand gestures and head (avatar) whilst wearing the sight-blocked HMDs. **iii)** A mapping between interaction and feedback. As no haptic feedback was supported in LeMos, additional augmented visual feedback is provided for user's interaction with the music interface in LeMos. For example, the button (grid/dot) changes its position and size accordingly in response to user's touch, mimicking the real-world experience.

Principle 2: Reduce Latency. Serafin et al. (2016) argued that ideally, all interactions should be smooth and minimum latency is preferred, and it is important for VRMIs to provide timely feedback as a response to users' interaction. The latency in LeMos includes two main aspects, which were addressed

respectively: **i) Local latency**, to shorten the latency between real movements and audiovisual feedback, higher frame rate (i.e. frames per second, FPS) was required. To do so, without improving the capabilities of PCs, the codes were optimised for higher efficiency. Also in LeMo II, as the FPS drops with an increasing number of musical interfaces, the maximum number of interfaces is set to 8 to enable a proper FPS. **ii) Latency between server and client**. Impacts of this latency were reduced by: (**a**) Getting the server synchronised takes time, thus in order to provide timely feedback, clients in LeMos were programmed to have high autonomy, enabling timely audiovisual feedback can be given right after local user's actions without waiting the synchronisation. By doing so, feedback can be provided with no impact from this latency and this latency only affects the delay that the users see other players' actions/movements, which is less perceptible and less detrimental. (**b**) The score-based music generation system also reduced the data that needs to be synchronised, thus reduced this latency, since what needs to be synchronised only includes the scores (and possibly some musical factors, e.g. tempo) and the position of the play-head. And then only changes need to be updated. E.g. dot in row 1, column 2 of interface 3 was switched off by player B, rather than updating the data set of the matrix of all dots, which even further reduced workload of synchronisation. (**c**) Reduce the synchronisation requirement. Currently the synchronisation of hand gesture is the most onerous synchronisation load, considering each player has two hands, both of which contain dozens of bones, all are moving constantly. We set the synchronisation rate to 19, which is lower than 24 (the lowest one of the popular FPS standards) but was found to be imperceptible. Note that this only affects updating the others' avatar hands, users still see their own virtual hands at a full FPS. Besides hands synchronisation, the data logger also records data locally to reduce the bandwidth requirement.

Principle 3: Prevent Cybersickness. This is more related to the VR hardware. LeMos apply mature commercial VR equipments to provide efficient tracking, and capable PCs to provide adequate FPS. Both LeMos apply one-to-one mapping between users' virtual and real translations and rotations, which is strongly suggested by Serafin et al. (2016). Apart from these, in LeMo II, as mentioned above, the number of music interfaces users can have is limited to 8 to constrain the consume of CPU resource and ensure a proper FPS.

Principle 4: Make Use of Existing Skills. Using metaphors derived from real world interactions are suggested by Serafin et al. (2016). In LeMos, the ways users touch a button or drag a slider are mimicking mouse clicks and mouse drags, which most people are familiar with in daily lives. Besides, the gestures introduced in LeMo II are also trying to mimic people's daily experience of manipulating physical objects, which might be be easily recalled and mapped

into VR. E.g. in the real world, objects usually start from nothing and then expand from small to large, in LeMo II users can simply move their hands together, pinch and stretch to create an object (music interface). This "stretch gesture" idea was initially inspired and mostly came from a VR demo called Blocks[11]. Another example is the gesture to delete music interface, the packed music interface is a semi-transparent sphere with a similar appearance with a balloon. Naturally, there are two common ways to vanish a balloon, either deflate it entirely or stretch it to explode it. LeMo II mimics a similar experience by allowing users to stretch the sphere to explode the interface to remove it. The third example is pinching inside a sphere (the music interface) to rotate and re-position the sphere, which mimics the daily experience of moving and rotating a real object. More details about the gesture design will be briefed in Section 5.1 and reflected in Section 7.4.1.

Principle 5: Consider Both Natural and "Magical" Interaction. Both types of interaction are applied in LeMos. E.g. re-positioning and rotating objects, and clicking button can be seen as are natural interaction. Button size changes in responding to a touch, pinch to draw 3D annotations, stretching to generate/remove bubble-like music interfaces are more of magical interaction as these are would not commonly happen in reality.

Principle 6: Consider Display Ergonomics. Three main design decisions in LeMo II echo this principle, superficially to address the narrow FOV and low resolution of the lenses of HMDs: **i)** the size of music interfaces in LeMo II was adjusted to a relatively small size, allowing users to see the whole interface whilst still being close enough to manipulate the interface. **ii)** The resolution of HMDs of HTC Vives is relatively low, which means it was hard to read text with small size. So apart from text description, symbols are also provided for the controllers. Compared with texts, symbols might be easier to read, understand and remember. **iii)** The 2-handed interactions are carefully designed and delivered in a way that they can be completed in a small range of space in front of the user, this is because users can only see a small range of field (due to the narrow FOV) and use own their virtual hands inside a small range of space (due to the limited tracking range of Leap Motion).

Principle 7: Create a Sense of Presence. One of the major characteristics that differentiates VR from other media is presence. A proper level of immersion is an end-result of good overall user experience. However, we argue presence is more a result to be measured and reflected rather than a design principle to follow. Both LeMos have not addressed this principle specifically, instead, measures regarding sense of self presence, other's presence, and other's activities were applied in the studies to reflect how well LeMos have done and

[11]Leap Motion Blocks: `https://gallery.leapmotion.com/blocks/`

how the designated independent variables of the study might have affected these sense.

Principle 8: Represent the Player's Body. As noted above in Section 3.5.2, avatars are supported in both LeMos. Gestures instead of controllers were purposely chosen to enable possible gesture-based interaction and communication. Considering no eye-tracking devices are applied in LeMos, to simulate a natural eye-contact experience, sunglasses were modelled and added to the avatar heads to hide the static eyeballs of the avatar head, seeing which might reduce the sense of realness and ruin the sense of other's presence.

Principle 9: Make the Experience Social. As mentioned above, both LeMos are SVEs, which support two players to make music together and communicate with each other in VE. Hence LeMos properly follow this principle.

3.8 Chapter Summary

This chapter has firstly discussed the methodology of measuring collaboration, and the data collection methods used in this thesis, including interview, questionnaire and data log. Then the required features of the prototype named LeMo have been briefed, followed by the description of the development process of LeMos. Finally how we addressed the design principles of VRMIs during the design and making process of LeMos has been discussed. Next chapter will present Study I, which was carried out based on LeMo I.

Chapter 4

Study I - Visual Approach

This chapter explores a multimodal approach to supporting creativity in collaborative music making in Virtual Reality. LeMo I enables people to communicate via visual representations including free-form 3D annotations instead of spoken communication, leaving their full auditory sense to experiencing the joint creation of music. Another feature supported by LeMo I is work identify, it enables people to differentiate their contributions, exploring whether this feature can contribute to awareness and therefore impact the collaboration. This chapter presents a study exploring how people used such visual tools in LeMo I to support their composition process and human-human interaction. Five classes of use of annotation were identified, three of which are particularly relevant to the future design of Sonic Interactions in Virtual Environments. A workshop paper[1] based on this study has been published by IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE).

4.1 LeMo I - A Virtual Reality Step Sequencer

As mentioned in Chapter 2, LeMo I is an SVE as well as a CMM system, which provides two people with a shared musical interface, within which they can make an 8-beat music loop together in virtual reality, see Figure 4.1. In LeMo I, people can draw 3D lines (annotations) simply by pinching their thumbs and index fingers together and moving their hand. These 3D lines are shared and can be seen by both collaborators. In this way people can make lines to convey approvals, ideas, doubts and so on, see an example in Figure 4.3 (right). People can discard all the 3D lines by turning both palms downward to avoid clutter or confusion. By doing so, all the existed 3D lines will fall on the ground

[1] See a short video clip regarding this study at: `https://goo.gl/W6a6jk`

Figure 4.1: LeMo I enables two players to work together on a music loop in VR.

immediately and disappear in 10 seconds, leaving the user a fresh empty space. There is no frequency of use nor time limit for people to add or discard lines.

The virtual environment is where people interact. The virtual surrounding of LeMo I contains three main parts (see Figure 4.2c): foreground (the inner play stage and the peripheral floor, bright lines at the edge of the inner stage indicates user the safe walking area), middle-ground (huge piano buttons in the surroundings) and the background (the sky-box).

Instead of replicating naturalistic shared music making experience, the aim of building LeMo I is providing a shared Collaborative Music Making system, in which the uses of different cues (e.g. visual annotation cue) for collaboration can be explored. The music interface contains a cube matrix, a play line, and a reference system, see Figure 4.2a. The moving play line indicates which step of the musical step sequence is currently being played - each step is represented as one column of possible notes (note B, A, G and so on). To activate or deactivate a note, users touch the button by moving their hands as their physical motion will be mapped to their virtual hands. To provide rich visual feedback and to mimic people's daily experience with pressing physical buttons, the buttons change colour when pressed and also move back and forward to mimic the movement of a physically pressed button.

As mentioned in Section 3.5.2, an avatar is provided for both players. Each avatar includes a head and a pair of hands, see Figure 4.1 and 4.2b. LeMo I uses the HTC Vive and Leap Motion to map real head and hand movement to players' avatars including (i) head position and rotation tracking and (ii) hand gesture tacking including hand position, orientation, and finger tracking. Since the avatars are synchronised in real time, collaborators in LeMo I can see each other's position and gestures via observing corespondent avatars.

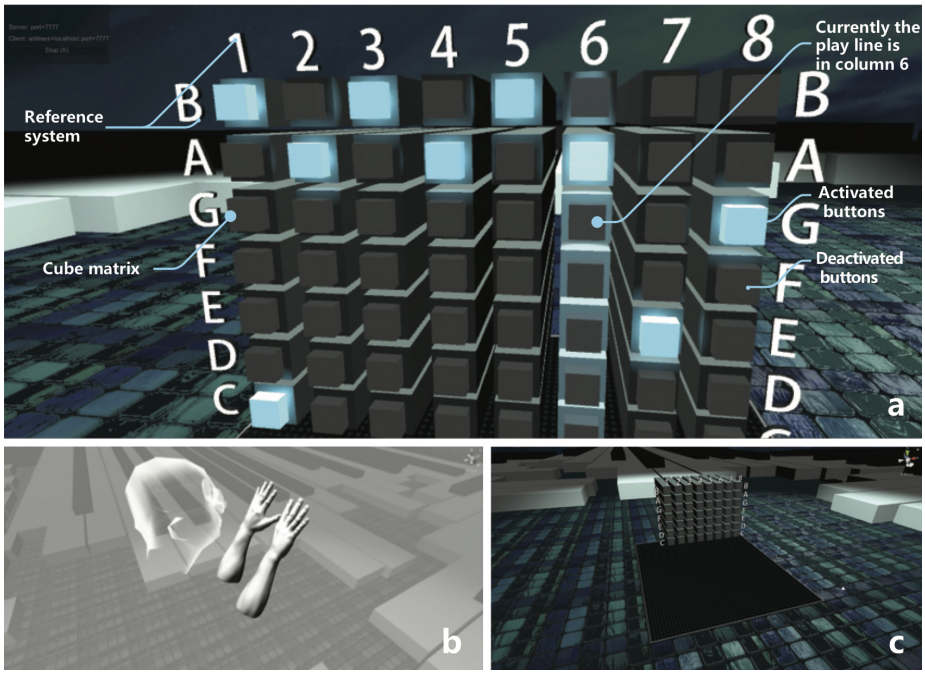


Figure 4.2: Three parts of LeMo I, including the music interface (a), the avatars (b), and the virtual environment that inhabits the former two parts (c).

4.2 Experiment

Although people are physically co-located when using LeMo I, we purposefully do not support spoken communication in this Study. This is because the creative content is in the sound domain and we are interested in exploring how to design systems which foreground the creative uses of sound whilst using complementary modalities to manage the creative process. Based on the related work reviewed in Section 2.2.3 (embodiment in VEs) and Section 2.3.2 (annotations in groupware), identity and annotation, as two visual cues, might provide

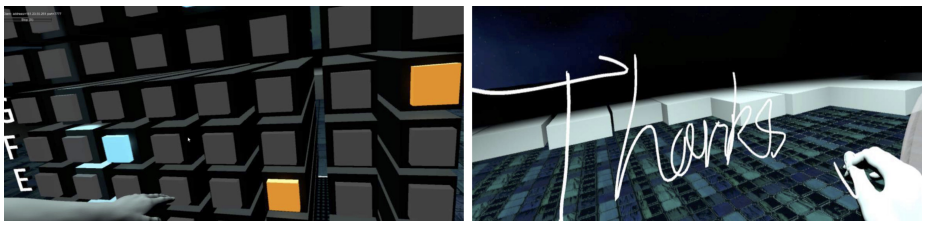


Figure 4.3: When *Work identity* is enabled (left), buttons activated by Participant A turned blue, those activated by Participant B turned orange. Hereafter *Work identity* will be referred to as *Work ID*; When *3D annotation* is enabled, participants could make 3D lines (right).

support for CMM in SVEs. Hence, we propose two hypotheses.

H1 - Work ID can provide additional cue for identity of contributors, and can hence provide support for CMM in SVEs by improving mutual awareness..

H2 - Supplementary communication channel (e.g. 3D annotation) can provide support for CMM in SVEs.

4.2.1 Independent Variables

The two design cues - *Work ID* and *3D Annotation* - are the independent variables in this experiment.

Work ID. Participants' contributions were distinguished by hue, i.e. each participant was assigned a unique hue, see Figure 4.3. Whereas without *Work ID* feature, all participants' contributions turned blue.

3D Annotation: With *3D Annotation* feature, participants can make 3D line markers, see Figure 4.3. Note: the term *3D annotation* used in this chapter covers not only texts, but also symbols and drawings made by using the *3D Annotation* feature.

To explore these hypotheses, the two factors – 3D annotation and work ID were manipulated to create four different conditions. These 4 conditions are:

Condition 1 - *Work ID + 3D Annotation* - referred to as $\mathbf{C_{ID\&Anno}}$;

Condition 2 - *Work ID + No 3D Annotation* - referred to as $\mathbf{C_{ID\_only}}$;

Condition 3 - *No Work ID + 3D Annotation* - referred to as $\mathbf{C_{Anno\_only}}$;

Condition 4 - *No work ID + No 3D Annotation* - referred to as $\mathbf{C_{none}}$.

4.2.2 Dependent Variables

Self-reports (see the questions in Table 4.1), data logging and semi-structured interview were used to obtain data. Specifically, we created the following metrics, measuring collaboration from three aspects: awareness, communication and performance.

Measures of awareness: Participants' reports of self-presence, sense of collaborator's presence, and sense of collaborator's activity.

Measures of communication: Participants' reports of the non-verbal communication quality (Note that to encourage participants' usage of the supplementary communication channel (*3D Annotation*), direct vocal communication was not allowed during the collaboration).

Measures of performance: i) Tacit coordination measure: collaborators' report of their feelings of the tacit cooperation they had with their partners; ii) Contribution measure: the number of contributions people made to the joint product (see Bryan-Kinns & Hamilton 2012); iii) Mutual modification measure: the number of changes they make to the others' work, cf. (Bryan-Kinns &

Table 4.1: Post-Session Questionnaire (PSQ).

Measure	Question
Awareness Measures	PSQ1 (sense of self-presence) - In the virtual world, I had a sense of "being there" PSQ2 (sense of collaborator's presence) - In the virtual world, I could strongly feel someone was there collaborating with me together PSQ3 (sense of collaborator's activities) - I had a clear sense of what he/she was trying to do
Communication Measures	PSQ4 (communication quality) - We had a high-quality non-verbal communication
Performance Measures	PSQ5 (tacit coordination) - I had a feeling, at some points, my interaction partner created notes according to mine PSQ6 (concern of task performance) - I activated or deactivated buttons without without any concerns/worries PSQ7 (performance assessment) - How satisfied are you with the piece of loop music you two finally created

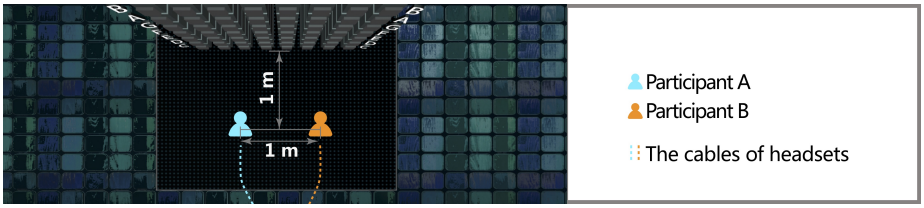


Figure 4.4: The starting positions of participants.

Hamilton, 2012); iv) Concern degree measure: collaborator's reports of the concern level of switching buttons, the higher, the less brave they are to do modifications.

Other measures: i) Content quality measure: participants' report of the satisfaction of the outcome; can reflect the collaboration quality (Bryan-Kinns & Hamilton, 2012); ii) Intimacy measure: how physically close are the participants during the collaboration. This can possibly change during different levels of collaboration.

4.2.3 Participants and Procedure

Thirty-two participants (16 pairs) were recruited via emails and posters to take part in a study of how they used visual annotations to support collaborative music making in LeMo I[2]. A quarter of the participants (25%) had not used VR before, 37.5% of them had tried it only once, nearly a third (28.5%) of them played 2-5 times and nearly 10% played VR more than 5 times. Only 2 rated themselves as music experts. Twelve pairs of participants were familiar with their study partner prior to the study.

After reading and signing informing sheets and consent forms, each pair of participants first received a tutorial of how to use LeMo I and then undertook

[2]The Queen Mary Research Ethics Committee granted ethical approval to carry out the study within its facilities (Ethical Application Ref: QMREC1592).

Table 4.2: Results of System Logged Data

Measure	Results of Factorial ANOVA Tests				
	Effect	*DFn*	*DFd*	*F*	*p*
AA1 - no. of note additions	Work ID	1	31	0.1138	0.7381
	3D Annotation	1	31	22.82	4.070e-05
	Work ID:3D Annotation	1	31	1.880	0.1801
AA2 - no. of deleting self-owned notes	Work ID	1	31	0.7084	0.4064
	3D Annotation	1	31	15.78	3.941e-4
	Work ID:3D Annotation	1	31	0.7563	0.3912
AA3 - no. of deleting other's notes	Work ID	1	31	2.876	0.0999
	3D Annotation	1	31	4.763	0.03678
	Work ID:3D Annotation	1	31	0.8510	0.3634
AA4 - Average distance between subjects	Work ID	1	15	0.2278	0.6401
	3D Annotation	1	15	0.08204	0.7785
	Work ID:3D Annotation	1	15	1.550	0.2322

a task-free trial of LeMo I for 5 minutes, during which they could change music notes and make annotations, helping them get familiar with LeMo I. After that, each pair undertook four sessions of composing music, each lasting 5 minutes. See participants' starting position in Figure 4.4. Each session covers an experimental condition and the sequence of the conditions were fully randomised to counterbalance the learning effect. The study ended with a semi-structured interview (around 5 minutes) in which each pair of participants had a talk with a researcher, the audio was recorded and post-hoc transcribed.

4.3 Results

4.3.1 System Logged Data

For the 4 measures developed based on System-Logged data, as this experiment is 2x2 within-subject factorial design experiment, Factorial ANOVA tests were conducted as the omnibus tests to compare the main effects of *Work ID* and 3D annotation and the possible interaction effect between *Work ID* and 3D annotation, see Table 4.2. As in this experiment, each factor only has two levels, and no interaction effect was identified between the two factors, no post-hoc tests were needed to reveal the source of the significant differences. The following sections report the detailed results of the system logged data.

Number of note additions - As shown in AA1 of Table 4.2, repeated measures ANOVA reveals participants had significant less note additions when 3D annotation was available ($F(1,31) = 22.82$, $p < 0.001$). The possible reason for this is that it took participants considerable amount of time to draw annotations, which could be used to make notes.

Number of deleting self-owned notes - Similarly, participant deleted

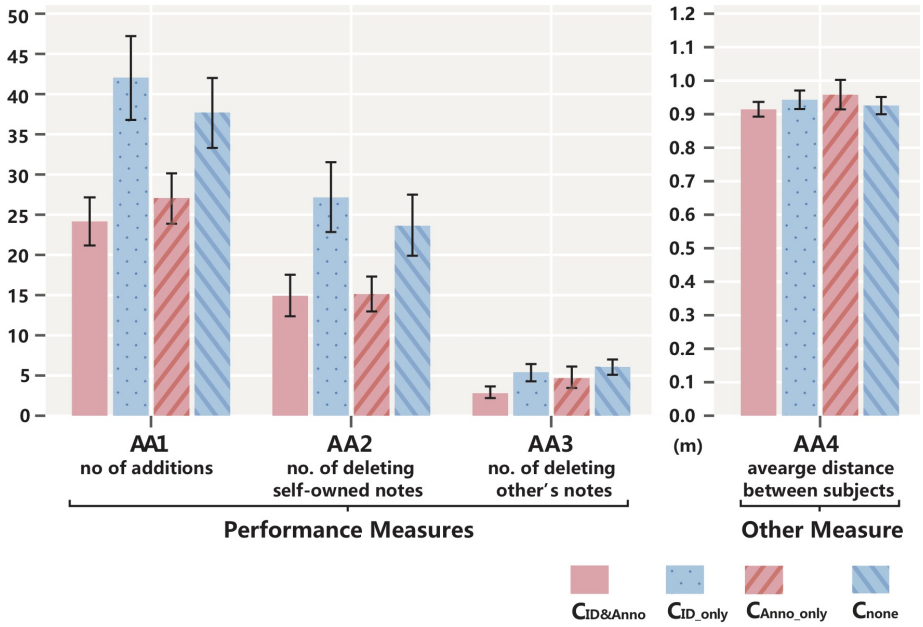


Figure 4.5: Means for system logged data: number of additions, deletions, and average distance between the collaborators.

significantly less self-owned notes when 3D annotation was available ($F(1,31) = 15.78$, $p < 0.001$), see AA2 of Table 4.2. This is probably due to the same reason – participants spent time on making annotations, which reduced the time they spend on note editing.

Number of deleting other's notes - Again similarly, as shown in AA3 of Table 4.2, participants deleted significantly less other's notes when 3D annotation was available ($F(1,31) = 4.763$, $p < 0.05$), which is probably due to a similar reason.

Average distance - The distance between the two collaborators was logged every 2 seconds. For each condition, a mean was then calculated for the pair. as shown in AA4 of Table 4.2, the Factorial ANOVA revealed no significant differences among the four conditions, indicating no significant influences from the two factors on the distance between collaborators.

4.3.2 Post-Session Questionnaire

Bar plots (Figure 4.6) were drawn and Factorial ANOVA tests were run to compare the ratings of the Post-Session Questionnaire. The results revealed neither significant impacts from the two factors (*Work ID* and *3D Annotation*) nor interaction effects between the two factors on the measures listed in Table 4.1.

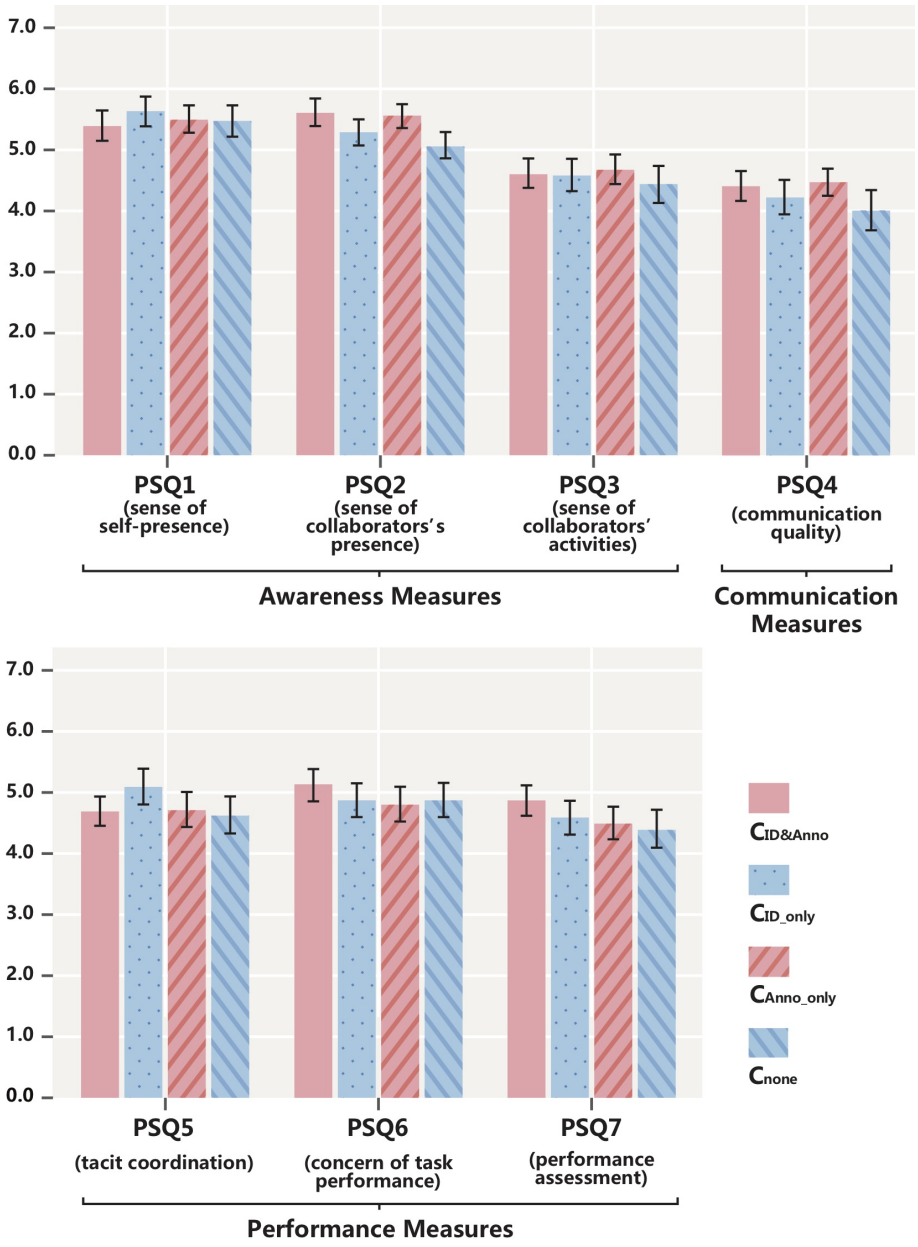
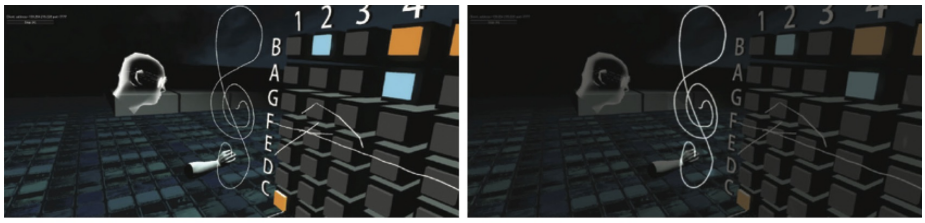


Figure 4.6: Mean ratings for the Post-Session Questionnaire.



before highlighting after highlighting

Figure 4.7: To increase the visibility of the annotation and increase their legibility outside the VR, all the annotations in later figures have been highlighted by darkening the background and brightening the annotation lines.

4.3.3 Annotations

Note: participants' annotations are highlighted in the images in this chapter to improve their legibility outside the VR, see an example in Figure 4.7. Seventy-eight annotations were post-hoc identified and categorised by the researchers according to the annotations for Mutual Engagement classification scheme (aME classification, see Bryan-Kinns 2011): presence, making it happen, quality, social, localisation. The aME classification was developed to analyse people's communication through annotation in the distributed web-based and phone-based music making systems - Daisy, which differ from LeMo I in that these systems were two-dimensional non-immersive user interfaces with no avatar representations. However, the shared and mutually modifiable nature of the real-time collaborative music making in LeMo I shares many collaborative features with Daisy which makes the classification scheme useful as a starting point for understanding the use of annotations in LeMo I. The following sections report on the kinds of annotations participants used when making music together in the LeMo I, and later sections reflect on these annotations and the utility of the aME classification scheme for SVEs.

4.3.4 A Long Conversation

Before reporting on the kinds of annotations found in use of LeMo I, it is worth noting that almost all the annotation-based conversations that emerged were very short (one or two conversational turns). However, there was one exception illustrated by the 5 conversational turns in Figure 4.8.

4.3.5 Presence

The concept of presence has been defined and interpreted in different ways (e.g. Heeter 1992; Slater & Wilbur 1997; Witmer & Singer 1998; Slater 2009). Presence is a subjective experience (Heeter, 1992; Slater et al., 1994) which can

Figure 4.8: A long conversation formed of six short annotations.

Figure 4.9: Presence annotation: "XiaoB" (a) and "it me" (b).

greatly affect collaboration (Frécon et al., 1999; Romano et al., 1998) - having knowledge of oneself and those we are working with is important in collaboration. An earlier study found many participants in distributed music making used annotations as a way to express and query presence, helping participants know about each other's existence (Bryan-Kinns, 2011). In this study based on LeMo I, only two users used annotations to convey presence, one wrote "*XiaoB*" (the participant's name) and the other wrote "*it me*" to tell the collaborators their presence and identity, see Figure 4.9. The reason that fewer people used annotations to convey presence might be the avatars provided a sense of presence and identity not available in the original Daisy studies. Avatars intuitively show the collaborators where they are, what they are doing, and where they are looking. Another possible reason is that the collaborators were co-located (though in VR) and, that they already met in the real world before entering the virtual space of LeMo I.

4.3.6 Making It Happen

Annotations were used to support the process of collaborative music making in several ways explored below: (i) Turn taking; (ii) Composition thoughts; (iii) Working area arrangement (iv) Confusion expressions.

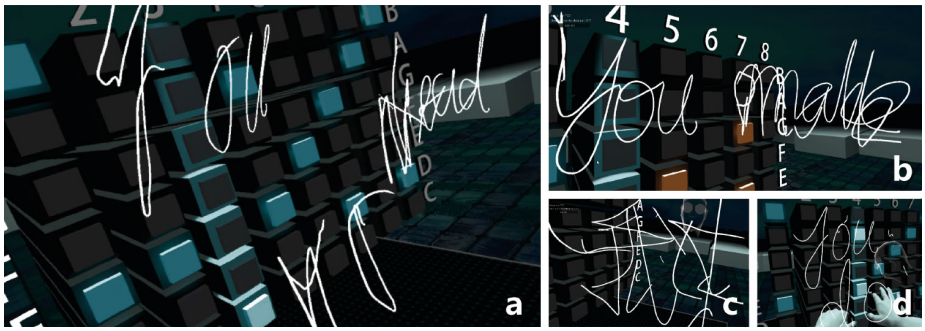


Figure 4.10: Turn taking annotations: "you go ahead" (a); "you make" (b); "I make" written in Chinese (c); "you do"(d).

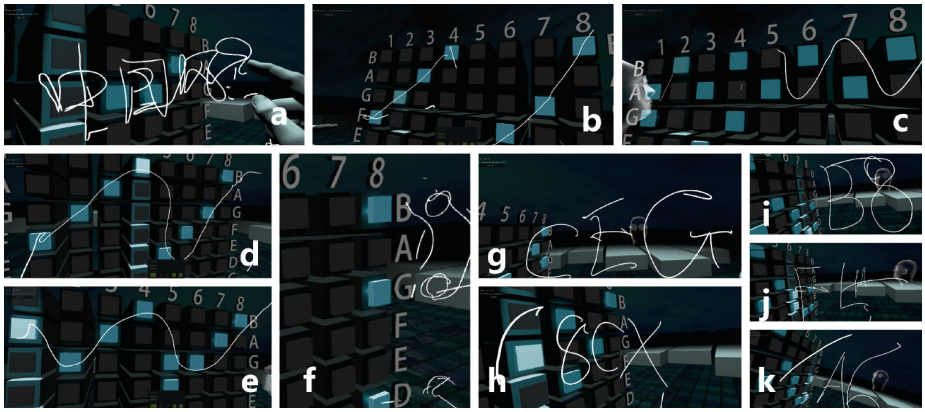


Figure 4.11: "Chinese style?" written in Chinese (a); Patterns formed by notes (b, c, d, e); Note markers (f); References of notes (g, h, i, j, k).

Turn Taking

Although LeMo I allows simultaneous editing of the shared musical loop, at some points participants took turns to contribute the musical notes, and used annotations to manage the process. As shown in Figure 4.10, participants wrote "*I make*" or "*you do*" to switch who had the active role. By doing so, the active person could either require or give away full control of the music interface until they agree a turn change - note that there was no explicit ownership control of the musical interface, so in these cases, participants were self-managing their access to the shared musical loop.

Composition Thoughts

Some annotations emerged were to express composition ideas, ranging from the highest level - music style, to the medium level - patterns formed of notes, and to the most specific - single notes.

Figure 4.12: Annotations for working area arrangement.

Figure 4.11b, c, d show participants sketching out composition ideas by drawing lines aligning with possible notes on the grid. This is slightly more specific communication of musical ideas than the suggestion of "*Chinese style*?" (Figure 4.11a). These annotations were drawn before activating the corresponding buttons to make a plan and share the plan, possibly so that the partner can help with the construction of the sequence of notes. If these compositional sketches were drawn afterwards (e.g. Figure 4.11e), they were used to demonstrate a musical idea. In both cases, this kind of annotation may have helped participants to formulate and understand the collaborative music plan/idea better.

More directed use of annotations in composition is illustrated in Figure 4.11f where the participant made three dot-markers near the column reference system (rows B, G, D specifically), asking the partner to make notes in these three columns which resulted in the partner adding these notes to the shared musical loop. A similar case is shown in Figure 4.11h, in which the partner was asked to make notes in row C, E, and G. Participants also directly wrote the reference to ask partners to change specific notes, see Figure 4.11h, i, j, k.

Area and Position Arrangement

Participants also used annotations to divide the working area and to manage their own work focus in the VE. Figure 4.12a shows an example in which a horizontal line was drawn, dividing the music interface into two parts, each for one participant. The pair did compose within their own working area after the line was drawn. A text annotation "*Switch*" was then used to switch positions (i.e. to swap from top to bottom), see Figure 4.12b. These annotations may have helped participants to manage their working areas and space.

Confusion Expressions

Participants used annotation to write "*what*" or question mark to presumably express confusion about their partners' activities given that such annotations were made directly after their partners drew, wrote and changed notes, or made gestures. Figure 4.13 illustrates typical indicators of confusion.

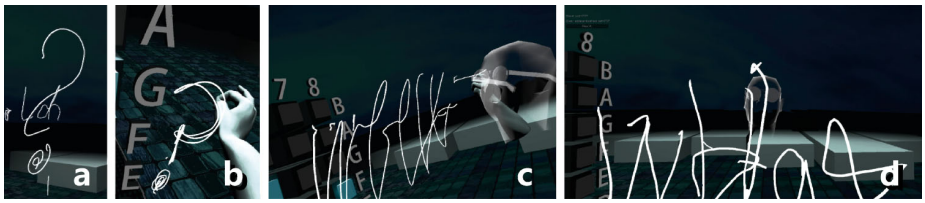


Figure 4.13: Confusion annotations.

4.3.7 Quality

When creating the music loop, reflecting and exchanging the ideas of the quality of the piece is crucial to smooth the cooperation and ensure a final output with good quality. In LeMo I, participants used annotations to express and exchange their judgements of the quality. These annotations are usually short words or simple shapes, either positive (e.g. "*OK*", "*Nice*", "*Cool*", "*Good*", heart shape) or negative (e.g. "*No*"), as illustrated in Figure 4.14. It may be that some of the confusion expressions such as "*?*" were actually indicators of queries of quality, not just queries about the process. It is also interesting to note that positive words may convey different meanings when temporal relationships change. For example, a "*yes*" written shortly after a note addition means the writer's satisfaction with the addition while an "*OK*" write much later with a certain addition has fewer relation with the addition and possibly means more satisfaction about the whole piece. These emerging annotation-based judgements help collaborators exchange feelings about the piece being made, reduce the idea variation and strengthen the cooperation on the activity.

4.3.8 Social

Beyond music making and process management, annotations were also used for non-task related purposes as illustrated in Figure 4.15 and Figure 4.16. Figure 4.16 shows detailed steps of a social drawing activity stated by one participant, whose partner then saw this and joined in with the drawing activity and they finished the drawing together. It is interesting to note that five human doodles appeared, two of which were drawn collaboratively. The possible reasons for its frequent emergence can be that participants were inspired unknowingly by the kinetic avatar or people just naturally love to draw faces. Although social annotations did not contribute to the music directly, making these lighthearted drawings, as a social interaction, develops a relationship between the collaborators.

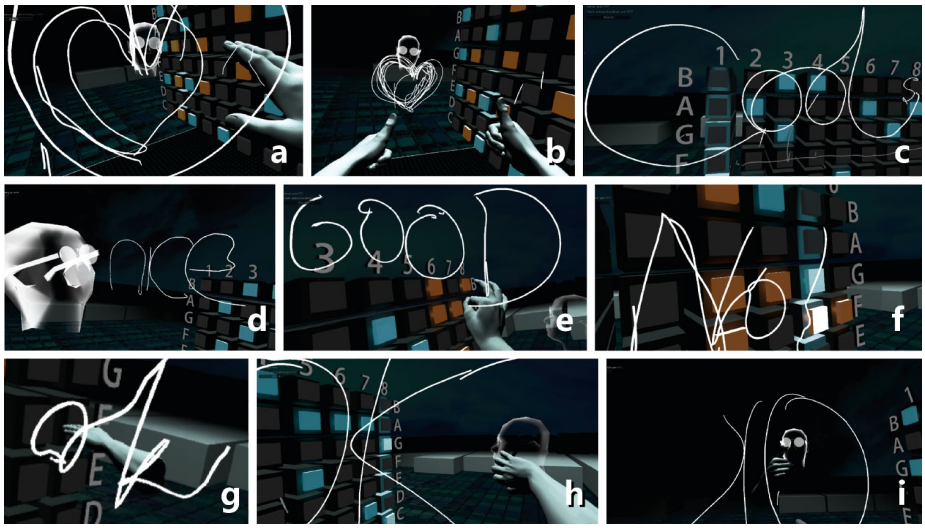


Figure 4.14: Quality Annotations.

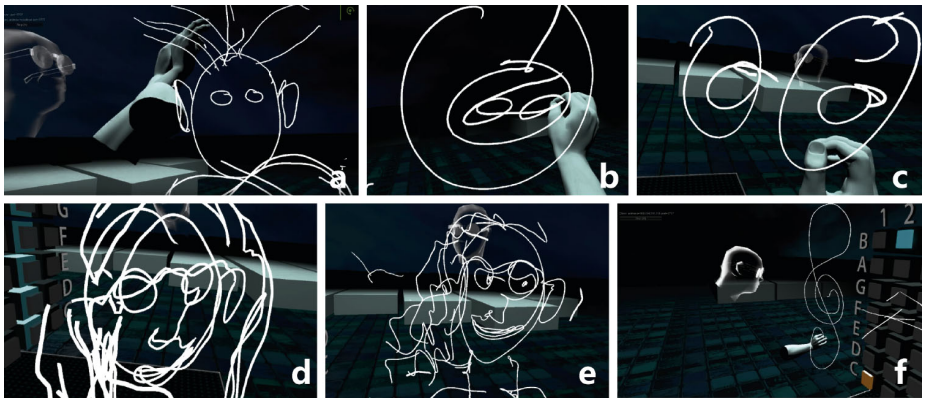


Figure 4.15: Annotations for social purposes.

4.3.9 Localisation

Bryan-Kinns (2011) identified the frequent use of annotation as a localisation cue (mainly by drawing arrows), but in LeMo I we only found one case of this use of annotations. In this case, the participant drew an arrow and, from the review of the interaction, successfully obtained their partner's attention, as illustrated in Figure 4.17. However, in this case the arrow may have been more to attract attention to the activity rather than to highlight a specific part of the joint creation. It may be that annotations are not used for localisation in LeMo I as people can simply draw each other's attention to parts of the VE by waving their hands and then pointing to specific locations.

Figure 4.16: Annotations for social purposes.

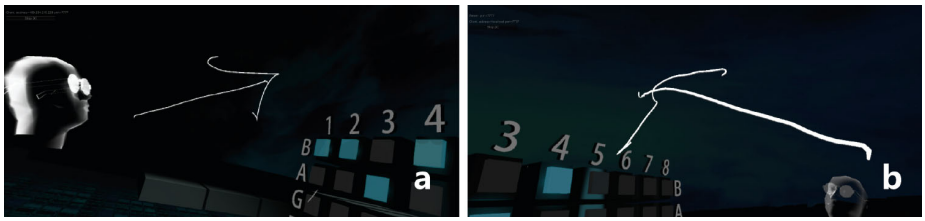


Figure 4.17: A participant drew an arrow (a), and this successfully drew their partner's attention to the intended area (b).

4.3.10 Interviews

Post-task interviews with participants revealed more reflective insights into the use of the annotations. The interviews were transcribed (around 5,000 words of transcription) and a thematic analysis (see Braun & Clarke 2006; Yin 2017) of the transcription was undertaken. The thematic analysis started with a reading through of the transcript, then an inductive analysis of the data was performed, and relevant patterns were collapsed into codes. Next, these codes were combined into overarching themes, which were then reviewed and adjusted until they were appropriate for the codes. All the codes and coded segments are available in Appendix A.3. In total, 63 coded segments and 3 overarching themes emerged from the thematic analysis. The 3 themes are: (i) *3D annotation*'s advantages and disadvantages, (ii) *work ID*'s advantages and disadvantages and (iii) reporting LeMo I system. Next, the advantages and disadvantages introduced by *3D annotation* and *Work* ID will be reported.

3D Annotation's Advantages. Many participants described that they had a positive feeling when they could write something to support their communication. They reported annotations were used to make "*signs and symbols*" to support composition, or to "*create drawing together [...] like a physical warm up*". Participants also reported that annotations exceeded vocal communication in some ways, e.g. "*with the lines, [they] could just circle the notes to say that was [note] G and go back to [note] C, from that perspective, drawing was more effective*". Many participants reported that they successfully understood each

other's intentions via the annotations, e.g. one participant drew a line and "*used the line to affect the partner*", guiding their partner to move notes to lower positions, the partner fully understood and reported they "*did the changes*". Other examples mentioned are showing satisfaction by "*writing an OK*" or using "*Hi*" for greetings.

3D Annotation's limitations. Meanwhile, writing and reading in 3D space were reported by participants to be quite different from the real world and these differences caused inconvenience and problems. For instance, the 3D nature of the annotations reduced their readability, it only "*makes sense to [them] from [their] perspective[s], because it was 3D*". For ease of identifying, "*[they] need to stand where the person wrote it stood*". Furthermore, making annotations was reported to be time-consuming, and "*when [they] finish[ed] it, it [did] not make sense*" anymore. Also, the low accuracy of movement tracking led to annotations being drawn at quite large sizes, which then led to a limitation of "*how much [they could] write*". Finally, participants reported that it was hard to notice each others' annotation activities, a participant "*waved hands to [their partner], but [the partner] did not see*", the participant "*had to wave hands [closer], directly in front of [the partner]*" to draw their attention to the annotations, so to get the annotations read. This was probably due to the narrower field of view (FOV) in VR compared with real life, i.e. around 100 horizontal degrees of FOV within HTC Vive headsets *vs* about 200 degrees binocular FOV in real life, see RLensLab[3] and (Hunt, 2018).

Work ID led to a clearer ownership and understanding - Some participants reported when *Work ID* was activated, they had a better feeling of the "*ownership*" (participant 6B) of the work and could better develop and see their own composition idea as well as what their collaborator's intended to do. *Work ID* functioned as a memory aid, helping to reduce the intrusiveness between each other's work and make the music pattern clearer. Some participants believed the *Work ID* did not matter because they administrated several columns and already had a rough sense of the ownership.

The concern of performing changes upon the work. Opinions upon this differ greatly. Some participants believed the concern degree depended on the relationship, e.g. lower concern when they know each other very well. Some reported it depended on the expertise of their partner, e.g. more respects and less willingness to change other's work when the other one is a music expert. Some participants added and deleted notes more freely when *Work ID*, because *Work ID* enabled them to know clearly the ownership of the notes and helped them get rid of worrying about deleting other's work accidentally. While some

[3]RLensLab: `https://vr-lens-lab.com/field-of-view-for-virtual-reality-headsets/`. Accessed: 2019-09-06.

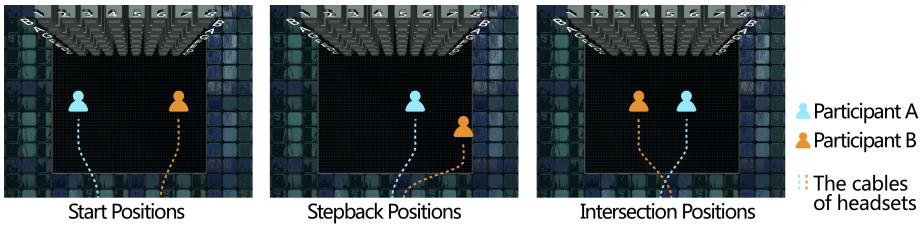


Figure 4.18: Start position, step back position and intersection position.

others believed they felt "*more brave when everything is in the same colour*" (participant 14B).

4.4 Reflection

The section reflects the impacts of adding *Work ID* and *3D annotation*, and propose implications for SVEs supporting collaboration.

4.4.1 Impacts of Adding Work ID

Results above show that *Work ID* has no significant impact on *awareness*, PSQ 1, PSQ 2 and PSQ 3 measure its impacts on awareness. The first possible reason could be that the embodiment part of LeMo I (the head and hands) already provided rich information about identity and activity, including who is there, where are they, what they are doing, who did the change and when. Another possible explanation could be the existence of segmentation of the working area. The pair started the experiment with one participant on the left and one on the right, in all cases, the pair did not exchange their positions (see Figure 4.18). Instead, participants preferred stepping back if their partner needed operation space. In other words, the participants kept their positions (left or right) throughout the sessions, which led to a work area segmentation. This segmentation possibly provided a sense of the ownership of columns and buttons, which reduced the impact of *Work ID*.

Nor did *Work ID* have a significant impact on the *communication*, which is not strange as ID's main function is showing authorship of contributions instead of being a communication channel.

In terms of *task performance*, although the result of self-report indicates that *Work ID* did not affect the concern of switching buttons significantly, interview results revealed *Work ID* did affect participants differently. For some of the subjects, they felt more free of making changes to each other's work under when *Work ID* is not deactivate, because they feel less ownership of the buttons. Thus, for them, making changes to each other's work would be less

disrespectful when there is no *Work ID*. Whereas, for some others, they felt more comfortable with making changes under the condition *Work ID*, as *Work ID* provided a clear cue of button ownership, helping them eliminate the risk of changing other's buttons accidentally (especially when their partners have a higher specialism in music field). All of these people were trying to be respectful of each other's contributions. This is also proved by the result of the system logged data L3 (Table 4.5). The result of L3 indicates that when $C_{ID\&Anno}$;, participants deleted other's notes significantly less than C_{Anno_only}, showing that subjects became less willing to delete other's notes when *Work ID* was activated.

4.4.2 Impacts of 3D Annotation

Similar to the findings in Bryan-Kinns (2011), the annotations that emerged in use of LeMo I could be classified into four types: presence, making it happen, quality, and social, but unlike the aME classification, localisation appears to be managed through avatar interaction. This similarity shows that *3D Annotation*s may serve similar functions in an immersive collaborative music making system as they do in a 2D non-immersive CMM system. However, much fewer annotations are used to convey presence compared with the findings of Bryan-Kinns (2011) which may be because avatars already contributed toward the presence very well, or due to longer time it took to make *3D annotations*, or may due to the physical collocation of participants with LeMo I compared to the Daisy studies which were distributed online. The length of the musical loop in LeMo I is 8 beats whereas in the Daisy studies the length was 48 beats which may have also had an effect on the kinds of annotation produced as the LeMo I loop was simpler and required less temporal organisation. Regardless of these issues, using aME to classify annotations in a study of CMM in an SVE rather than a 2D user interface indicates that the annotation classification scheme is applicable beyond the Daisy systems it was previously used to evaluate.

For Sonic Interaction Design of Virtual Environments, the findings of this exploratory study indicate that 3D graphical annotations of a virtual environment can support a creative activity where the co-produced sound is prioritised over other modalities - CMM in this case. Conversation was prevented during the creative process to allow us to explore how to support collaboration without interrupting or interfering with the music being created by collaborators. The step sequencer used in LeMo I was intentionally simple to allow initial exploration of the role of annotations without conflating this with the complexity of an interface. For richer and more complex sonic creation and exploration in VR we suggest that annotations could usefully support communication about the process, quality, and also social aspects of interaction without compromis-

ing the joint product being produced. It may facilitate a foregrounding of the creative sound product to such an extent that the sounds created are able to use the full width of the sound domain at the exclusion of all other parts of the human-human interaction necessary for collaboration.

Whilst the annotations of LeMo I supported co-creation of music in LeMo I, they did generate some issues. More specifically, making annotations and viewing them were reported to be very different from daily experience. Participants needed to get used to controlling strokes by pinching and releasing fingers. Besides, compared with writing or drawing with a real pen, the LeMo I has less accuracy in supporting these. To increase the readability of written contents and sketches, participants tended to write or draw in a bigger size, which resulted in a limitation of how much they can write and draw. But on the positive side, the larger size made it possible to write and draw together, which expanded the range of annotating action, making it less personal but more socially friendly and more accommodating to multiple people. Another unexpected problem found in this study was that 3D annotations can, of course, be viewed from many angles so written text is often reversed for a participant's collaborator, especially if they write in the space between themselves. This clearly decreases the readability of the annotations. Some participants wrote in reverse to try to compensate for this issue, see an example shown in Figure 4.14h, i. Future development of the use of annotations in VR would need to explore how this mirroring could be addressed.

4.4.3 Design Implication

Based on these findings of 3D Annotation, we propose the following design implication for SVEs focusing on supporting CMM:

Although *3D Annotation* has shown its potential in making signs and simple texts, its weakness in its current form can not be neglected. Thus, we recommend SVEs to contain *3D Annotation* only as a subsidiary tool to support communication, aiding the usage of main communication channels (audio and visual communication). This is especially suggested when the co-produced sound is prioritised over other modalities and avoiding any impacts on it is required.

4.5 Chapter Summary

This chapter has detailed LeMo I, and explored how two design features - *Work ID* and *3D Annotation* - were used to support CMM. Four key kinds of use of annotations were identified: making it happen, quality, and social activities.

Previously identified uses of annotations for presence and localisation were not found as commonly as in previous studies which may be because the LeMo I avatars offered good support for these aspects of collaboration. Insights into the differences between annotations in a VR medium and non-VR media have been given. The next chapter will present Study II, which explores space usage in CMM in SVEs.

Chapter 5

Study II - Spatial Approach

As mentioned in the introduction chapter, Shared Virtual Environments (SVEs) have been extensively researched for education, entertainment, work, and training, yet there has been limited research on the creative aspects of collaboration in SVEs. This raises questions about how to design virtual working spaces to support collaborative creativity in SVEs. In this chapter, we present a study of LeMo II, in which 42 users composed music together using three different virtual working space configurations. Results indicate that (i) two types of territory and working configurations emerged during collaborative composing, (ii) when made available to them, personal working spaces were extensively used, and were considered to be essential to successful collaborative music making, (iii) a publicly visible personal working space was preferable to a publicly invisible one. Based on these findings, three corresponding design implications for Shared Virtual Environments focusing on supporting collaborative creativity are given. A conference paper[1] based on this study has been published in ACM Creativity and Cognition 2019.

5.1 LeMo II - More Freedom

Recall that LeMo I was built to explore the use of 3D annotations and work ID in collaboration in SVEs and to fill the gap between CMMs and SVEs. LeMo I provides one music interface with fixed location, i.e. users could not have more than one music interface or re-position the music interface. These restrictions may limit the amount and range of the interactivity of the users. These constraints need to be broken to investigate the uses of space in CMM in SVEs. As such, we built LeMo II, which is an extensively modified version of LeMo I. The main modifications are LeMo II enables players to generate, re-position

[1]See a short video clip regarding this study at: `https://youtu.be/nk781TFleZI`

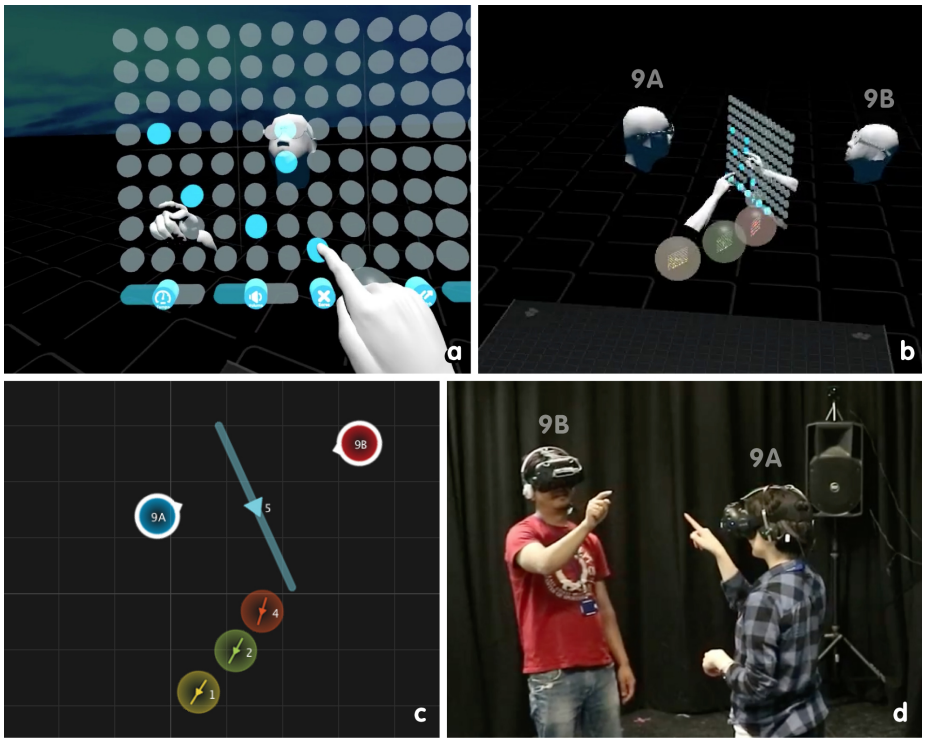


Figure 5.1: A pair of participants (Group 9, participant 9A and participant 9B) are editing notes on the same interface. First person perspective of 9A (a); third person perspective (b); arrangement plot generated by a reanimation system (c); real world experience scene (d).

and remove music interfaces, players can now have up to 8 interfaces. More freedom are provided in controlling music features (tempo, pitch, instruments and volume, 8-beat loop replaced with 16-beat loop).

As briefed in Section 3.6 and shown in Figure 5.1a, b, LeMo II also has three key elements: **(i) Music interface.** LeMo II allows users to generate, remove, position and edit virtual music interfaces, which have two modes: *sphere* and *matrix* (Figure 5.2d). Users can generate up to 8 *spheres* with pinch and stretch gesture, see Figure 5.2a. Both the *sphere* and the *matrix* can be switched in between, re-positioned or removed by manipulating the *sphere* or the pop button of the matrix with corresponding gestures, see Figure 5.2d. The *matrix* interface contains a grid of 16 x 8 dots, with controllers at the bottom. Each row represents the same pitch, forming an octave bottom to top. Users can edit notes by tapping the dots. A vertical play-line repeatedly moves from left to right playing corresponding notes. In this way, each interface generates a 16-notes music loop. Three controllers (tempo, volume and pitch) and two functional buttons (erase and switch) are located at the bottom of the *matrix*

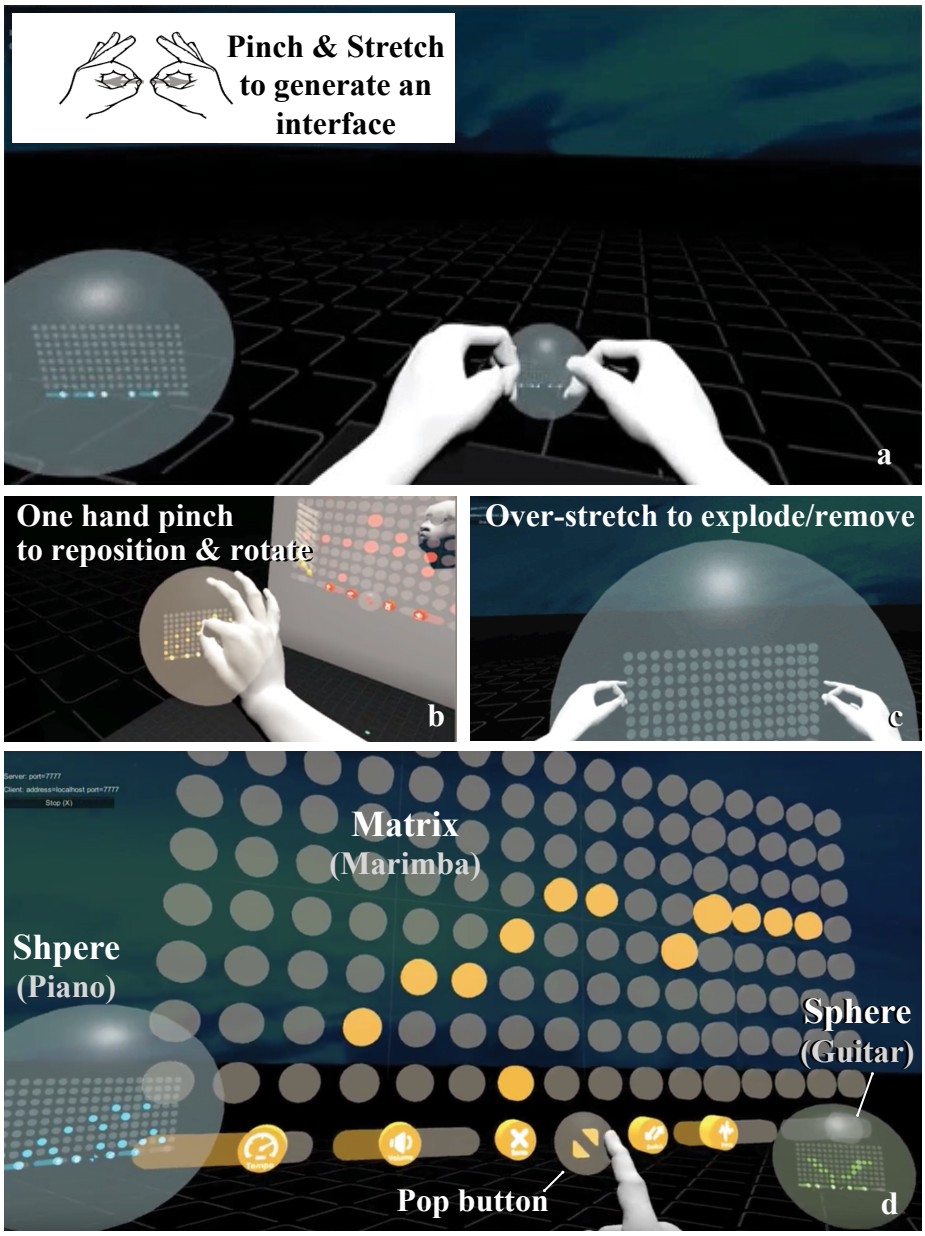


Figure 5.2: (a) The gesture to generate a new music interface; (b) The gestures to re-position and rotate an interface; (c) The gesture to remove an interface; (d) *Matrix* (opened interface) and *sphere* (packed interface), double click the pop button to switch in between.

interface. **(ii) Avatars.** Each user has an avatar, including a head and both hands (Figure 5.1). Avatars are synchronised with users' real movements in real time, including position and rotation of head, and gestures. **(iii) A virtual space** that includes a grey stage with a grid pattern (Figure 5.1a, b and Figure 5.2a). Three types of stage configuration were designed for this study and will be detailed later.

Similar to LeMo I, LeMo II also has: spatialised audio so that users can hear where the sounds come from and the volume drops with distance; A data-log system to log user's interaction (e.g. users' heads' position and rotation, musical note edits); A voice notification system to facilitate the experiments, e.g. in experimental scenario users will hear "1 minute left" and "end of session" notifications.

5.2 Experiment

In creative group-work, enabling people to shift between individual creativity and tightly coupled collaboration is needed (Dourish & Bellotti, 1992; Heath et al., 2002). When it comes to musical related collaboration, to achieve this goal, we need to provide users with spaces with not only exclusive access but also audio exclusivity. Namely, only the owner can access and hear music interface inside that space (no necessary effects on vocal communication), so the owner can develop their own ideas without affecting others. Studies have shown adding personal workspace is helpful to collaboration, and a visibility to co-workers' workspace is preferred but had no explicit effect on how the personal workspace is used and can possibly draw some negative effects (Fencott & Bryan-Kinns, 2010). We would like to explore whether this applies to group and personal workspace in a VE setting, and whether the addition of personal workspace affects the way participants using the space and may impact their collaboration. We developed three hypotheses in keeping with research on collaboration in non-immersive media such as tabletops:

H1 - Different types of territories (personal/group territory) will emerge during collaboration in SVEs, and people's interactive behaviour will change according to their location, cf. (Scott et al., 2004; Xambó et al., 2013).

H2 - Providing personal spaces will facilitate efficient collaboration in SVEs, cf. (Fencott & Bryan-Kinns, 2010).

H3 - Transparent personal space will be preferred, cf. (Fencott & Bryan-Kinns, 2010).

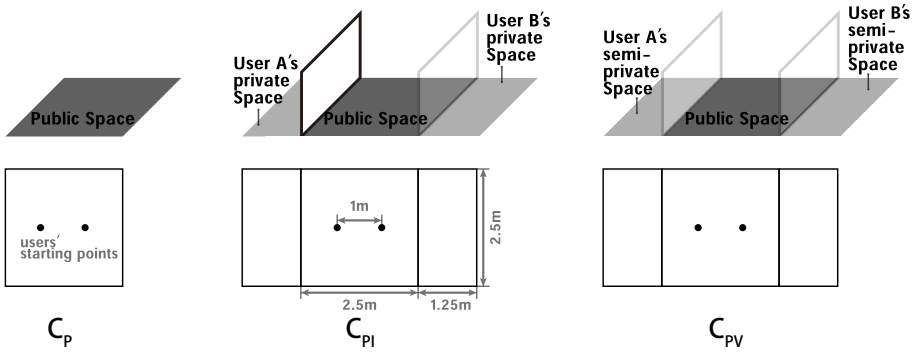


Figure 5.3: Three settings of spaces of the experiment, directional view(upper), top view (bottom).

5.2.1 Independent Variable

Spatial configuration is the independent variable in this experiment. To investigate the hypotheses we designed three space configurations as the independent variable levels, as shown in Figure 5.3, these include:

Condition 1 - **P**ublic space only (referred to as C_P): where players can generate, remove or manipulate *Spheres*, and have equal access to all of the space and the music interfaces.

Condition 2 - **P**ublic space + Publicly **I**nvisible personal spaces (C_{PI}): in addition to the public space (in C_P), each user is also provided with a personal workspace that can only be accessed, heard and seen by the owner. Note when being inside personal spaces, users can still hear and see what's happening in the public space, Figure 5.3 (middle) shows the setting from user B's view.

Condition 3 - **P**ublic space + Publicly **V**isible personal spaces (C_{PV}): in addition to the public space (in C_P), each user is also provided with a personal workspace that is visible to their coworkers.

5.2.2 Dependent Variables

To identify how users use the space and the effect of personal spaces, series of dependent variables were developed. The research question of this chapter can be thought of as an two-fold question: (i) how different space configurations can impact users' collaboration; and (ii) how participants use the spaces during collaboration. As such, the the three themes (awareness, communication and performance measures) for measuring collaboration were used, and some other metrics based on system logged data were developed and used to form the fourth theme (space and territory measures) to measure how participants use the space and moved in the the space. These measures were grouped into questionnaire

and activities assessments. Next, the contents of participants reports and the system logging will be detailed following the themes respectively.

Participants Reports

We developed questionnaires to identify participants' subjective assessment of the conditions and their experience of the collaboration. Awareness and Sense of presence are closely related in virtual reality. The Igroup Presence Questionnaire (IPQ) developed by Schubert et al. (2001) was used to inform the design of questions about awareness, specifically including self-awareness, awareness of others, and awareness of others' activities. Questions regarding communication and performance were adapted from the Mutual Engagement Questionnaire (MEQ, Bryan-Kinns 2013). The rest of the questions were designed to question people's preference for conditions. The questionnaire included questions on:

(1) Awareness measures: i) Sense of self-presence, ii) sense of co-worker's presence and iii) sense of collaborator's activities.

(2) Communication measures: quality of communication, which may vary as the visibility of spaces can possibly affect the embodiment and nonverbal communication.

(3) Performance measures. Performance measure here including two parts: i) content generation, including the feeling of self's contribution and the feeling of others' contribution; ii) Satisfaction of the outcome: the satisfaction of the final music created reflects the quality of collaboration, cf. (Bryan-Kinns & Hamilton, 2012; Bryan-Kinns, 2013).

(4) Other measures: Preference: preference of the conditions, to see if users have an overall subjective preferences towards the settings.

Both post-session questionnaire and comparison questionnaire were applied to cover these measures. Specifically, the post-session questionnaire (PSQ) shown in Table 5.1, was developed to be filled after participants experiencing each condition to collect participants' feedback of each condition before it fading out, and the comparison questionnaire (CQ, Table 5.2) was developed to be filled at the end of the experiment to ask participants to compare conditions. PSQ also contains three open-ended questions to collect more in-depth subjective experience, see Table 5.4.

Activity Assessments

Most measures of activity assessments were developed to record and explore how participants utilise the space and form territory during the collaboration. However some of them are measuring the attention paid between collaborators and content generation. They covered the following 3 themes:

Table 5.1: Results of Post-Session Questionnaire[a] and results of Wilcoxon Rank Sum Test (two-tailed).

Measure	Question	C_P	C_{PI}	C_{PV}	C_P *vs* C_{PI}	C_P *vs* C_{PV}	C_{PI} *vs* C_{PV}
		M (SD)	*M (SD)*	*M (SD)*	*p (W)*	*p (W)*	*p (W)*
Awareness Measures	PSQ1 [b] (sense of self-presence) - In the virtual world, I had a sense of "being there"	8.17	8.83	8.55	0.06046	0.3014	0.3206
		(1.78)	(1.45)	(1.58)	(679.5)	(769.5)	(989)
	PSQ2 (sense of collaborator's presence) - My collaborator was there, collaborating with me together, all the time	8.02	8.26	7.95	0.4742	0.8159	0.3473
		(1.83)	(1.86)	(1.81)	(803.5)	(908)	(985)
	PSQ3 (sense of collaborator's activities) - I had a clear sense of what my collaborator was doing	6.88	7.10	7.24	0.6642	0.4606	0.7236
		(2.36)	(2.16)	(2.36)	(833.5)	(800)	(842.5)
Communication Measures	PSQ4 (communication quality) - How would you rate the quality of communication during the session	7.07	8.02	7.83	0.04531	0.1259	0.6553
		(2.25)	(1.87)	(1.99)	(661)	(713)	(931.5)
Performance Measures	PSQ5 (performance assessment) - How satisfied are you with the final piece of music created in this session	7.21	7.93	7.88	0.07462	0.1034	0.8158
		(1.91)	(1.39)	(1.48)	(686.5)	(702)	(908)
	PSQ6 (amount of contribution) - The amount of your contribution to the joint piece of music	6.38	6.76	6.93	0.1866	0.1037	0.7499
		(1.86)	(1.85)	(1.81)	(736)	(702.5)	(846.5)
	PSQ7 (amount of contribution) - The amount of your collaborator's contribution to the joint piece of music	7.14	7.19	7.21	0.6330	0.8339	0.8839
		(1.72)	(1.84)	(1.69)	(829)	(858.5)	(898.5)
	PSQ8 (quality of contribution) - The quality of your contribution to the joint piece	6.69	7.10	7.00	0.3753	0.5285	0.834
		(2.23)	(1.76)	(1.81)	(784)	(812)	(905.5)
	PSQ9 (quality of contribution) - The quality of your collaborator's contribution to the joint piece	7.31	7.57	7.64	0.4967	0.4538	0.9344
		(1.85)	(1.55)	(1.51)	(807)	(799.5)	(872.5)
Other Measures	PSQ10 (preference) - The addition of personal spaces in this is very helpful to the task	-	6.19	6.88	-	-	0.2666
		(-)	(2.67)	(2.33)	(-)	(-)	()758.5

[a] With 10-point-Likert scale, 1 indicates no fulfilment at all with the description of the questionnaire and 10 indicates a full fulfilment.

[b] The index numbers are for reference purpose, not reflecting the order of the questions in the questionnaire used in the study.

(1) Space and territory measures: i) number of uses of personal space, time length of using personal space, and average duration of each use of personal space; ii) distribution of participants' location during the session; iii) the sizes of personal and group territory if they emerge; iv) distribution of participants' interactions; v) colocation (cf. Bryan-Kinns 2013).

(2) Communication measures: Attention: i) time length of participants' paid attention toward each other's location; ii) number of times of participants paying attention to each other's location. [2]

(3) Performance measures: Contribution: i) number of interface additions; ii) number of note edits.

5.2.3 Participants

Students at the author's university were recruited through group emails. Each participant was compensated 10 GBP for their time. Twenty-one pairs of participants took part[3]. Twenty-five males, 17 females, aged from 22 to 42 ($M = 29$, $SD = 4.2$), 11.9% had no VR experience before, 16.7% tried VR once, 59.5% played 2-5 times and the rest 11.0% played more than 5 times or frequently. Half (21 participants) played a musical instrument and average experience of composing collaboratively is 2.6 in a 10-point Likert scale (1 for no experience and 10 for highly extensive experience). Slightly more than half (52.4%) knew their experiment partner very well prior to the study, a third rated (33.3%) their partner as stranger, the rest, nearly a sixth (14.3%) met their partner before but not know well.

5.2.4 Procedure

After reading and signing the information form and consent form, each pair of participants first received an explanation of music interface of LeMo II (see Figure 3.5). Then one experimenter demonstrated all the types of interaction gestures supported in LeMo II. By linking the demonstration with the first-person view shown on monitors, participants understood and grasped the tricks of how to play LeMo II. Next, participants were asked to try all the gestures during a 5 - 15 minutes trail, the trial ended once they were confident enough of all the gestures. The time length of tutorial session is flexible to ensure participants with diverse musical knowledge could grasp LeMo II. Participants were then asked to have three sessions of composing music together, each last-

[2]It should be noted that as no eye tracking device was applied in LeMos, LeMos identifies the attention by checking collisions between users' cone of visions and their collaborator's perspectives.

[3]The Queen Mary Research Ethics Committee granted ethical approval to carry out the study within its facilities (Ethical Application Ref: QMREC1872).

ing 8 minutes. To avoid the impact of adding of personal spaces and have a pure observation on how participants form their own proximity in the public space, all pairs started with C_P, and then C_{PI} and C_{PV} in a random sequence. Each session ended with a short questionnaire (Q1-Q10 of PSQ). The comparison questionnaire (CQ) and a short interview were carried out at the end of experiment.

5.3 Results

Next, the results of participant reports (PSQ + CQ), and Activity Assessments will be detailed respectively.

5.3.1 Participant Reports

To compare the conditions against each other, Wilcoxon Rank Sum Tests were run on the post-session questionnaire data, and Binomial Tests were run on the comparison questionnaire data. No significant difference was found in the former tests (see Figure 5.4), and several differences were found in the latter tests (Table 5.2). Next, these results will be detailed.

Awareness Measures

i) **Self-presence.** In all conditions, participants reported a high level of sense of self-presence (C_P: $M = 8.17$, $SD = 1.78$; C_{PI}: $M = 8.83$, $SD = 1.45$; C_{PV}: $M = 8.55$, $SD = 1.58$, in a 10-point Likert Scale), indicating a proper level of immersion, which forms a solid base for this VR study. ii) **Sense of collaborator's presence -** Participants reported high ratings on sense of collaborator's presence in all conditions (all $M > 7.95$). In comparison questionnaire, a significant portion (20 out of 42) of participants reported in C_{PI} they had the least strong sense of their collaborators' presence, see CQ2 of Table 5.2. iii) **Sense of collaborator's activities -** Participants reported a proper level of sense of collaborator's activities in all conditions (C_P: $M = 6.9$; C_{PI}: $M = 7.1$; C_{PV}: $M = 7.2$), no significant difference was found. However, in comparison questionnaire, significantly many (20 out of 42) of participants rated C_P as the session in which they felt least difficult to track collaborator's activities (Binomial Test, $0.48 > 0.33$, $p = 0.0384$, 1-sided[4]), and significantly few (7 out of 42) participants rated

[4] 1-sided test was used instead of 2-sided test, this is because the proportion 0.48 (20 out 48) is bigger than the presumed 0.33 (1/3). Hence we are testing whether the conditions are so different, resulting in having significant more participants than expected choosing the condition. In comparison, 2-side test only suits when the hypothesis does not have a direction, i.e. when p equals the presumed p ($1/3 = 0.33$), and k equals the presumed $k52/3 = 14$).

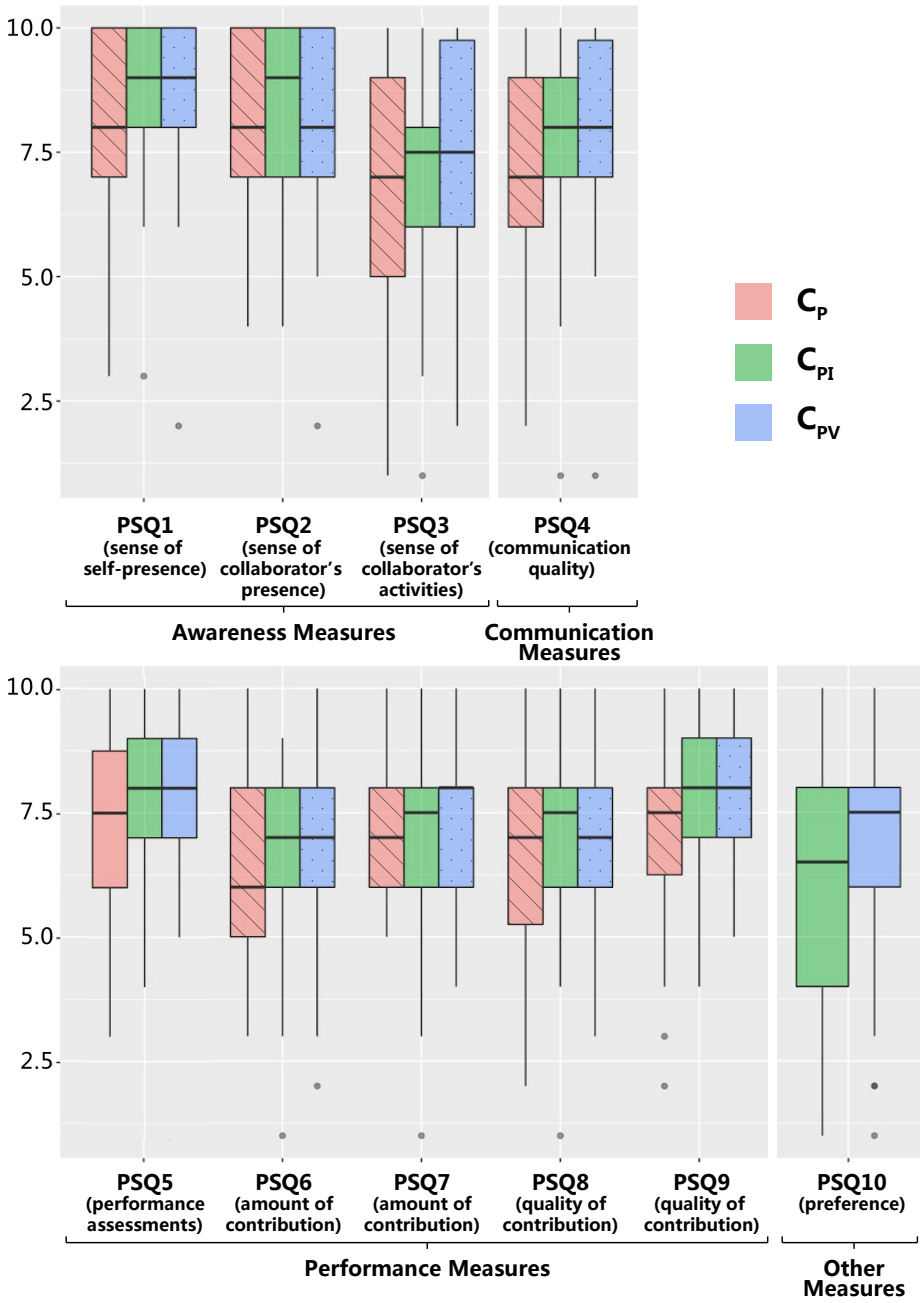


Figure 5.4: Box-plots of ratings gathered from Post-Session Questionnaire (PSQ).

Table 5.2: Results of Binomial Test of Comparison Questionnaire (CQ)[b]

Measure	Question description	Option	C_P		C_{PI}		C_{PV}	
			k	p	k	p	k	p
Awareness Measures	CQ1 [b]- Which session did you find most difficult to track collaborator's activities							
		hardest	14	1.00	19	0.0726	9	0.0667
		second hardest	8	0.0317	16	0.307	18	0.127
		least hardest	20	0.0384	7	0.0132	15	0.428
	CQ2 - Which session did you have the strongest sense that your collaborator was there working with you together?							
		strongest	18	0.127	11	0.209	13	0.442
		second strongest	11	0.209	11	0.209	20	0.0384
		least strongest	13	0.442	20	0.0384	9	0.0667
Communication Measures	CQ3 - Which session did you have the best quality of communication							
		best	13	0.442	13	0.442	16	0.307
		second best	10	0.125	17	0.205	15	0.428
		least best	19	0.0726	12	0.317	11	0.209
Performance Measures	CQ4 - In which session, you made the music you were most satisfied with?							
		most satisfied	8	0.0317	16	0.307	18	0.127
		second most satisfied	12	0.317	14	1.00	16	0.307
		least satisfied	22	0.00835	12	0.317	8	0.0317
	CQ5 - Which session do you feel you made the most contribution							
		the most	7	0.0132	13	0.442	22	0.00835
		second most	12	0.317	18	0.127	12	0.317
		least	22	0.00835	11	0.209	9	0.0667
	CQ6 - Which session do you feel your collaborator made the most contribution							
		the most	11	0.209	16	0.307	15	0.428
		second most	14	1.00	15	0.428	13	0.442
		least	17	0.205	11	0.209	14	1.00
Other Measures	CQ7 - Which session had the best setting for creating music collaboratively							
		best	10	0.125	11	0.209	21	0.0187
		second best	16	0.307	13	0.442	13	0.442
		least best	16	0.307	18	0.307	8	0.0317
	CQ8 - Which session did you find most difficult to cooperate with collaborator							
		most difficult	15	0.428	17	0.205	10	0.125
		2nd most difficult	13	0.442	16	0.307	13	0.442
		least difficult	14	1.00	9	0.0667	19	0.0726

[a] Lower-tailed test when k<14, two-tailed test when k=14, upper-tailed test when k>14.
[b] The index numbers in this table are only for reference purpose, not reflecting the real order of the items listed in the questionnaire used in the study.

C_{PI} as the session in which it is the least difficult to track collaborator's activity (Binomial Test, $0.17 < 0.33$, $p = 0.0132$, 1-sided).

Communication Measures

Participants reported a good quality of communication in all conditions (C_P: $M = 7.07$, $SD = 2.25$; C_{PI}: $M = 8.02$, $SD = 1.86$; C_{PV}: $M = 7.83$, $SD = 1.99$). No significant difference was found.

Performance Measures

When asked their satisfaction with the output in all conditions, participant seemed to be more satisfied with the music piece they produced in C_{PI} and C_{PV} compared with the one made in C_P (C_P: $M = 7.21$, $SD = 1.91$; C_{PI}: $M = 7.92$, $SD = 1.38$; C_{PV}: $M = 7.88$, $SD = 1.48$), Wilcoxon Rank Sum test shows a trend towards a significance (C_{PI} vs C_P: $W = 1077.5$, $p = 0.07462$; C_{PV} vs C_P $W = 1062$, $p = 0.1034$). In comparison questionnaire, significantly few participants (8 out of 42) believed it was in C_P that they made the most satisfied music ($0.19 < 0.33$, $p = 0.0317$, 1-sided), and significantly many participants (22 out of 42) chose C_P as the session in which they made the least satisfied music ($0.52 > 0.33$, $p = 0.00835$, 1-sided), and significantly few participants (8 out of 42) believed it was in C_{PV} that they produced the least satisfied music ($0.19 < 0.33$, $p = 0.0317$, 1-sided), see results in Table 5.2.

Significantly many participants (22 out of 42) reported they had the least sense of self-contribution in C_P ($0.52 > 0.33$, $p = 0.00835$, 1-sided) and a significant small number of participants believed in C_P they had the strongest sense of self-contribution ($0.17 < 0.33$, $p = 0.0132$, 1-sided). By contrast, a significant proportion of participants (22 out of 42) reported they had the strongest sense of self-contribution in C_{PV} ($0.52 > 0.33$, $p = 0.00835$, 1-sided). No significant difference was found in terms of the sense of co-workers' contribution.

5.3.2 Other Measures

A significantly big portion of participants (37 out of 42) reported that there were differences between these conditions ($0.88 > 0.5$, $p = 4.434e\text{-}07$, 1-sided). Participants felt positively to the addition of both C_{PI} ($M = 6.19$, $SD = 2.67$) and C_{PV} ($M = 6.88$, $SD = 2.33$), though no significant difference was found between the two conditions. In comparison questionnaire, C_{PV} was believed to have the best setting by significantly many participants (21 out of 42; $0.5 > 0.33$, $p = 0.0187$, 1-sided), and a significantly small number of participants (8 out of 42) rated C_{PV} as least best setting ($0.19 < 0.33$, $p = 0.0317$, 1-sided).

Table 5.3: Results of activity assessments and results of Wilcoxon Rank Sum Test.

Measure	Description (Unit)	C_P		C_{PI}		C_{PV}		C_P *vs* C_{PI}		C_P *vs* C_{PV}		C_{PI} *vs* C_{PV}	
		M	*SD*	*M*	*SD*	*M*	*SD*	*p*	*W*	*p*	*W*	*p*	*W*
Communication Measures	AA1 - Time length of attention on collaborator (second)	143.05	118.33	52.53	62.03	98.25	60.76	0.000147	888	0.248	673	0.00103	310
	AA2 - Times of drawing attention on collaborator (-)	47.29	25.91	15.82	11.80	30.59	14.98	3.52e-07	993.5	0.00259	824	5.70e-05	249.5
Performance Measures	AA3 - No. of note edits done in group territory (-)	44.71	69.88	12.29	16.11	7.17	10.54	0.0696	197	0.0101	218	0.373	169.5
	AA4 - No. of note edits done in personal territory (-)	39.94	38.53	74.12	49.13	86.12	55.23	0.037	83.5	0.00925	68.5	0.593	128.5
	AA5 - No. of note edits done in other's territory (-)	29.24	42.68	20.71	25.17	25.24	46.04	0.444	167	0.626	159	0.902	140.5
	AA6 - No. of music interface additions (-)	4.5	2.43	3.85	1.76	3.79	1.95	0.293	663	0.215	678	0.871	591.5
	AA7 - No. of music interface additions in personal space (one)	-	-	2.18	1.47	2.06	1.48	-	-	-	-	0.743	604.5
	AA8 - No. of music interface additions in public space (-)	4.5	2.43	1.68	1.37	1.74	1.76	1.52e-07	1001	5.93e-07	981.5	0.874	591
	AA9 - No. of note edits (-)	127.82	63.19	120.74	49.56	129.56	51.39	0.825	596.5	0.754	552	0.540	527.5
	AA10 - No. of note edits done in public space (-)	127.82	63.19	41.44	31.60	44.21	30.07	6.54e-09	1051.5	2.75e-08	1031.5	1.00	577.5
	AA11 - No. of note edits done in personal space (-)	-	-	79.29	49.08	85.5	51.64	-	-	-	-	0.936	571
Space and Territory Measures	AA12 - No. of uses of personal spaces (-)	-	-	4.5	2.36	4.18	4.83	-	-	-	-	0.0316	751.5
	AA13 - Time length of using personal spaces (second)	-	-	218.96	92.35	214.84	81.03	-	-	-	-	0.841	561
	AA14 - Average duration of each entry of personal space (second)	-	-	68.02	56.33	130.79	97.87	-	-	-	-	0.0181	829
	AA15 - Size of group territory (m^2)	0.90	0.81	0.44	0.40	0.41	0.46	0.0447	203	0.00251	210	0.654	158
	AA16 - Size of personal territory (m^2)	1.41	0.53	2.88	0.69	2.62	0.80	2.64e-14	43	1.05e-09	80	0.3	663
	AA17 - Size of personal territory fallen in public space (m^2)	1.41	0.53	1.44	0.61	1.39	0.49	0.410	510	0.849	594	0.324	659
	AA18 - Size of personal territory fallen in personal space (m^2)	-	-	1.44	0.30	1.24	0.45	-	-	-	-	0.0411	745
	AA19 - Colocation (metre)	0.99	0.31	2.31	0.52	2.29	0.47	8.57e-10	0	3.43e-09	2	0.760	135

Table 5.4: Results of Open-Ended Questions

Participant No. - Transcription (grouped according to 3 themes)
Advantages of providing personal spaces:
P_{1A} - [I could] try the effect of different settings without interrupting my collaborator.
P_{2A} - Had private space to test ideas in [C_{PI} and C_{PV}].
P_{5A} - [C_P] is good for people to work together, but without private space, I feel difficult to create my own idea.
P_{6A} - The [personal space] is nice to have. We prepared private rhythms and melodies and then showed them to each other.
P_{9B} - [Having] own space...allowed us to compile our own piece of music in comparison to joint and single space in session 1. This enabled us to work on our separate compilation and merge to create the final piece, which was good!
P_{14B} - Very good to focus on one thing before sharing.
P_{15A} - I like doing something together but I also need my own work space.
Disadvantages of providing personal spaces:
P_{1B} - For me it was same since I did not use the private space.
Preference on publicly visible/invisible personal space:
P_{3A} - I was able to see my collaborator in...session C_{PV}.
P_{5A} - [C_{PI}] was too private, I cannot see my partner's job. It feels not comfor-, table, [C_{PV}] was fine...it provides both privacy and teamwork equally.
P_{5B} - I feel more communicated with my collaborator during [C_P] and [C_{PV}].
7A - Session 3 and [C_{PV}] is the most helpful one.
P_{8B} - We can see what my collaborator is doing [in C_{PV}].
P_{9A} - [C_{PV} is] very distracting and prefer audio communication only.
P_{10A} - The way we can see each other but not hear from each other [C_{PV}] is most efficient.
P_{17A} - In [C_{PI}] we definitely felt more isolated from each other, but I am kind of used to this...on in isolation.
P_{18A} - [C_{PI}] provides good private and public space, which allows us to work individually or cooperatively...In [C_{PI}], I had the weakest sense of communication because of the private space being opaque.
P_{19B} - [C_{PI}] help[ed] me to think on my own, without too much disturbance...

5.3.3 Activity Assessments

In this section we report on measures focusing on the participants' interactive activities, see Table 5.3. Wilcoxon Rank Sum tests were run to compare conditions against each other. Note that 17 out of 21 groups used personal spaces in both C_{PI} and C_{PV}, however, there were 4 groups, one member of which did not use personal spaces one of these conditions. As this section focuses on investigating how participants' activity might change due to the use of personal space, the data of these 4 groups were abandoned in this section. Activity Assessments does not have any items that assessing the awareness, and assessing awareness has been covered by self-reports. Next, the results of AA will be reported from three aspects: (i) communication measure, (ii) performance measure and (iii) space and territory usage.

Communication Measures

Activity Assessments has two items related to communication, specifically, attention paid to collaboration, see AA1 and AA2 in 5.3. Throughout the 480-second session, participants had their attention toward their collaborators' location for 143.05 seconds in C_P and 98.25 seconds in C_{PV}, both of which are significantly longer than that of C_{PI} (52.53 seconds). Participants oriented their attention toward their collaborator significantly more times in C_P ($M = 47.29$), compared with that of C_{PI} ($M = 15.82$) and C_{PV} ($M = 30.59$; both p <0.01). Although they did significantly more times in C_{PV} than C_{PI} (Wilcoxon Rank Sum Test, $W = 249.5$ $p = 5.70e\text{-}05$).

Performance Measures

(1) Number of music interface additions. Participants generated on average 4.5 interfaces in C_P, 3.85 in C_{PI}, 3.79 in C_{PV} on average, no significant difference was found. In C_{PI} and C_{PV}, the personal spaces were available, participants had some of interface additions done in personal space. Specifically, in C_{PI}, out of the 4.5 interfaces, they added 1.68 in public space, and 2.18 in personal space; In C_{PV} 1.74 were added in public space and 2.06 in personal space, no significant difference was found between these two conditions or between the public space and personal space in each condition, see more detailed statistics in Table 5.3. (2) Number of note edits. No significant difference was found in terms of the sum number of note edits (C_P: $M = 127.82$; C_{PI}: $M = 120.74$; C_{PV}: $M = 129.56$). However, when classified by types of spaces, participants had significantly more note edits in personal spaces (C_{PI}: $M = 79.29$; C_{PV}: $M = 85.5$) than they did in public space (C_{PI}: $M = 41.44$; C_{PV}: $M = 44.21$; Wilcoxon Rank Sum Tests, both $p < 0.001$, see detailed statistic in Table 5.3).

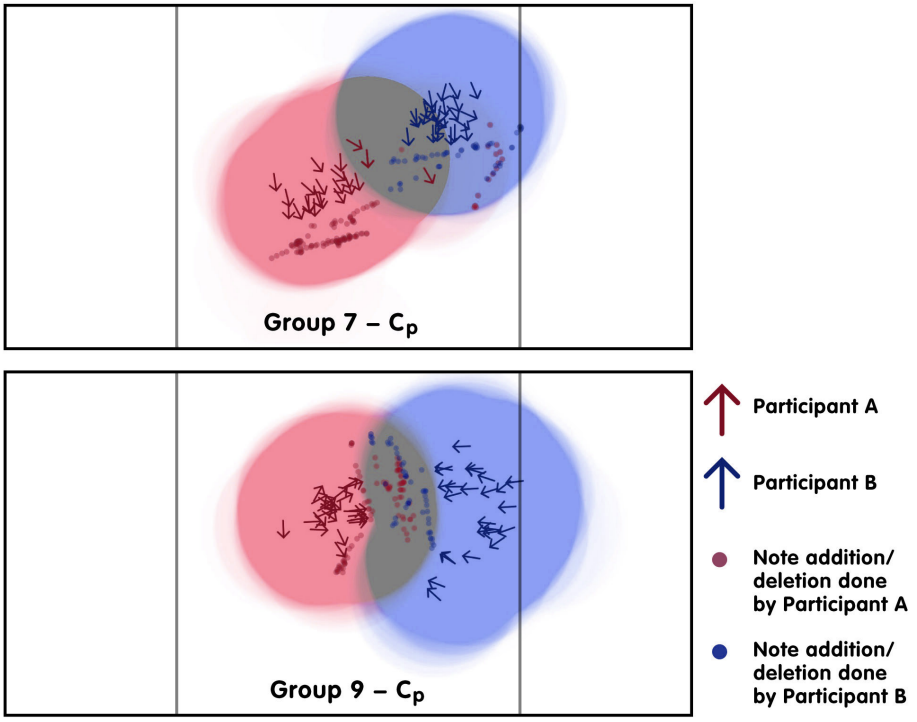


Figure 5.5: Visual trace of the participants' positions, directions and musical note edits (group7 and group 9 in C_P).

Considering participants spent significantly shorter time in personal spaces than in public space, this means participants were more productive in making notes edits in personal spaces.

Space and Territory Usage

In the 8-minute (480-seconds) session, although there is no significant difference in terms of the time length of using personal spaces (218.96 seconds in C_{PI} and 214.84 seconds in C_{PV}), participants did enter personal space significantly more times in C_{PV} ($M = 4.5$) compared with C_{PI} ($M = 4.18$; Wilcoxon Rank Sum Test, $W = 751.5$, $p = 0.0316$). This results in a significant shorter stay of each entry in C_{PI} ($M = 68.02$) compared with C_{PV} (M=130.79; $W = 327$, $p = 0.001807$). Although participants spent a certain amount of time in personal space, they still stayed significantly longer in public space (261.04 in C_{PI} and 265.16 in C_{PV}; C_{PI}: W = 769, p = 0.01878; C_{PV}: W = 780, p = 0.01279).

(1) Formation of territory To illustrate how participants use the space, their locations, directions and musical note edits were plotted on a top view of the stage, see Figure 5.5, we call these plots visual traces. Specifically, the arrows

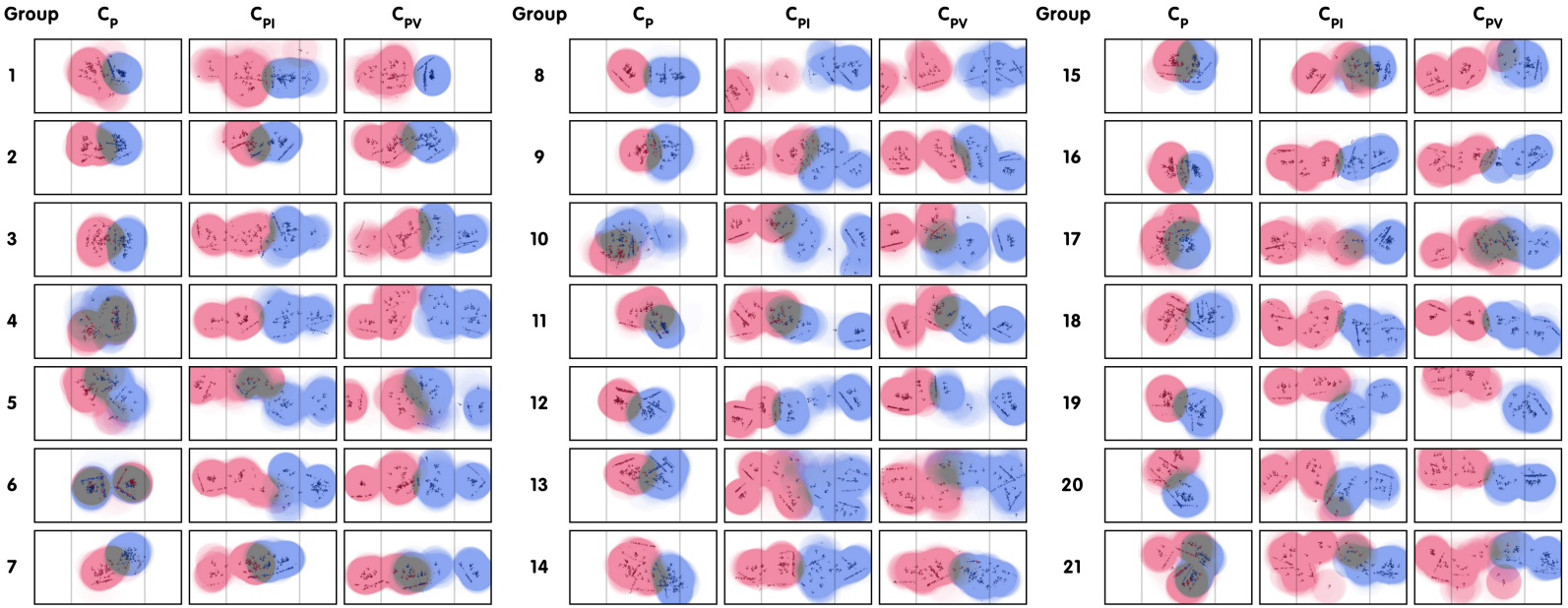


Figure 5.6: Visual traces of participants' locations and interactions.

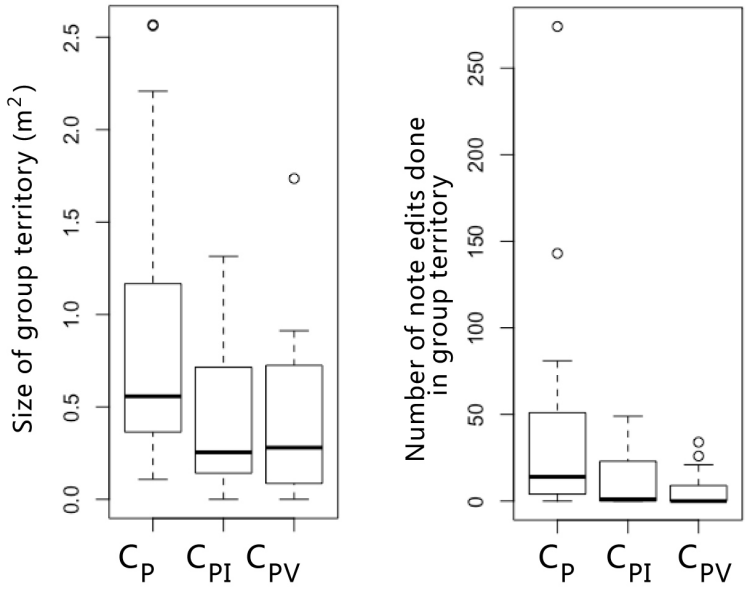


Figure 5.7: Box-plot: size of group territory (left); number of note edits done in group territory (right).

were participants' locations at 20-second interval for ease of reading the diagram, and dots are the position of participants hand when editing note on music interface. Personal territory was defined as a workspace close to the person and group territory was defined the central area or spaces between collaborators (Scott et al., 2004; Xambó et al., 2013; Scott & Carpendale, 2010), the area within a 0.6 metre radius of the participants' locations (locations here are at 1-second interval) were dyed with different tint colours (green for participant A personal territory, and red for B). We chose 0.6 metre as this also falls into the range of close phase of personal distance and it permits one participant to touch each other or the same music interface (Hall, 1966), most of the participant's music edits also fell inside this range.

As tint colour applied, the greener/redder the colour is, the more presence the corresponding participant shown in that location. The overlap is dyed with grey, indicating appearances of both participants. Figure 5.6 shows the visual traces of all the groups, it can been seen that in C_P, apart from three groups (Group 4, 6 and 21), a significant proportion of groups (18 out of 21) developed fixed personal territory (green area for A and red area for B) and group territory (Binomial Test, $0.86 > 0.5$, $p = 0.00074$, 1-sided). The visual traces also proved a significant proportion of groups (17 out of 21) did use personal spaces when available ($0.81 > 0.5$, $p = 0.003599$, 1-sided).

(2) Sizes of territory and edits done classified by territories. As shown in

Table 5.3 and Figure 5.7 (left), participants formed a significant larger size group territory in C_P ($M = 0.90$ m^2), than C_{PI} ($M = 0.44$ m^2) and C_{PV} ($M = 0.41$ m^2), an inspection of the visual traces (Figure 5.6) might also verify these results. Participants had an average 44.71 group edits (edits done in group territory) in C_P, which is significantly more than that of C_{PV} ($M = 7.17$; $W = 218$, $p = 0.0101$), and a near-marginal significantly more than that of C_{PI} (12.29, $W = 197$, $p = 0.0696$), also see Figure 5.7 (right).

When personal spaces available, participants had significantly more music edits done in personal territory (both $p < 0.05$, more details in Table 5.3) and formed a significant larger size of personal territories in C_{PI} and C_{PV} compared with C_P (both $p < 0.001$, more details in Table 5.3). However, if we deduct the part they formed inside personal spaces, there is no significant difference. We also found although the size of personal territory fallen in personal space in C_{PV} is significantly smaller than that of C_{PV} (C_{PI}: $M = 1.44$; $M = 1.24$ in C_{PV}; $W = 745$, $p = 0.00411$), the amount of music edits done in personal territories inside personal space in C_{PI} ($M = 74.12$) is similar to that of C_{PV} ($M = 86.12$).

(3) Colocation. No significant difference was found between the colocation of C_{PI} ($M = 2.3$ m, $SD = 0.52$) and C_{PV} ($M = 2.3$ m, $SD = 0.47$). However, compared with C_P (M = 0.99 m, SD = 0.31), the average distance between participants (colocation) is significantly larger in C_{PI} and C_{PV} (C_P vs C_{PI}: Wilcoxon Rank Sum Test, $W = 0$, $p = 8.57\text{e-}10$; C_P vs C_{PV}: $W = 2$, $p = 3.428\text{e-}09$). The main reason can be that personal spaces are on the opposite sides of the public space, and participants stayed away from each other when working at personal spaces.

5.3.4 Interviews

More insights about the spatial configurations and how they affected the collaboration were revealed from the interview. The interviews audio were firstly transcribed into around 18000 words and a thematic analysis was undertaken. See more about thematic analysis research method in Braun & Clarke (2006); Yin (2017), section 4.3.10. Note: the interviews of 7 groups (group 7, 8, 10, 11, 13, 15, 19) were carried out in Chinese, audio recordings of these groups were firstly transcribed and then translated to English. Different from self-report and activity assessments, where metrics can be grouped into predefined groups, the process of interviews is less structured, hence here we report the interview results according to the themes emerged from the thematic analysis. In total, 41 codes and 4 overarching themes emerged from the thematic analysis, see Figure 5.8. These 4 themes are: (i) reporting learning effects; ii) reporting music

quality; (iii) reporting conditions; (iv) reporting LeMo II system. All the codes and coded segments are available in Appendix B.6. Next the four themes will be reported one by one.

Reporting Learning Effects

Seventeen coded segments contributed by 15 participants[5] from 11 groups are reporting effect of the session sequence. (i) Feeling of wonder - P_{3A} reported feeling "very excited" in the first session as this was the first time of experiencing such a thing. (ii) Learning effect - With limited experience learnt in the tutorial, 4 participants (P_{9A}, P_{3A}, P_{9B}, P_{12A}) reported a feeling of difficulty in the first session. For example, P_{9B} (Participant 9B) reported:

> *...Initially when...you're asking people to do experiment so you put...them in the same space and [say] okay, there's things [for you to do], so it makes it quite difficult...*

With progression of each session, participants reported they got more experienced, better at how to interact within the system (P_{1B}, P_{4A}, P_{8B}, P_{10A}). Therefore, in later sessions, they felt more confident, comfortable and know better what to do (P_{17A}, P_{18B}, P_{19A}), they felt more unrestrained and were able to create things more freely (P_{13A}, P_{13B}) and could even "muck around a bit more" (P_{17A}). As a result, the second and third sessions turned out to be much better (P_{9B}).

Reporting Music Quality

Fifteen coded segments from 10 participants are reporting the quality of music outpu. Six participants reported they produced the best music in C_{PI} and 5 reported this was produced in C_{PV} while no participants reported C_P generated the best music. On the contrary, 3 participants believed they made the worst music in C_P whilst no participants reported this happened in C_{PI} and only 1 reported it was made in C_{PV}. This is probably due the learning effect, as the sequence of sessions was not fully-randomised. Namely, C_P always came the first for all the groups and participants were less experienced in that condition compared with the other two conditions. Note that two participants reported the quality of productions was not necessarily closely related to the configurations, luck might also be an important factor (P_{13A}, P_{15A}).

[5]Specifically, they are Participant B in group 1 (referred to as P_{1B}), P_{3A}, P_{3B}, P_{4A}, P_{4B}, P_{8B}, P_{9A}, P_{9B}, P_{10A}, P_{12A}, P_{12B}, P_{13A}, P_{13B}, P_{17A}, P_{18B}, P_{19A} .

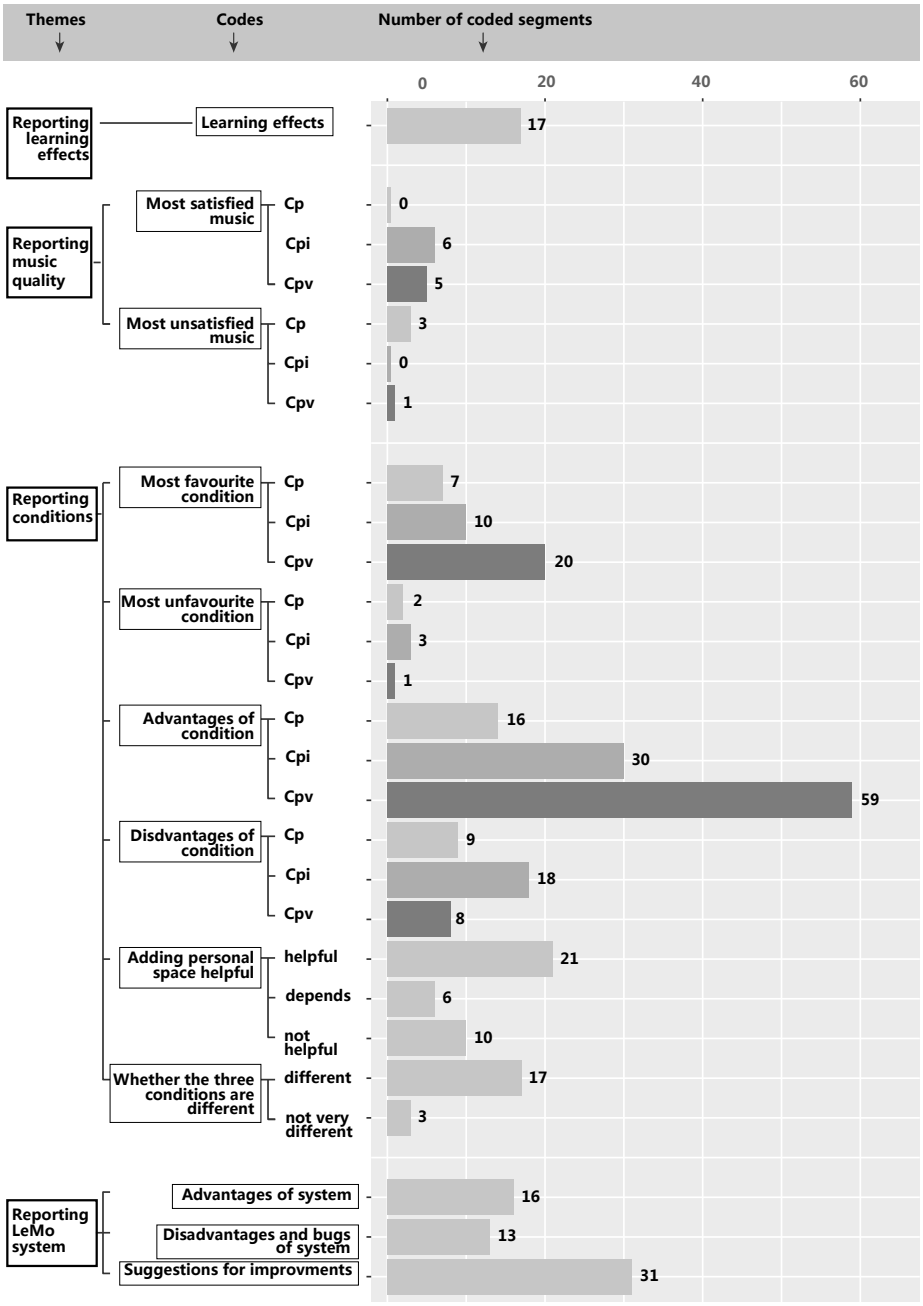


Figure 5.8: Ingredients of the thematic analysis of the interview, in total, there are 4 themes, 12 codes, and 335 coded segments.

Reporting Conditions

There are 256 coded segments reporting conditions, covering six codes: (i) helpfulness of adding personal space; (ii) feeling that the three differences are different; (iii) most favourite condition; (iv) least favourite condition; (v) advantages of condition; (vi) disadvantages of condition. Next, results of these codes will be presented in detail.

Whether conditions are different. There are 17 coded segments reporting that participants did feel the three conditions were different, 3 coded-segments reporting that the conditions were not very different, and no coded-segments reporting the conditions were the same. More details behind the differences will be revealed in the following subsections briefing the advantages and disadvantages of the conditions.

Whether adding personal space is helpful. There are 21 coded segments reporting it was helpful to add personal space, outnumbering those reporting it depends (6 coded segments), and those reporting unhelpful (10 coded segments). More reasons behind the helpfulness of personal space will be revealed in detail in the following subsections that reporting the advantages and disadvantages of the conditions.

Preference of conditions - Twenty coded segments contributed by 16 participants from 13 groups reported C_{PV} to be the best setting, by comparison, C_P was reported to be the best by 7 coded segments contributed by 7 participants from 6 groups, and C_{PI} was reported to be the best by 10 coded segments contributed by 10 participants from 8 groups. To summarise, more participants and more coded segments are reporting C_{PV} to be the best setting, indicating a preference on C_{PV}. Six coded segments from 6 participants were reporting the worst setting, 2 participants reported C_P to be the worst, 3 believed it was C_{PI} and 1 believed it was C_{PV}. However, all these numbers are relatively small, and cannot be used to indicate meaningful findings.

Advantages of C_P - Staying Together - Sixteen coded segments are regarding C_P's advantages. C_P has advantages in terms of getting "constant feedback" (Participant A in group 2) and "quick feedback" (11_B), feeling of working "together" (P_{3A}, P_{8B},P_{16A}) and less feeling of loneliness (P_{11B}). Participants were able "to see what [their collaborator] was doing instantly and do the discussion directly" (P_{11B}), it was "easier to negotiate" (P_{5B}) and "better understanding" of each other's intentions (P_{5A}), allowing participants to "both see and hear" (P_{8B}), the communication was "much more" and "better" (P_{21B}). With vague ownership of music interfaces, compared with C_{PI} and C_{PV}, P_{10A} felt "least stressful in [C_P], because [they] did not have anything to be responsible for". These advantages are mainly the result of the fact that only one space

was provided and people had to work together and possibly closer.

Disadvantages of C_P - Messy - Nine coded segments are reporting C_P's disadvantages. P_{9B} reported "things did not work out very well" in the same "common space". The word "messy" appeared several times to describe situations in C_P. "[It] was messy, there were all kinds of sounds [messing up] (5A). "It will get messy [if people want to] try things in the public space...[because their] partner is trying things at the same time" (P_{9A}), "[it was] too messy, hers mixed with mine, too chaotic" (P_{15A}), "I felt it was too messy"(P_{17A}). Similarly, P_{10B} reported that in C_P, the inference between each other "affected the efficiency". Without providing personal space, P_{5A} reported they could "only passively participate" as their partner P_{5B} played a "dominant" role in the creation. P_{9A} reported C_P to be"the least collaborative session", which is quite different from other participants, recall that three participants reported a feeling of togetherness in C_P. P_{9A} reported C_P led less collaboration because they believed the addition of personal space required people to communicate more to keep the pair on the same page, which promoted the collaboration, and thence the promotion disappeared without providing personal space.

Advantages of C_{PI} - Thirty coded segments were reporting C_{PI}'s advantages. These comments reveal three major advantages of C_{PI}. **(i) Providing a chance to work independently without interfering each other.** This advantage is mainly a result of the addition of personal spaces, and hence is mostly shared with C_{PV} as well because C_{PV} also provides personal spaces. For instance, P_{2A} "felt more comfortable taking risks" when trying out the ideas in C_{PI} and C_{PV}. With personal space available, P_{4A} and P_{10A} were able to "work dependently" and P_{4A} could avoid "bother[ing] [their] collaborator about his work"any more. P_{9A} and P_{9B} reported they could avoid affecting each other. This became especially true when one participant was more dominant in the collaboration, e.g. while 5B being dominant in the collaboration, P_{5A} could only "passively participate in" C_P. After introducing personal spaces in C_{PI} and C_{PV}, P_{5A} could "complete some of [their] ideas in [their] space". **(ii) Less distraction and easier to understand compared with C_{PV}.** P_{10A} reported the limited visibility of each other's avatars helped them to "focus on creation" (P_{10A}), they became "very concentrated" (P_{11A}), C_{PI} avoided "any visual distractions" (P_{21A}), they didn't "need to distract to see what colour [their collaborator] was using" (P_{11B}). What they saw was what they could hear, "if you cannot see it you shouldn't hear it" (P_{4B}), this plain feature of C_{PI} made it "simpler" (P_{21A}) and "easier" (P_{21B}) to understand. **(iii) Limited visual cues force collaborators to communicate and collaborate more.** People might be thinking "that if you see each other it's easier, but it's not because [they] then didn't feel the need to talk...[whist they] didn't see that much,

so [they] had to communicate more." P_{12B}. Similarly, P_{21A} and P_{21B} reported "not being able to see each other forced [them] to talk more and resulted more communication and more collaboration.

Disadvantages of C_{PI} - Eighteen coded segments report C_{PI}'s disadvantages. The opaque wall blocked others seeing inside the participant's personal spaces caused the lack of visual cue, which led two negative impacts: **(i) Difficulty of tracking collaborator's activities.** P8B reported in C_{PI} they did "not know how many music pieces the other was doing and what types of instruments he was using", P_{15A} and P_{15B} reported C_{PI} was "too closed", they had no clue of what "instruments each other used and what each other chose". P_{5A} reported that "without seeing what my collaborator was doing at the opposite side, [they] did not know what she was doing, and then [they] felt less enthusiastic in creation". **(ii) Reduced sense of collaborator's presence.** The continuity of sense of collaborator's presence was interrupted. P_{1A} felt "a bit weird that some at some point, [their collaborator] was gone, behind the wall". Similarly, P_{5A} reported, while their collaborator was inside the personal space, they" seem to be working alone, and then suddenly there is a person out there". For P_{15A}, this invisibility influenced the sense of collaborator's presence. With less feeling of each other's presence, P_{10A} reported they just "[made] something carelessly with no cooperation".

Advantages of C_{PV} - There are 59 coded segments reporting C_{PV}'s advantages, more than the sum of what C_{PI} and C_P got.

> *I felt like for some reason, we balanced working publicly, collaboratively and privately the right amount or better at least than in the other situations.* (P_{17B})

> *From the perspective of creation, or team creation, [C_{PV}] is better. That is, everyone can have a development space to complete their own ideas fairly, and then put the ideas in the centre to compare whose piece is better or worse.* (P_{5A})

According to the interview results, **C_{PV}** has strengths in these 5 aspects: **(i) A space to develop ideas freely.** This advantage is mainly a result of adding personal space. This advantage is shared with C_{PI}, which also provides personal spaces. With the addition of personal space, participants reported feeling "more comfortable taking risks" (P_{2A}), by enabling them to "do on their own thing" (3A), they could avoid bothering their collaborator as they might create some messy sounds during the composing (P_{4B}, P_{9B},P_{11A}), "no one would disturb you, so you can focus on creating" (P_{10B}). The got a chance to complete "some of [their] ideas in [their] space". **(ii) Better sense of collaborator's activities.** Being able to constantly see and track collaborator's activities was also

reported to be beneficial. Participants reported they could "see what the other person was working" (P_{1B}), 7 participants (P_{3A}, P_{4A}, P_{5A}, P_{13A}, P_{15B}, P_{17B}, P_{20B}) reported similar views. They could see which instrument was being used (P_{2B}, P_{4A}, P_{7A}, P_{8B}, P_{11B}, P_{12B},P_{18A}, P_{20B}), having a rough sense of collaborator's production (P_{10B}, P_{11B}), so they can avoid repeating it (P_{4A}, P_{7A}). They could compare the quality of music even without hearing (P_{10B}), "have a better control about what instruments [they] should choose, more convenient for you to coordinate" (P_{13A}). "The visual feedback" was helpful to get back on track "in case somebody forgets and deviates from what has been decided"(P_{20B}). **(iii) Better sense of collaborator's presence.** Being able to see the to see the other person's avatar possibly led to a stronger sense of collaborator's presence. P_{13B} and P_{13A} reported a stronger "feeling of togetherness" compared with C_{PI}, 15B reported "being able to see is more intuitive". **(iv) Promoted communication.** P_{17A} reported they "communicated better in C_{PV}, definitely, because [they] had like a clear [feeling of] what was going on", P_{13B} reported they "could discuss" while inside personal space in C_{PV}. P_{20A} reported they could "talk about what [they] do and [they] gain and generate new ideas based on what [they are] doing, and the conversation pushes forward." **(v) Other points.** P_{2B} felt in C_{PV} "the space are larger", P_{15A} reported a stronger feeling of security and less feeling of uncertainty. P_{2A} reported less stressful to go into the personal space compared with C_{PI}.

Disadvantages of C_{PV} - Eight coded segments are reporting C_{PV}'s disadvantages: **(i) Distracting.** P_{9B} felt C_{PV} to be "slightly distracting" because they were trying to see each other's work and "unconsciously trying to adjust" the work, and P_{9A} reported they "couldn't help to look at what [P_{9B} was] doing and try to adjust it, so it makes sense with [P_{9B}'s] music". Similarly, P_{19B} reported being able to see each other's activity inside personal space made C_{PV} slightly and visually "messy". **(ii) Reduced communication.** P_{12B} reported the belief that "no need to communicate" as visual cue was enough to substitute the reduced communication, which was very important and should not be weakened. **(iii) Complicated to understand.** Opposite to the C_{PI}'s feature - what participants saw is what participants could hear - in C_{PV}, even participants could see the interface, participants possibly wouldn't be able to hear it, this was reported hard to understand by P_{4B}.

Reporting LeMo II System

Participants reported the advantages (16 coded segments) and disadvantages and bugs (13 coded segments), and suggestions for improvements (31 coded segments) with the design and technical part of LeMo II. Since these segments

are not directly related to the scope of this chapter or thesis, these segments will not be detailed here.

5.4 Discussion

Next, we will firstly discuss how collaboration was impacted from three perspectives: awareness, communication and performance, and then discuss the impact on the space and territory.

5.4.1 Impacts on Collaboration

Measures captured the impacts that spatial configurations pose on the collaboration from three perspectives: awareness, communication and performance. Next, they will be discussed respectively.

Awareness

While in C_P and C_{PV}, they felt *"more comfortable with the system. They can see what my collaborator is doing"*. C_{PV} "*provides both privacy and teamwork equally.*" By contrast, in C_{PI} One participant reported "*[he] definitely felt more isolated from each other, but [he was] kind of used to this in writing music which [he] then [gave] to other people to work on in isolation*". Another reported C_{PI} " helped" *[her] to think on [her] own, without too much disturbance from the other.* "The additional visual cues in C_{PV} can also result in an overloaded cognitive information. So as suggested by some participants, it would be good if people have a toggle to change the visibility.

Results also indicate that participants went to personal spaces significantly more times and stayed shorter for each entry in C_{PI} compared with C_{PV} (Table 5.3), the reason can be that the lack of visibility of personal spaces in C_{PI} made participants had to go back to public space more frequently to update each other's condition to form a proper level of shared knowledge of the group work. While in C_{PV}, the visibility ensured them to work a bit longer independently with a proper knowledge of what their collaborators were doing inside personal spaces.

Communication

The visibility difference between C_{PI} and C_{PV} also impacted the communication. Participants had significantly shorter length and fewer times of drawing attention toward their collaborator's location in C_{PI} compared with C_{PV} (AA1

and AA2). The comparison questionnaire results (CQ1 and CQ2), also indicates that invisible personal space C_{PI} C_{PI} led to a weaker sense of coworker's presence and activities. Besides this, participants also reported "C_{PI} *was too private, [they] cannot see [their] partner's job and [they] did not feel comfortable*".

By contrast, in C_{PV}, participants could see each other even when their collaborator being inside personal spaces, thus they felt "*more communicated*". In other words, C_{PV} delivered a better communication. As a possible result, C_{PV} is rated as the most attractive one CQ7 in Table 5.3). Participants reported C_{PV} to be a good space setting for working either "*individually or cooperatively*" and reported they had the "*weakest sense of communication with the partner*" and had to "*talk more*".

Performance

Performance was also influenced in two ways:

A More Efficient Workspace - Results of interview has shown, without personal spaces added, C_P was reported to be "messy" (P_{5A}, P_{9A}, P_{15A}, P_{17A}) and "chaotic" (P_{15A}). Participant 10B reported the inference between each other "affected the efficiency". Besides these, results of activity assessments show in C_{PI} and C_{PV}, when personal spaces available, participants had more note edits done in personal spaces although they spent significantly more time in public space, indicating that participants were more productive in personal spaces. One reason can possibly be that when being inside the public space, participants spent time on communication and discussion and in personal space, they focus more on producing. Another reason is that when participants were inside their personal spaces, "[they] could create something new without disturbing [their] collaborator, and vice versa", the disappearance of auditory disturbance and distraction made participants more focused on the development of ideas. Compared with C_{PV}, the invisibility of personal spaces in C_{PI} also removed the visual disturbance of co-worker's activities, and some participants reported the removal was good, while some others felt "the way [they] can see each other but not hear from [work of] each other [was] most efficient".

A Chance to Explore Ideas Freely - Earlier research focusing on privacy and awareness in collaborative music (Fencott & Bryan-Kinns, 2010) shows participants interpreted personal space as an"area for experimentation and development, and participants often described their personal space as an area to 'prepare', 'sketch', 'test' and 'draft' contributions". Similarly, as shown in Table 5.4, and as reported in the subsection "*preference*" of section "*participant report*", most precipitants held a positive attitude to the additions of personal

spaces in both C_{PI} and C_{PV}. Participants reported although C_P is good for people to work together, a workspace to work on their own is necessary (P_{1A}, P_{5A}, P_{6A}, P_{9B}, P_{14B}, P_{15A}, see Table 5.4). This is because the additional personal space helped them to "focus on personal composition before sharing it", made it easier to "create own idea", and made it possible to "create something new without disturbing [their] collaborator and vice versa", see more detailed opinions in Table 5.4. To conclude, using personal spaces temporarily enables producing ideas independently, which may also increase the variety of diversity of final output, e.g. one participant reported "not being able to hear my collaborator's work added an interesting dynamic to the piece". In this way these creations acted as useful intermediates which were then discussed, revised, and combined into the final group piece in public space.

5.4.2 Impacts on Space and Territory Measures

Note in this section, rather than all the three conditions, we look specifically into the data of C_P, as C_P was designedly experienced prior to the other two conditions to investigate the emergence of territories without impacts from the addition of personal spaces. As reported in results, in C_P all the groups formed a group territory and 18 out of 21 groups formed personal territory with relatively fixed positions. In the visual traces shown in Figure 5.5 and 5.6, it can be seen that most of the interaction dots fall inside the personal territory with the same colour or inside the group territory, indicating participants did most note edits inside their personal territory or group territory, they did fewer note edits inside each other's territory, which can also be proved by Table 5.3. We interpret this as participants had an ownership of their personal territories, music interfaces inside the personal spaces can be seen as personal interfaces, which limits other's access to these interfaces.

Participants formed not only personal territory, but also group territory, which matches the types of territory emerged in tabletop based collaboration (Scott et al., 2004). While personal territory served to ease people's individual activities, group territory served a different function. Participants had equal access to group territory and the music interfaces located in group territory, there was no clear ownership of these interfaces, which might possibly ease the concern of editing on it. As a result, more music edits fell in group territory (44.71 in group territory *vs* 39.94 in personal territory *vs* 29.24 in other's territory). Participants seldom entered other's personal territories and fewer edits were done in other's personal territory, the majority of edits fell in either the personal territory or group territory. These territory-related behaviours match

Taylor's claim that territorial behaviour occurs during human interaction in territories, they build territoriality which gains them some degree of excludability of use, responsibility for and control over activities in these sites (R. B. Taylor, 1988).

The results showed that people not only build territoriality in real, but they also build it in virtual, and the territoriality serves a similar function. Thus, H1 is supported.

Two Types of Group Territory Configurations

By inspecting the C_P part of Figure 5.6, 2 types of territory configurations can be identified:

(i) Side-by-side work configuration. Seven groups (Group 5, 7, 10, 11, 12,15, 20) used this arrangement for working together, see Figure 5.6. Take Group 7 (Figure 5.5) as an example, participants faced roughly the same direction towards the music interfaces, and did note edits side by side. Inside the group territory, most of their note edits fell on one side of the group territory. Working side-by-side is common in daily working settings. Although due to distances or due to coworker's body as an obstacle, this configuration does not provide an equal access to the all parts of the shared working interface, being able to see the interfaces from the same side enables both participants to have a similar perspective on the interfaces, which is important for having a shared knowledge of co-work undergoing. We think the reason for the emergence of this configuration is the flat music interface of LeMo, leading participants borrowed their daily life experiences. See Figure 5.3b, although the interface is 3D, it is still quite flat in *Matrix* mode.

(ii) Face-to-face work configuration. As shown in Figure 5.6, participants of five groups (Group 3, 9, 16, 17, 18) worked face to face with music interfaces between them, see Group 9 (Figure 5.5) as an example. Different from side-by-side, the note edits spread more evenly in the group territory. In the real world, people do talk face-to-face, but people seldom work or interact with a vertical physical interface between them since a transparent interactive work interface is very rare in the real world and an opaque one will block their sight and face to face communication. While the VE gives the opportunity of creating semi-transparent interfaces (Figure 5.1a), which enables participants to do so without obstructing sight between collaborators, thus, participants borrowed their experience of talking face-to-face. This configuration ensures both participants an equal access to the interfaces, enable them to see each other and the music interfaces at the same time, which can be helpful to track each other's activities. They can also interact with music interfaces without colliding

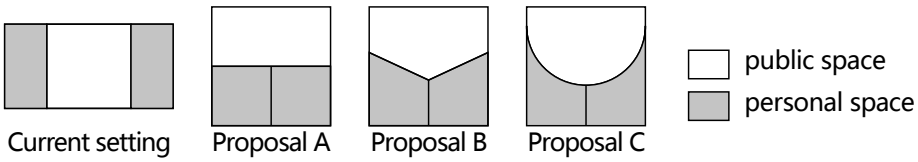


Figure 5.9: Proposals of different space settings.

with each other's bodies (the avatars). However, this configuration has two limitations: (i) The interfaces need to be specifically designed to be directionless, ensuring it can be manipulated from more than one direction; (ii) Participants' view of the interfaces is reversed from each other's, which can be an obstacle to a shared knowledge of the co-work, e.g. with the music interface of LeMo, participants perceived the play-line moves oppositely.

5.4.3 Negative Effects of Introducing Personal Space on Territory

Results have shown the addition of personal spaces led to a shrunken size of group territory, decreased number of note edits done inside group territory and a bigger co-located positions. Participants did not rotate or switch positions anymore (happened in Group 4, 6 and 21, see Figure 5.6), all the groups formed a fixed personal territory in the public space. Results have shown group territory became significantly weaker (Table 5.3), e.g. we can also see from Figure 5.6 that Group 8, and Group 19 in C_{PV} did not form a group territory, group 13 formed a much weaker group territory(the grey colour becomes much lighter). We argue the reasons of this can be two-fold, first, when personal spaces available, participants spent a considerable amount of time in the personal spaces, led to less presence in the public space, resulting in less chance to form group territory. The second reason is the locations of personal spaces are at the opposite sides of the public spaces, for ease of using personal spaces, participants tend to manipulate the interfaces somewhere near their personal spaces. The measure of colocation (colocation is significantly greater in C_{PI} and C_{PV} compared with C_P). An increased distance between each other resulted in a smaller overlap of personal territory, namely a smaller group territory.

However, we believe this negative effect is in part because in the current setting, personal spaces distribute on the opposite side of the group space, resulting in a long distance between participants' accesses of personal space. See 3 proposals shown in Figure 5.9, shortening the distance between the collaborators' personal spaces can possibly reduce the negative effect yielded by the current setting.

5.4.4 Key Findings

In summary, the following are key findings from the results:

- Impacts on collaboration in terms of awareness and communication - the invisible personal space resulted in reduced awareness and less communication, and this is probably the cause of participant's preference for visible personal space.
- Impacts on collaboration in terms of performance - Additional personal spaces supported individual creativity and increased efficiency with the cost of shrunken group territory and decreased number of note edits in group territory.
- Observation of the use of space and formation of territory - The study identified the emergence of group territory and personal territory. At the same time, two working configurations were also found: *side-by-side work configuration* and *face-to-face work configuration*.

5.5 Design Implications

According to these three key findings concluded above, we propose three design implications for Shared Virtual Environments (SVEs) focusing on CMM:

(1) When personal space is provided, personal space with public visibility is preferred and thus should be prioritised, cf. (Fencott & Bryan-Kinns, 2010).

(2) Similar to existed findings in non-immersive media (Fencott & Bryan-Kinns, 2010; Scott et al., 2004), where possible, users should be provided with personal spaces to support their efficiency and creativity. However the possible disadvantages of adding personals space should be aware of.

(3) The virtual shared working interfaces should be designed to suit *side-by-side work configuration* when direction forms an important factor of the working content (e.g. reading texts or diagrams), whereas transparent directionless interfaces can be applied to suit *face-to-face work configuration* to achieve a better face-to-face communication and an equal access to the shared interfaces.

5.6 Chapter Summary

This chapter has presented a study, which explored how three different configurations impacted collaboration from three aspects (awareness, communication and performance) and how the space was used. By using system logging tools, the study also identified the emergence of group and personal territory during in

collaboration in SVE and identified two types of configuration of group territory. Thee key findings have been concluded and three-point implications have been proposed for future design of SVEs focusing on supporting collaborative tasks based on the findings. The next chapter will present Study III, which tests more forms of personal spaces, exploring if the disadvantages of introducing personal spaces can be alleviated.

Chapter 6

Study III - Spatial Audio Approach

The previous chapter explored how the territory emerges during collaborative music making in the SVE LeMo, identified the advantages and disadvantages introduced by adding personal spaces. In this chapter presents a following study based on LeMo II, in which 52 users composed music in pairs using four different virtual working space configurations will be presented. This study aims to understand designing virtual space better and minimise the disadvantages of introducing personal space that have been identified by Study II. Key findings indicated by results include: (i) providing personal space is an effective way to support collaborative creativity in SVEs; (ii) personal spaces with a fluid light-weight boundary could provide enough support, worked better and was preferable to ones with rigid boundaries; (iii) a configuration that provides a movable personal space was preferred to one that provided no mobility. A journal paper written based on this study had been published in PeerJ Computer Science.

6.1 Motivation

The real world envelops us with space that we share with others, and in this surrounding environment, people perceive rich sensory information about objects and events happening around us. Using this information, people interact with this outer world around us via inference, manipulation, and exploration. In a similar fashion, people interact with each other. In other words, space can be seen as a material of human activity (Raffestin, 2012), and it has a great influence on social activity, e.g. the size of space limits what kind of actions can

be performed, the fill material of a space limits how far people can see or hear, and the proximity between bodies and objects in a space limits their scope of influence.

Digital virtual spaces have existed in different forms for several decades. One of the earliest examples can be video games, e.g. *Star Trek* created in early 1970s provides a computational space that players can visit and experience through text descriptions on a computer screen, see Case et al. (1990). Though these non-immersive media can involve people to a very high level and generate the experience of flow, few of them have enabled people to interact in a natural way that is similar to the way that people experience real-world interactions, e.g. inputting information using keyboards and clicking a mouse. These interactions in non-immersive media have very different properties to real-world interactions (Gaver, 1992). In contrast, Virtual Reality (VR) provides a novel space for multi-sensory experience (Turchet et al., 2018), and enables people to see, hear, and even interact with a virtual space naturally. It offers the potential for people to coordinate collaborative activities in a much more similar way to the real world, presenting people an opportunity to collaborate in virtual space in a more natural way in comparison to non-immersive digital media.

Whilst VR has become a hot topic and has been researched intensively, little attention has been paid to human-human interactions in Shared Virtual Environments (SVEs), with even less being paid to addressing the creative and collaborative aspects of these interactions. This raises a number of interesting questions: is there any difference between collaboration in SVEs and real-world collaboration? How should SVEs be designed to support collaborative creativity? Having a better understanding of the role of space and territory within creative collaborations might provide a strong starting point, since real-world collaborations make use of space (Raffestin, 2012) and the demarcation of personal and shared territory is a spatial strategy to affect, influence and control resources and access (Sack, 1983). Hence an effective arrangement and utilisation of a working space can possibly be a crucial factor to a successful collaboration in SVEs too. Thus it would be interesting to design and test more working space configurations to see if they can minimise the side effects of introducing personal spaces and provide more fluid support to creative collaboration in SVEs.

6.2 Experiment Design

Acoustic Attenuation - Sound attenuates as a result of diminishing intensity when travelling through a medium. This feature of sounds enables humans to use their innate spatial abilities to retrieve and localise information and to aid performance (cf. Billinghurst & Kato 2002). Whilst it is hard to adjust the

acoustic attenuation of a real medium (e.g. the air) to enhance its potential, within VR, as the audio is simulated, an augmented spatialised sound can be simulated purposely. E.g. there are many packages available online for simulating a realistic spatialised sound Oculus Native Spatializer[1], Superpowered VR Audio Spatializer[2], Google VR Audio Spatializer[3] , and Steam VR Audio Spatializer[4], Spatialized Audio Rendering for Immersive Virtual Environments (Naef et al., 2002). Research has been done on investigating the impacts of spatialised sounds on user experience in VR (Hendrix & Barfield, 1996b). However, little research explores how the spatialisation of sound may affect or aid collaboration (e.g. CMM). Considering sound is both the primary medium and the final output of the creative task (Study II), by affecting the audio, different settings of acoustic attenuation can possibly affect the collaboration differently. With the ability to modify the simulated acoustic attenuation in an immersive virtual environment, a sonic privacy can possibly be created by augmenting acoustic attenuation, this privacy can then possibly be used as personal space supporting individual creativity in CMM.

6.2.1 Hypotheses

Research has suggested users should be allowed to work individually in their personal spaces at their own pace, cooperatively work together in the shared space, and smoothly transition between both of the spaces during collaboration (Greenberg et al., 1999; Shen et al., 2003; Sugimoto et al., 2004). In Study II, following this implication, three different spatial configurations (C_P,C_{PI} and C_{PV}) were built and tested to see their different impacts of these spatial configurations on collaborative music making in SVEs. The results showed adding personal space to be helpful in supporting collaborative music making in SVE, since it provides a chance to explore individual ideas, and provides higher efficiency. However, several negative impacts also showed up along with the addition of personal space, e.g. longer average distance between participants, reduced group territory and group edits. This might due to: (i) the separated stationary locations of the personal spaces, which meant users had to leave each other to use them, causing a longer distance between participants and less collaboration; (ii) the rigid boundary between public space and personal space made users more isolated, resulting in a higher sense of isolation. Thus designing some new types of personal territory might be able to eliminate these disadvantages, and to provide a more flexible, more fluid experience to the collaborators. To increase the

[1] `https://developer.oculus.com/documentation/audiosdk/latest/concepts/book-ospnative-unity/`

[2] `https://superpowered.com`

[3] `https://developers.google.com/vr/reference/ios-ndk/group/audio`

[4] `https://valvesoftware.github.io/steam-audio/`

flexibility, enabling users to use personal space anywhere on the stage in SVE might be a solution and the flexibility might positively affect the collaboration. As such, H1 was developed.

To make the shift between personal and public spaces more fluid, inspired by the implication that the separation between public and personal workspace should be gradual rather than too rigid (Greenberg et al., 1999), the attenuation feature could possibly be applied to form a gradual personal space, enabling a fluid transition between personal space and public space. This is because the sound is both the primary medium of collaborative tasks and the final work of CMM (Men & Bryan-Kinns, 2018), thus by manipulating acoustic attenuation, sonic privacy can be produced. For instance, different levels of attenuation can lead to different levels of sonic privacy, and a high level of sonic privacy may play a similar role of personal space, thus H2 was developed. Additionally, the acoustic attenuation, rather than a personal space with rigid separation from public space, enables a gradual shift between personal and public workspace, which may possibly increase the fluidity of the experience and support collaboration better (cf. (Greenberg et al., 1999)). Thus H3 was developed. Below are the three hypotheses:

H1 - Personal space with mobility provides better support for collaboration than personal space with no mobility.

H2 - Attenuation can play a similar role to personal space with rigid form in CMM in SVE, providing collaborators with a personal space and supporting individual creativity during the collaboration.

H3 - Acoustic attenuation provides a fluid transition (no hard borders nor rigid forms) between personal and public spaces, which supports collaboration better compared to conditions with rigid borders.

6.2.2 Independent Variable

Spatial configuration is the independent variable in this experiment. To investigate these three hypotheses, four space configurations were developed as the independent variable levels, as shown in Figure 6.1, including:

Condition 1: **Pub**lic space only (referred to as $\mathbf{C_{pub}}$): where players can generate, remove or manipulate Spheres, and have equal access to all of the space and the music interfaces. As no personal space is provided, a shift between public and personal space does not exist, and users cannot shift to personal space.

Condition 2: Public space + **Aug**mented Attenuation Personal Space (referred to as $\mathbf{C_{aug}}$). In addition to $\mathbf{C_{pub}}$, the sound attenuation is augmented. The volume of audio drops much faster, creating sonic privacy, which can be

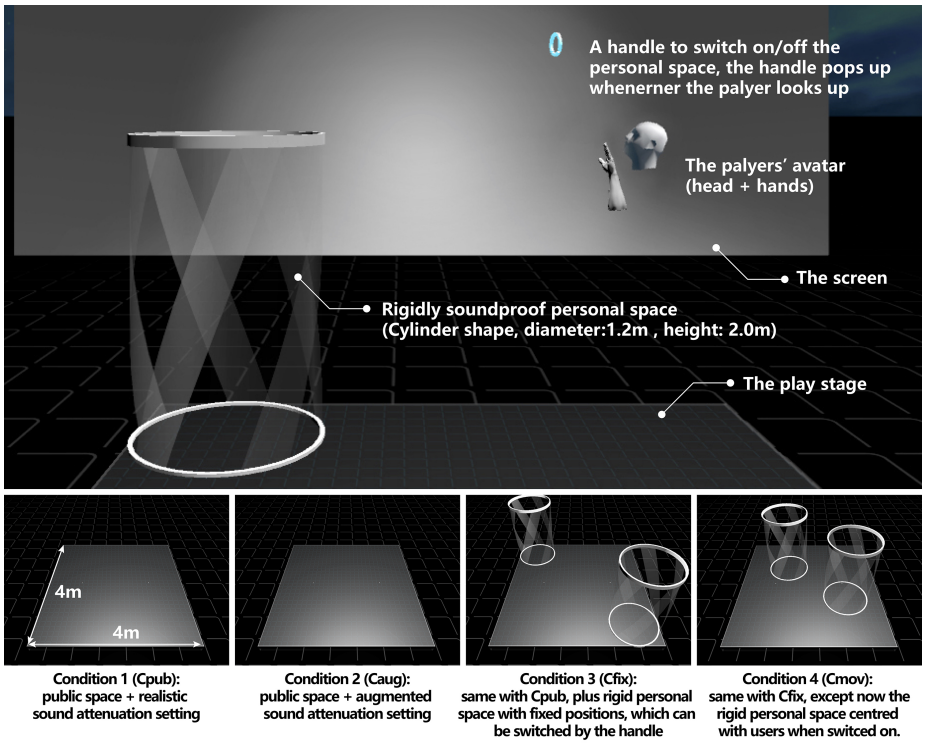


Figure 6.1: Top view of the four experimental condition settings.

seen as a personal space. As the volume changes gradually with the changes in distance, the shift between personal space and public space is gradual.

Condition 3: Public space + **Fix**ed Personal Space (referred to as $\mathbf{C_{fix}}$). In addition to $\mathbf{C_{pub}}$, each user is now provided with a personal space located at the corner of the stage (see Figure 6.1), which works like an acoustically solid boundary between public space and personal space. In other words, the shift between personal space and public space is now rigid. Users have a handle to activate/deactivate the personal space, the handle appears automatically over their head when they look up.

Condition 4: Public space + **Mov**eable Personal Space (referred to as $\mathbf{C_{mov}}$). Every feature of this condition is the same as $\mathbf{C_{fix}}$, except now the personal space appears centring the user's current head's position when being triggered.

Note the sound attenuation in C_{pub}, C_{fix}, C_{mov} are set to mimic the real sound attenuation in the real world rather than no attenuation at all, making the conditions less artificial and more natural.

6.2.3 Dependent Variables

To identify how users use the space and the effect of adding augmented sound attenuation or personal space, a series of dependent variables were developed, which can be split into *Participant Reports* and *Activity Assessment*.

Participants Reports

Questionnaires were used to collect participants' subjective assessment of the conditions and their experience of the collaboration. Most items of the questionnaire inherited from Study II. The rest of the questions were designed to question people's preference for conditions. The questionnaire included questions on:

(1) Awareness measure: i) sense of co-worker's presence and ii) sense of collaborator's activities. Note sense of self-presence is not included in the questionnaire because based on results from previous studies, we found LeMo was capable to support a high level of self-presence, and the chance that the difference can introduce great impact on this is very low.

(2) Communication measures: quality of communication, which may vary as the visibility of spaces can possibly affect the embodiment and nonverbal communication.

(3) Performance measures: i) the satisfaction of the final music created reflects the quality of collaboration, cf. (Bryan-Kinns, 2013; Bryan-Kinns & Hamilton, 2012). ii) Contribution, including the feeling of self's contribution and the feeling of others' contribution.

(4) Other measures: preference of the conditions, to see if users have subjective preferences towards the settings.

Both a PSQ-Session Questionnaire (PSQ, see items in Table 6.1) and a Comparison Questionnaire (CQ, see items Table 6.2) were made cover these measurements. We used both to fully take their advantages in collecting data. PSQ questionnaire were filled after each session to collect data before participant's feeling fading out and CQ was filled after all the conditions being experienced, so participants could rank the conditions according to a certain metric.

Activity Assessments

To access the characteristics of collaboration, based on the system-logged data of activity in the collaboration, the following measures were developed:

(1) Awareness measures: i) time participants spent paying close or ordinary attention to collaborator; ii) number of times paying close or ordinary attention to the collaborator. Strictly speaking, here "paying attention" means "facing toward the collaborator's avatar" as no eye tracker was involved in this study.

(1) Communication measures: As mentioned before, LeMo enable users to record users' head rotation and position. As such, we were able to record their attentions. Attention can be regarded as a way of non-verbal communication. Strictly speaking, here "paying attention" means "facing toward the collaborator's avatar" as no eye tracker was involved in this study. More specifically, here the attention measure include: i) time participants spent paying close or ordinary attention to collaborator; ii) number of times paying close or ordinary attention to the collaborator.

(3) Performance measures: i) number of musical note edits; ii) number of note additions and deletions; iii) number of mutual note modifications. Here mutual note modifications indicate an edit on a certain note, the last update of which was performed by the collaborator, cf. (Bryan-Kinns et al., 2007).

(4) Space and territory measures. Time and amount of use of personal space (only in condition C_{fix} and C_{aug}): i) number of uses of, ii) length of time of using, and iii) average duration of each use of personal space. Location and territory: iv) distribution of participants' locations and interactions; v) the sizes of personal and group territory if they emerge; vi) note edits fallen in different types of territory; vii) average distance between participants, cf. colocation in Bryan-Kinns (2013).

Participants and Procedure

Fifty-two participants (26 pairs) were recruited for this study[5] via emails sent to group lists within the authors' school. All participants were aged between 18 and 35, with a mean age of 23.00 ($SD = 4.37$). Each participant was compensated 10 GBP for their time. Participants' mean rating of personal musical theory knowledge is 3.92 ($SD = 2.50$) on a 10-point Likert scale, where higher values indicate increased knowledge. Twenty-four (46.15%) play one or more instruments, and the remaining 28 (53.85%) do not. Participants' experience of collaboratively composing music is 2.13 ($SD = 1.56$) on a 10-point Likert scale, where higher values indicate increased experience. Regarding familiarity with computers, 31 (59.62%) participants rated themselves as computer "experts", 20 (38.46%) chose "intermidiate", and only 1 participant (0.02%) chose "beginner". Twenty participants (38.46%) had tried VR 2-5 time before, 20 (38.46%) had only tried once, the remaining 12 (23.08%) had no VR experience before. Thirty-seven participants knew their collaborators very well prior to the experiment, 3 met their collaborators several times, but did not know well, the remaining 12 did not know their collaborators at all prior to the experiment.

[5]The Queen Mary Research Ethics Committee granted ethical approval to carry out the study within its facilities (Ethical Application Ref: QMREC2005).

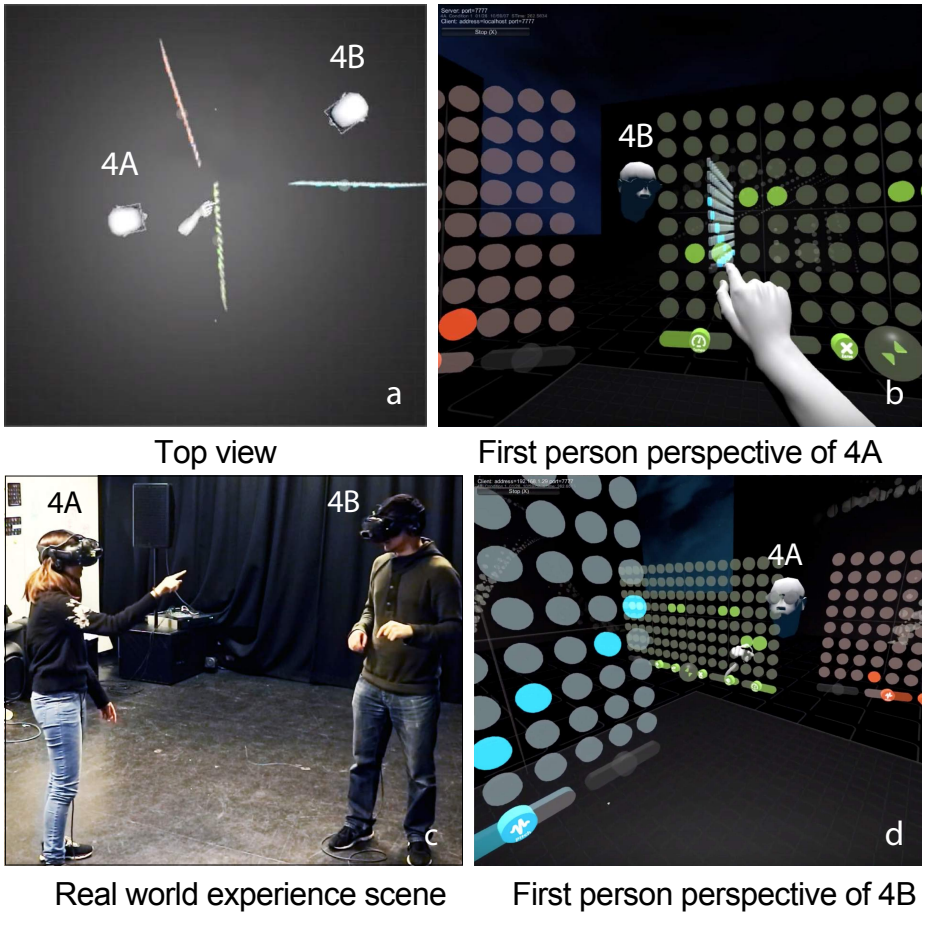


Figure 6.2: Participant 4A and 4B are creating music together.

After reading the information form and signing the consent form, each pair of participants first received an explanation of the music interface of LeMo II. Then one experimenter demonstrated all of the interaction gestures supported in LeMo II. By linking the demonstration with the first-person view shown on monitors, participants had a chance to learn how to play LeMo II. Then, participants took a trial (5-15 minutes) to try all the ways of interaction. The trial ended once they were confident enough of all available gestures. The length of time of the tutorial session was flexible to ensure participants with diverse musical knowledge could grasp LeMo II. Participants were then asked to have four sessions of collaboratively composing music that was mutually satisfying and compliments an animation loop, each session lasts 7 minutes based on the pilot study and a previous study (Men & Bryan-Kinns 2018, i.e. Study I), 7 minutes were found to be sufficient for the task. Check the experimental scene in Figure 6.2. To avoid the impact of adding personal spaces and have a pure observation on how participants form their own proximity in the public space, all four conditions were experienced in a fully randomised sequence to counterbalance the learning effect. In total four animation loops were introduced to trigger participants' creativity, each to be played in one experimental session on four virtual screens surrounding the virtual stage. These clips were played in an independently randomised sequence to counterbalance impacts on the study. Each session ended with a Post-Session Questionnaire (PSQ, see Table 6.1). After all the four sessions finished, the Comparison Questionnaire (CQ, see Table 6.2) and a short interview were carried out at the end of the experiment. A data-log system logged time-stamped data from events generated by users' interactions and movements: positions and rotations of head, index finger, thumb finger, the manipulation with musical interface (addition/deletion/re-positioning of musical interfaces, addition and deletion of music notes), usage of personal space (activation/deactivation of personal spaces).

6.3 Results

Next, the following sections will detail results collected by participant reports and activity assessments.

6.3.1 Participant Reports

This section reports on the results of the questionnaires. Ratings of Post-Session Questionnaires were refined to counterbalance the learning effect and then analysed with Wilcoxon Rank Sum Tests (Table 6.1). Binomial tests were run to see if the number of ratings for each option was significantly different than would

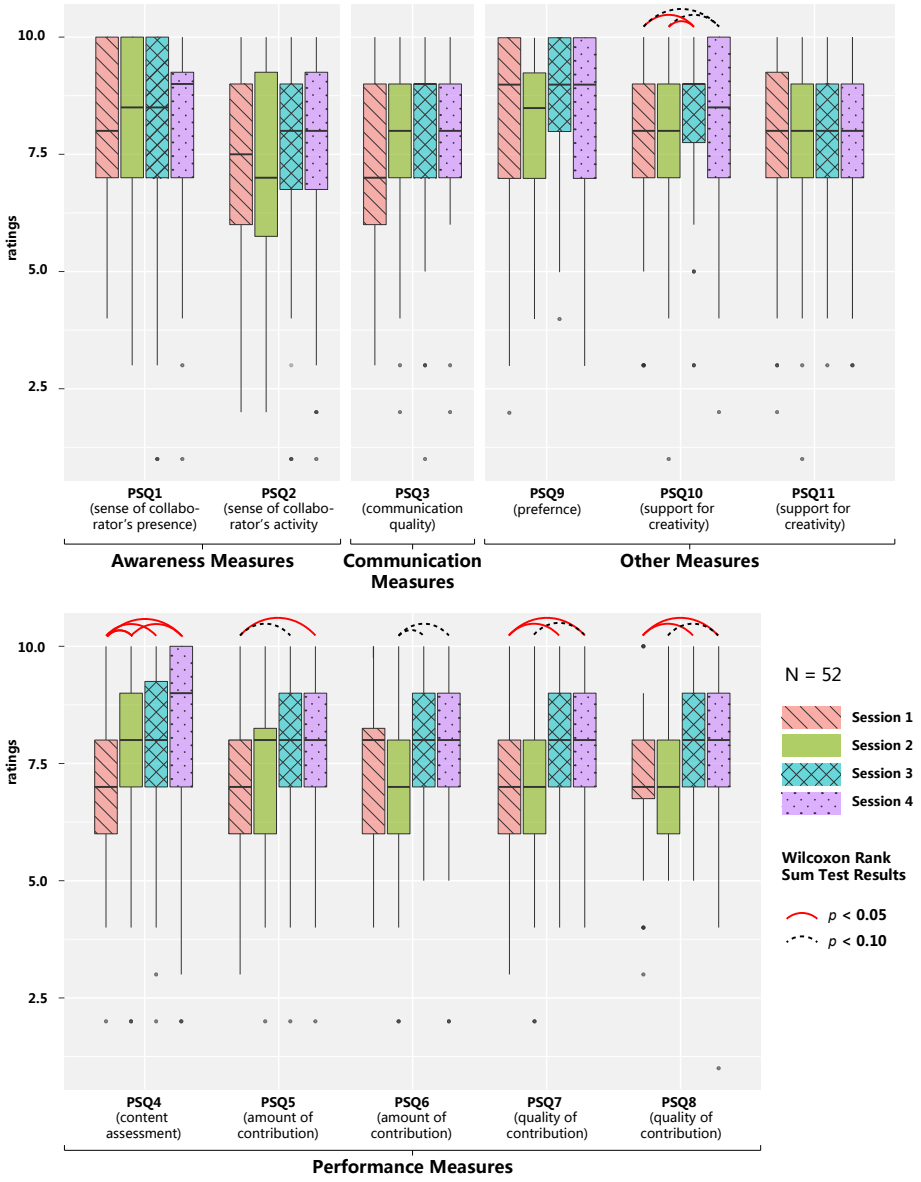


Figure 6.3: Results of Post-Session Questionnaire (N = 52) for all sessions; arcs show significant (solid line) and marginal-significant (dotted line) differences between conditions, indicating possible ordering effects.

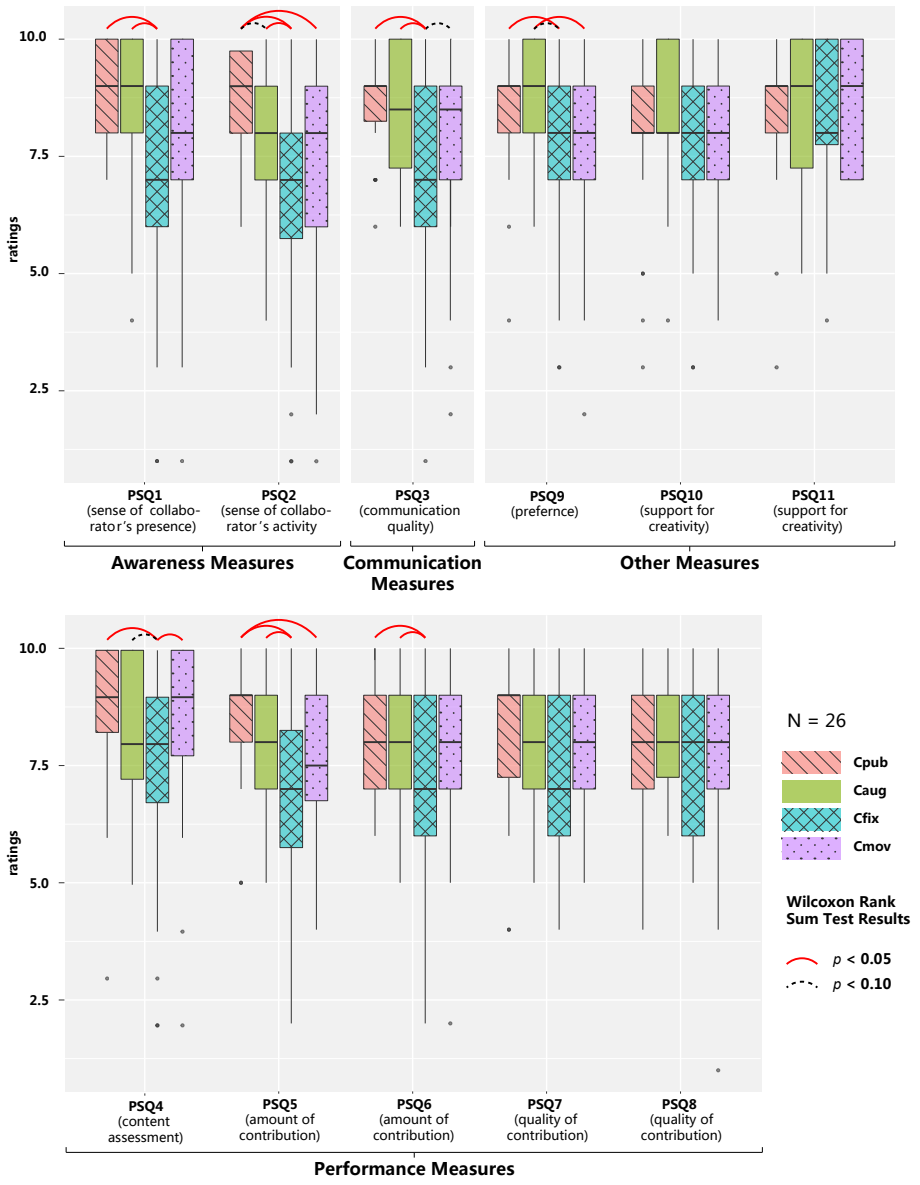


Figure 6.4: Results of Post-Session Questionnaire (N = 26), data grouped by experimental conditions. Only data collected in the latter two sessions are included; arcs showing significant (solid line) and marginal-significant (dotted line) differences between conditions.

Table 6.1: Results of Post-Session Questionnaire[a] and results of Wilcoxon Rank Sum Test (two-tailed)[b].

Measure	Questions	C_{pub}	C_{aug}	C_{fix}	C_{mov}	C_{pub} *vs* C_{aug}	C_{pub} *vs* C_{fix}	C_{pub} *vs* C_{mov}	C_{aug} *vs* C_{fix}	C_{aug} *vs* C_{mov}	C_{fix} *vs* C_{mov}
		M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*
Awareness Measures	PSQ1[c] (sense of collaborator's presence)- I always had a strong feeling that my collaborator was there, collaborating with me together, all the time										
		8.91	8.54	7.07	7.93	0.7961	0.004813	0.1636	0.01946	0.3229	0.1368
		(0.92)	(1.68)	(2.52)	(2.26)	(298.5)	(450)	(377.5)	(497)	(420)	(302)
	PSQ2 (sense of collaborator's activity) - I had a clear sense of what my collaborator was doing										
		8.73	7.96	6.50	7.29	0.08094	0.0003856	0.03436	0.02786	0.5095	0.176
		(1.20)	(1.54)	(2.52)	(2.49)	(368.5)	(487.5)	(414.5)	(489.5)	(402)	(310)
Communication Measures	PSQ3 (communication quality) - How would you rate the quality of communication between you and your collaborator during the session										
		8.68	8.50	7.04	8.04	0.7644	0.004494	0.3038	0.01038	0.5404	0.05073
		(1.09)	(1.36)	(2.25)	(1.97)	(300.5)	(450.5)	(359)	(510)	(399)	(274)
Performance Measures	PSQ4 (content assessments)- How satisfied are you with the final piece of loop music you two created in this session										
		8.64	8.38	7.21	8.32	0.4287	0.005155	0.5557	0.05449	0.803	0.02163
		(1.73)	(1.50)	(2.22)	(1.96)	(323.5)	(448.5)	(337.5)	(473.5)	(349.5)	(254)
	PSQ5 (amount of contribution) - The amount of your contribution to the joint piece of music is										
		8.41	8.15	6.96	7.50	0.4776	0.009236	0.03928	0.04281	0.166	0.4489
		(1.44)	(1.46)	(2.15)	(1.67)	(320)	(439.5)	(412)	(479.5)	(443)	(346)
	PSQ6 (amount of contribution) - The amount of your collaborator's contribution to the joint piece of music is										
		8.18	8.23	7.29	7.61	0.8486	0.08916	0.4025	0.06406	0.3008	0.4739
		(1.26)	(1.39)	(1.96)	(1.97)	(276.5)	(394)	(350.5)	(469.5)	(423)	(348.5)
	PSQ7 (quality of contribution) - What do you think of the quality of your contribution to the joint piece of music is										
		8.05	7.81	7.36	7.86	0.319	0.1031	0.4648	0.3596	0.2829	0.8599
		(1.70)	(1.41)	(1.68)	(1.53)	(333.5)	(390)	(345)	(416.5)	(327)	(353.5)
	PSQ8 (quality of contribution) - What do you think of the quality of your collaborator's contribution to the joint piece of music is										
		7.73	8.19	7.54	7.75	0.3496	0.5636	0.6459	0.1143	0.6992	0.3559
		(1.52)	(1.20)	(1.50)	(2.05)	(241.5)	(337.5)	(284.5)	(453.5)	(386)	(336.5)
Other Measures	PSQ9 (preference) - I enjoyed the spatial configuration of this virtual world very much										
		8.27	8.65	8.18	8.07	0.2622	0.9358	0.3863	0.3863	0.2165	0.8010
		(1.61)	(1.60)	(1.87)	(1.86)	(233)	(303.5)	(311.5)	(412.5)	(433.5)	(407.5)
	PSQ10 (support for creativity) - I think the spatial configuration in this session was extremely helpful for creativity										
		8.55	8.77	7.61	7.82	0.5695	0.07706	0.1563	0.02372	0.06318	0.695
		(1.44)	(1.34)	(2.01)	(1.94)	(259)	(396.5)	(379)	(492)	(469)	(368)
	PSQ11 (support for creativity)- I feel like the spatial configuration in this session was extremely helpful to support the to support the development of my own ideas										
		7.82	8.35	7.71	7.75	0.5211	0.6456	0.5029	0.2172	0.1452	0.8999
		(1.92)	(1.50)	(1.88)	(1.62)	(255.5)	(331.5)	(342)	(434)	(446.5)	(400)

[a] With 10-point-Likert scale, 1 indicate no fulfilment at all with the description of the questionnaire and 10 indicate a full fulfilment.
[b] Note statistics here were calculated based on the data collected from the third and fourth session to counterbalance the learning effect.
[c] The index numbers in this table are only for reference purpose, not reflecting the real order of the items listed in the questionnaire.

Table 6.2: Results of Binomial Test of Comparison Questionnaire (CQ)[a].

Measure	Question description	Option	C_{pub}		C_{aug}		C_{fix}		C_{mov}	
			k	p	k	p	k	p	k	p
Awareness Measures	CQ1[b] (sense of collaborator's activities) - Which session you found most difficult to track collaborator's activities?									
		most difficult	7	0.03317	12	0.4469	20	0.02205	13	1.000
		least difficult	22	0.004691	14	0.4262	8	0.06971	8	0.06971
	CQ2 (sense of collaborator's presence) - Which session did you have the strongest sense that your collaborator was there working with you together									
		strongest	27	2.807e-05	17	0.1322	2	5.277e-05	6	0.01368
		least strongest	4	0.001378	7	0.03317	28	8.12e-06	13	1.000
Communication Measures	CQ3 (communication quality) - Which session did you have the best quality of communication between your self and your collaborator									
		best quality	20	0.02205	17	0.1322	4	0.001378	11	0.3232
		worst quality	6	0.01368	13	1.000	25	0.0002698	8	0.06971
Performance Measures	CQ4 (content assessment) - In which session, you made the music you were most satisfied with?									
		most satisfied	16	0.2089	12	0.4469	10	0.2146	14	0.4262
		least satisfied	13	1.000	9	0.1292	21	0.01054	9	0.1292
	CQ5 (contribution) - Which session do you you feel you made the most contribution to the joint piece									
		most contribution	14	0.4262	12	0.4469	13	1.000	13	1.000
		least contribution	11	0.3232	13	1.000	13	1.000	15	0.3084
	CQ6 (contribution) - Which session do you you feel your collaborator made the most contribution to the joint piece									
		most contribution	11	0.3232	11	0.3232	16	0.2089	14	0.4262
		least contribution	15	0.3084	12	0.4469	18	0.07806	7	0.03317
Other Measures	CQ7 (preference) - In which session, you enjoyed the spatial configuration the most?									
		most enjoyed	10	0.2146	16	0.2089	10	0.2146	16	0.2089
		least enjoyed	15	0.3084	10	0.2146	20	0.02205	7	0.03317
	CQ8 (preference) - Which session had the best setting for creating a good piece of music collaboratively									
		best setting	16	0.2089	16	0.2089	8	0.06971	12	0.4469
		worst setting	13	1.000	10	0.2146	19	0.04298	10	0.2146
	CQ9 (coordination) - Which session did you find most difficult to cooperate with collaborator									
		most difficult	7	0.03317	12	0.4469	22	0.004691	11	0.3232
		least difficult	21	0.01054	14	0.4262	7	0.03317	10	0.2146

[a] Lower-tailed test when $k < 13$, two-tailed test when $k = 13$, upper-tailed test when $k > 13$.
[b] The index numbers in this table are only for reference purpose, not reflecting the real order of the items listed in the questionnaire.

be expected by chance, upper-tail, lower-tail or two-tailed tests were used accordingly, results are listed in Table 6.2. Next, how the learning effect on PSQ was counterbalanced will be briefed and then results will be reported.

Counterbalancing the Learning Effect

As aforementioned, a fully randomised order of experimental conditions was introduced to counterbalance the learning effect. However, it turned out many measurements in the Post-Session Questionnaire were still affected by the sequence to a certain extent, as shown in Figure 6.3, in which data from all groups were compiled according to how the group was ordered in the session sequence. Wilcoxon Rank Sum tests were run between each two conditions of every question. An orange solid arc indicates a significant difference between two bars ($p < 0.05$), and a grey dotted arc indicates a trend toward a significant difference ($p < 0.1$). The arcs show that results of four questions (especially PSQ4, PSQ5, PSQ6, PSQ7, PSQ8, PSQ10) are very sensitive to the sequential position of the session. Specifically, in later sessions, participants responded more positively to the 'helpfulness of the spatial configuration' (PSQ10), higher satisfaction with their output (PSQ4), and both more, and better quality of contributions by themselves and contributors (PSQ5, PSQ6, PSQ7, PSQ8). This is probably due to the learning effect which has a much stronger effect on these measures compared with the differences between experimental conditions. Given the limited experience some participants have had in VR and collaborative music making, learning effect could have possibly and greatly promoted participants' skills and knowledge in performing the task, resulting in a better feeling of the spatial configuration of the session, higher quality of output, more contribution with better quality in later sessions. This learning effect has been also mentioned by some participants in the interview. More details will be discussed in the later subsection "Interviews".

To better counterbalance the learning effect and habituation on PSQ, only data collected via PSQ in later two sessions (session 3 and 4) will be retained at the expense of the halved sample size. Box-plots were then drawn (Figure 6.4) and Wilcoxon Rank Sum tests were run (Table 6.1) to compare the conditions against each other.

Awareness Measures

Results of PSQ1, PSQ2 and CQ2 reveal participants' sense of collaborators' presence and activities. C_{fix}'s ratings in PSQ1 and PSQ2 are significantly lower than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, all $p < 0.05$), indicating C_{fix} saw a lower sense of the collaborator's presence and activities. Similarly, when

being asked in which session they had the strongest sense of collaborators (CQ2), significantly many participants (27) chose C_{pub} ($0.52 > 0.25$, p = 2.807e-05, 1-sided), 17 chose C_{aug}, significantly few chose C_{mov} (chosen by 6) and C_{fix} (chosen by 2; Binomial Test, both $p < 0.05$; more details in CQ2 of Table 6.2). When questions changed to "least sense of collaborator's presense", ratings reversed, significantly many chose C_{fix} while significantly few chose C_{pub} and C_{aug} (Binomial Test, all $p < 0.05$; more details in CQ2 of Table 6.2). These results indicate that in terms of maintaining the sense of collaborator's presence, $C_{pub} > C_{aug} > C_{mov} > C_{fix}$.

Regarding the sense of collaborator's activities (PSQ2), a significantly weaker sense was reported in C_{fix} compared with C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$). C_{pub} also saw a stronger sense compared with C_{fix} (Wilcoxon Rank Sum Test, $W = 414.5$ $p < 0.0346$). No significant difference was found between C_{pub} and C_{aug} nor between C_{fix} and C_{mov}. Similarly, CQ1 of the Comparison Questionnaire reveals that significantly many participants reported tracking collaborators' activities in C_{fix} was the most difficult, and significantly many felt least difficult to do so in C_{pub} (Binomial Test, both $p < 0.05$). These indicate that C_{pub} seems to be easier for participants to track collaborators' activities, and C_{fix} is more difficult for them to do so.

Communication Measures

As listed in PSQ3, communication quality of C_{fix} ($M = 7.04$) is significantly lower than 8.68 of C_{pub} and 8.50 of C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$), and near-marginal significantly lower than 8.04 of C_{mov} (Wilcoxon Rank Sum Test: $W = 274$, 0.05073), see PSQ3 in Table 6.1 and PSQ3 in Figure 6.4. When asked to compare these sessions, significantly many participants reported the best communication quality was in C_{pub} and significantly few believed they had the best communication quality in C_{fix} (Binomial Test, both $p < 0.05$, see CQ3 of Table 6.2). Conversely, significantly few had the worst communication quality in C_{pub} and significantly many had worst in C_{fix} (Binomial Test, both $p < 0.05$, see CQ3 of Table 6.2). In summary, C_{fix} saw a relatively lower communication quality.

Performance Measures

Regarding the quality of the outcome, participants reported a mean rating 7.21 of output quality in C_{fix} (PSQ4 of Table 6.1), which is significantly lower than 8.64 in C_{pub}, and 8.32 in C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.05$), and quasi-significantly lower than 8.38 in C_{aug} ($W = 473.5$, $p = 0.05549$). Similarly, significantly many participants reported that they produced the least satisfying

piece of music in C_{fix} (Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided), see CQ4 of Table 6.2. In other words, C_{fix} tended to led to a music output with lower quality compared with other sessions.

Regarding the contribution, participants reported that they did a significantly larger amount of contributions in C_{pub} compared with C_{fix} ($W = 439.5$, $p = 0.009236$) or with C_{mov} ($W = 412$, $p = 0.03928$), and had done significantly more contribution in C_{aug} compared with C_{fix} ($W = 479.5$, $p = 0.04281$), see PSQ5. No significant difference was found in CQ5, which is also questioning the feeling of own contribution.

No significant differences were found in the ratings of the amount of the collaborators' contribution (PSQ6), except a trend reporting their collaborator had a lower amount of contribution in C_{fix} than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, $W = 469.5$, both $p < 0.1$). However, CQ6 reveals significantly few participants felt that their collaborator did the most contribution in C_{mov} (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided). These results indicate the addition of personal space in C_{fix} and C_{mov} possibly led to a weaker sense of collaborator's activities.

Other Measures

When asked the condition's helpfulness for creativity (PSQ10 in Table 6.1), on a 10-point Likert Scale, participants gave an average rating of 8.77 in C_{aug} which is significantly higher than 7.61 given in C_{fix} (Wilcoxon Rank Sum Test, $W = 492$; $p = 0.02372$). There are trends towards significances between participants' rating of C_{aug} and C_{mov} ($W = 469$; $p = 0.06318$), and between C_{pub} and C_{fix} ($W = 396.5$; $p = 0.07706$). These differences indicate that C_{aug} is better than C_{fix}, and possibly also better than C_{mov} in terms of supporting participants' creativity.

When asked to rate the helpfulness of spatial configurations to support personal idea development (PSQ11), the mean rating of C_{aug} ($M = 8.35$) is higher than that of the other three conditions (C_{pub}: $M = 7.82$; C_{fix}: $M = 7.71$; C_{mov}: $M = 7.75$), though no significant differences were found. CQ9 of Table 6.2 shows that C_{pub} was rated by significantly many participants (21 out of 52) to be the least difficult to cooperate with their collaborator (Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided), whilst significantly few participants rated C_{pub} as the most difficult one to do so (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided). On the opposite, C_{fix} was rated by significantly many participants as the most difficult (Binomial Test, $0.42 > 0.25$, $p = 0.004691$, 1-sided), and significantly few participants (7 out of 52) rated it as the least difficult (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided).

When asked the level of enjoyment of the spatial configuration (PSQ9), similar to PSQ11, C_{aug} received a higher rating ($M = 8.65$). However, no significant difference was revealed. In CQ7 of Table 6.2, when asked which session has the most enjoyable spatial configuration, out of 52 participants, both C_{aug} and C_{mov} were chosen by 16 participants, more than those choosing C_{pub} and C_{fix} (10 participants each), though no significant difference was found. When asked which session had the least enjoyable spatial configuration, a significant number of participants (20 out of 52) opted C_{fix} ($0.38 > 0.25$, $p = 0.02205$, 1-sided), and significantly few (7 out of 52) opted C_{mov} ($0.13 < 0.25$, $p = 0.03317$, 1-sided). Result of CQ8 in Table 6.2 indicates that significantly many participants (19 out of 52) believed C_{fix} is the worst setting for creating a good piece of music collaboratively. To summarise, the spatial configuration of C_{fix} is more disfavoured and that of C_{mov} is less disfavoured.

6.3.2 Activity Assessments

This section reports on measures focusing on the participants' interactive activities. All measures are listed in Figure 6.3, Wilcoxon Rank Sum tests were run to compare conditions against each other.

Communication Measures

(1) The time participants spent paying close attention to each other - Throughout the 420-second session, participants had their close attention toward their collaborators' heads for 14.04 seconds in C_{aug}, which is significantly longer than 7.19 seconds in C_{pub}, and 7.31 seconds in C_{fix} (Wilcoxon Rank Sum Test, both $p < 0.05$), and near-marginal significantly longer than 11.02 seconds in C_{mov} ($W = 1645$, $p = 0.05722$, see AA1 in Table 6.3).

(2) Participants oriented their close attention toward their collaborator for significantly different times, they did most of the time in C_{aug} ($M = 14.04$), this is significantly more than 7.19 in C_{pub}, 7.31 in C_{fix} and 11.02 in C_{mov} (Wilcoxon Rank Sum Test, C_{pub} vs C_{aug}: $W = 924.5$, $p = 0.005451$; C_{aug} vs fix: $W = 1945$, $p = 0.0001145$; C_{aug} vs C_{mov}: $W = 1683.5$, $p = 0.03122$). Wilcoxon Rank Sum Test result also shows participants paid their attention to their partner significantly fewer times in C_{fix} than they did in C_{mov} ($W = 1015$, $p = 0.02838$), see AA2 in Table 6.3. These results indicate the spatial configuration in C_{aug} significantly promoted participants to pay more close attention to their collaborator, whilst C_{mov} possibly promoted insignificantly and C_{fix} demoted insignificantly compared with C_{pub}.

(3) The time participants spent paying ordinary attention to each other. Different from the impact of spatial configurations on close attention, neither

C_{aug} nor C_{mov} significantly changed the way participants paying their ordinary attention, all fell inside the range from 70 to 80 seconds out of the 420-second session. C_{fix} greatly reduced participant's ordinary attention paid to each other, on average participants only spent 52.89 seconds on doing this, which is significantly shorter than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$) and near-marginal significantly shorter than C_{mov} (Wilcoxon Rank Sum Test, $W = 1081$, $p = 0.07865$), more details are shown in AA3 of Table 6.3.

(4) Similar to the time paying ordinary attention to each other, participants only drew an average of 32.87 times of ordinary attention to each other in C_{fix}, which is significantly lower than all the other three conditions (Wilcoxon Rank Sum Test, all $p < 0.001$), check detailed statistics in AA4 of Table 6.3), indicating the spatial configuration of C_{fix} greatly reduced participants' paying ordinary attention to each other.

Performance Measures

(1) Note edits (including note additions and deletions). On average, participants did 98.35 note edits in C_{fix}, which is significantly more than 77.13 of C_{pub}, 80.27 of C_{aug}, and 77.69 of C_{mov} (Wilcoxon Rank Sum Test, all $p < 0.05$). Note additions, as the main part of note edits, follow a similar pattern. The number of note additions in C_{fix} is significantly greater than that of C_{pub} ($W = 1026$, $p = 0.03429$), and near-marginal significantly greater than that of C_{aug} and C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.07$, check detailed statistics in AA6, Table 6.3). No significant difference was found in note deletions between conditions, this is probably due to the much smaller amount of deletions compared with the number of note edits and additions. These results indicate that participants had more musical edits, specifically note additions in C_{fix} than the other conditions.

(2) Mutual note modifications. C_{pub} saw the highest average number of mutual note modifications ($M = 4.37$, $SD = 4.42$), this is significantly more than C_{fix} ($M = 3.71$, $SD = 7.69$; Wilcoxon Rank Sum Test, $W = 1703.5$, $p = 0.01929$) and C_{mov} ($M = 2.44$, SD 3.92; Wilcoxon Rank Sum Test, $W = 1754.5$, $p = 0.007331$). C_{aug} has the second highest mean ($M = 4.23$, $SD = 5.57$), which is significantly more than C_{mov} ($W = 1687.5$, $p = 0.02514$), and near-marginal significantly more than that of C_{fix} ($W = 1068.5$, $p = 0.06614$). No significant difference between C_{pub} and C_{aug} or between C_{fix} and C_{mov} was found. These results indicate participants had more mutual modifications in C_{pub} and C_{aug} than C_{fix} and C_{mov}, which might indicate a closer collaboration.

(3) Number of note edits that fell into public/personal space. Note this measure is only applicable to rigid personal space, which were only available in C_{fix} and C_{mov}. Participants did 54.48 ($SD = 48.69$) note edits in public space,

Table 6.3: Statistics and Wilcoxon Rank Sum Test (two-tailed) of Activity Assessments (AA)

Measure	Description	C_{pub}	C_{aug}	C_{fix}	C_{mov}	C_{pub} *vs* C_{aug}	C_{pub} *vs* C_{fix}	C_{pub} *vs* C_{mov}	C_{aug} *vs* C_{fix}	C_{aug} *vs* C_{mov}	C_{fix} *vs* C_{mov}
		M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*	*p (W)*
Communication Measures	AA1 - Time spent paying close attention to collaborator (unit: second)[a]	7.19	14.04	5.51	9.43	0.01032	0.364	0.4725	0.001757	0.05722	0.1088
		(7.44)	(15.19)	(9.60)	(13.96)	(957)	(1492)	(1241)	(1833.5)	(1645)	(1105)
	AA2 - Times of paying close attention to collaborator[a]	9.31	14.79	7.31	11.02	0.005451	0.1591	0.446	0.0001145	0.03122	0.02838
		(8.33)	(11.16)	(7.48)	(11.62)	(924.5)	(1568.5)	(1234.5)	(1945)	(1683.5)	(1015)
	AA3 - Time spent paying ordinary attention to collaborator (unit: second)[a]	79.38	76.32	52.89	74.68	0.9818	0.03719	0.6655	0.02047	0.5695	0.07865
		(58.65)	(51.55)	(39.91)	(60.69)	(1356)	(1673)	(1419)	(1709)	(1440)	(1081)
	AA4 - Times of paying ordinary attention to collaborator[a]	45.02	46.77	32.87	44.98	0.7973	0.0006613	0.8888	0.0001704	0.6773	0.002744
		(19.23)	(19.42)	(13.55)	(20.67)	(1312)	(1876)	(1374)	(1930.5)	(1416.5)	(891)
Performance Measures	AA5 - No. of note edits	77.13	80.27	98.35	77.69	0.6988	0.02386	0.7599	0.03375	0.8965	0.02228
		(36.59)	(36.92)	(48.67)	(34.61)	(1292)	(1004)	(1304.5)	(1025)	(1372.5)	(1704)
	AA6 - No. of note additions	50.23	58.96	72.88	55.98	0.8301	0.03429	0.8149	0.06572	0.876	0.05591
		(27.12)	(30.03)	(40.93)	(25.31)	(1318.5)	(1026)	(1315.5)	(1068.5)	(1376.5)	(1646.5)
	AA7 - No. of note deletions	20.90	21.31	25.46	21.71	0.7108	0.243	0.8376	0.3308	0.9689	0.3323
		(14.46)	(12.94)	(18.39)	(15.15)	(1294.5)	(1172)	(1320)	(1202)	(1358.5)	(1501.5)
	AA8 - No. of mutual note modifications[b]	4.37	4.23	3.71	2.44	0.6452	0.01929	0.007331	0.06614	0.02514	0.7732
		(4.42)	(5.57)	(7.69)	(3.92)	(1422.5)	(1703.5)	(1754.5)	(1627.5)	(1687.5)	(1394.5)
	AA9 - No. of group edits (note edits done in group territory)	36.44	43.04	17.50	25.23	0.2913	0.001448	0.07839	4.043e-05	0.009044	0.1822
		(35.24)	(34.79)	(23.79)	(29.00)	(1189.5)	(1837)	(1621.5)	(1977)	(1751.5)	(1151)
	AA10 - No. of personal edits (note edits done in own personal territory)	40.50	37.10	80.62	52.42	0.9610	1.179e-05	0.0294	2.157e-06	0.02016	0.003695
		(44.81)	(38.42)	(51.89)	(38.81)	(1360)	(678)	(1017)	(623)	(994.5)	(1799)
	AA11 - No. of note edits done in other's personal territory	0.058	0	0.19	0.019	0.3267	0.1797	1.000	0.04343	0.3267	0.1686
		(0.42)	(0)	(0.89)	(0.14)	(1378)	(1275)	(1352.5)	(1248)	(1326)	(1431)
	AA12 - No. of note edits in public space	-	-	54.48	43.69	-	-	-	-	-	0.5051
		(-)	(-)	(48.69)	(34.96)	(-)	(-)	(-)	(-)	(-)	(1455)
	AA13 - No. of note edits in personal space	-	-	43.87	34	-	-	-	-	-	0.3869
		(-)	(-)	(40.10)	(25.37)	(-)	(-)	(-)	(-)	(-)	(1485.5)
Space & Territory Measures	AA14 - Size of group territory (unit: m^2)	0.3465	0.4331	0.2339	0.3103	0.152	0.1013	0.7099	0.005236	0.09421	0.2639
		(0.2443)	(0.2446)	(0.1878)	(0.1942)	(259)	(428)	(359)	(491)	(430)	(276.5)
	AA15 - Size of personal territory (unit: m^2)	0.4282	0.4547	0.7475	0.5067	0.9559	2.25e-12	0.02347	2.085e-10	0.04421	2.3e-08
		(0.1690)	(0.2193)	(0.1801)	(0.1894)	(1343)	(272)	(1003)	(374)	(1042)	(2212)
	AA16 - Average distance between collaborators (unit: metre)	1.11	1.12	2.19	1.28	0.8632	4.731e-11	0.2119	2.663e-10	0.08045	2.459e-08
		(0.38)	(0.38)	(0.58)	(0.41)	(348)	(26)	(269)	(34)	(242)	(616)
	AA17 - No. of uses of personal spaces	-	-	2.40	2.85	-	-	-	-	-	0.2912
		(-)	(-)	(1.95)	(2.14)	(-)	(-)	(-)	(-)	(-)	(1193)
	AA18 - Length of time of using personal spaces (unit: second)	-	-	128.60	112.19	-	-	-	-	-	0.4685
		(-)	(-)	(86.95)	(78.67)	(-)	(-)	(-)	(-)	(-)	(1464)
	AA19 - Average duration of each entry of personal space (unit: second)[c]	-	-	73.07	49.54	-	-	-	-	-	0.008019
		(-)	(-)	(56.55)	(44.83)	(-)	(-)	(-)	(-)	(-)	(1512)

[a] The difference between the close attention and the ordinary attention is the breadth and depth of FOV, FOV of close attention roughly covers 27 degrees (horizontally), 28 degrees (vertically) and 1.0 m (depth), whilst FOV of ordinary attention roughly covers 27 degrees (horizontally), 28 degrees (vertically) and 2.7 m (depth).

[b] Mutual note modifications include activation/deactivation, the last update of which was performed by the collaborator.

[c] Data of four participants (3B, 4A, 17B 18A) were excluded when calculating this metric as these participants did not use personal space, which made this metric not apply to them.

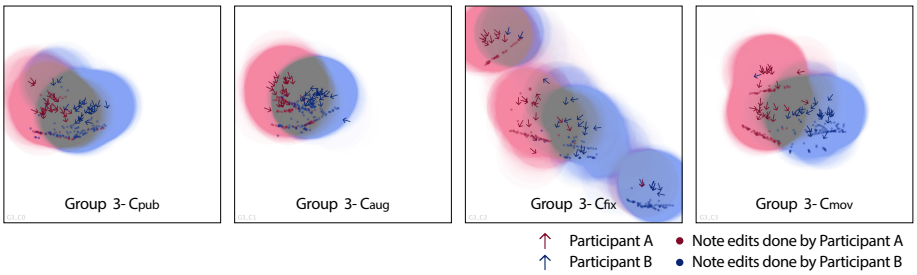


Figure 6.5: Illustrative example of visual traces of the participants' locations, directions and musical note edits (group 3). Arrows show participant's position and direction at 20 second intervals, dots show participant's hand's position while performing note edits. .

43.87 (SD = 40.10) note edits inside personal space in C_{fix}, these numbers reduced to 43.69 (SD = 34.69) in public space and 34 (SD = 25.37) inside personal space when it comes to C_{mov}. Although both numbers decreased, no significant differences were found between conditions.

Space & Territory Measures

To illustrate how participants used the space, based on the system-logged data, the positions and directions of participants' heads and musical note edits are plotted on a top view of the stage, see Figure 6.5 as an illustrative example of visual traces from arbitrarily selected group 3. Visual traces of all groups are shown in Figure 6.6. These figures were made based on system logged data, specifically, the arrows were participants' locations at 20-second intervals for ease of reading the diagram, and dots are the locations of participants' hands when making musical note edits. Research of table-top collaboration defines personal territory as a workspace close to the person and group territory as the central area or spaces between collaborators (Xambó et al., 2013; Scott et al., 2004; Scott & Carpendale, 2010). Following this definition, the area within a 0.6-metre radius of the participants' locations (locations here are at 1-second interval for higher accuracy) are coloured with different tint colour hues (red for participant A's personal territory, and blue for B's) to indicate territories. A distance of 0.6 metres was chosen as it falls into the range of close phase of personal distance, and permits participants to touch each other or the same music interface (Hall, 1966), most of the musical note edits also fell inside this range. The more intensely blue or red the area is, the more presence the corresponding participant had shown in that location. The overlap is coloured grey, indicating the appearance of both participants, which can be seen as group territory. Next, the following paragraphs will detail the territory-related results.

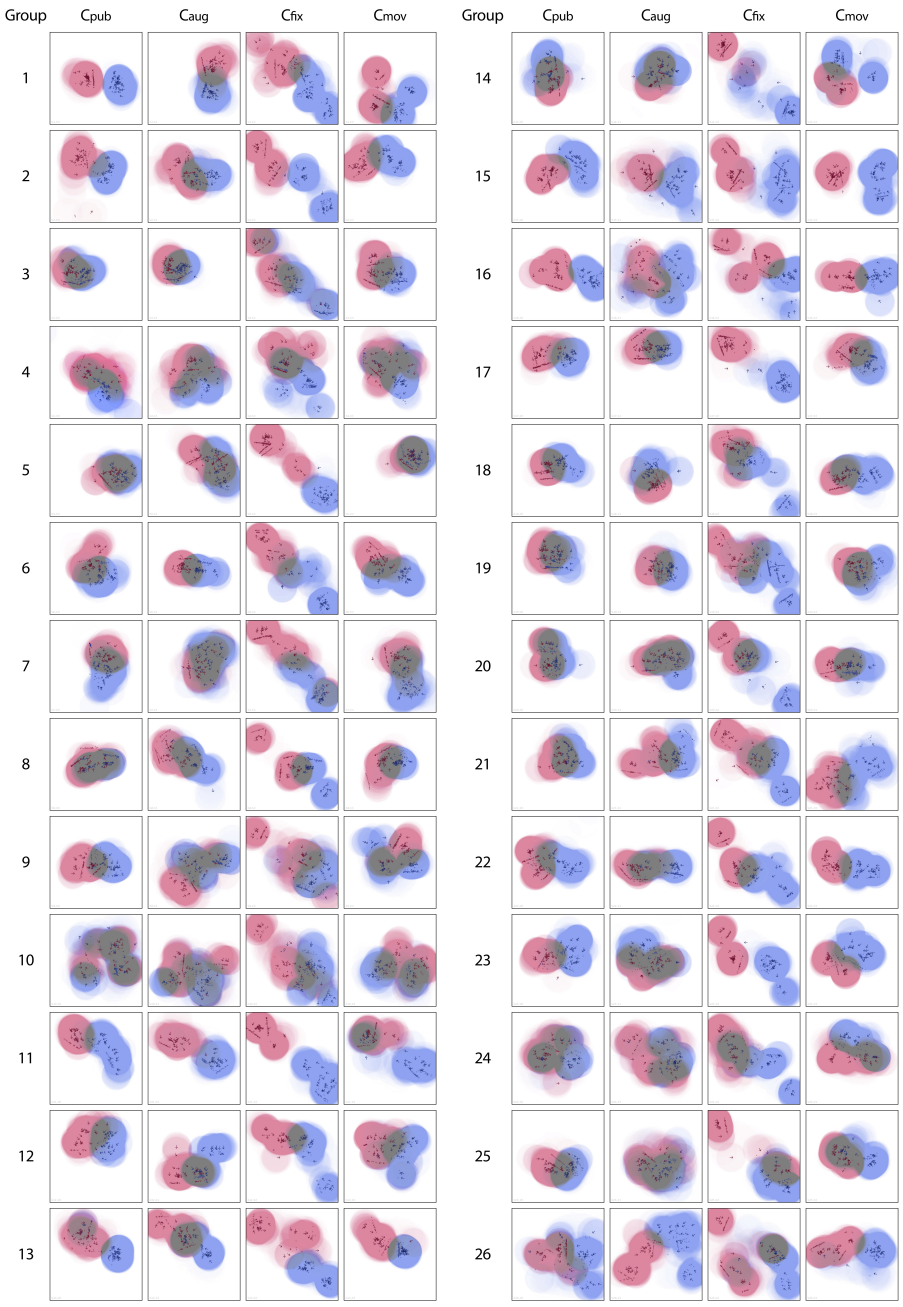


Figure 6.6: Visual traces - the participants' locations, directions and musical note edits shown on a top view of the stage (based on system-logged data of all groups).

(1) Sizes of group territory and group edits (edits fallen into group territory). By calculating the size of red/blue/grey area, the size of personal/group territory can be calculated. Specifically, participants formed an average of 0.3465 m^2 of group territory in C_{pub}, 0.4331 m^2 in C_{aug}, 0.2339 m^2 in C_{fix} and 0.3103 m^2 in C_{mov} (AA14 of Table 6.3). Results of the Wilcoxon Rank Sum Tests show that the size of group territory of C_{aug} is significantly larger than that of C_{fix} ($W = 491$, $p = 0.005236$), and near-marginal significantly larger than C_{mov} ($W = 430$, $p = 0.09421$). No significant difference was found between C_{pub} and C_{aug}. AA9 of Table 6.3 shows that participants had an average of 36.44 group edits in C_{pub}, which is significantly more than 17.50 of C_{fix} (Wilcoxon Rank Sum Test, $W = 1837$, $p = 0.001448$), and a near-marginal significantly more than 25.23 of C_{mov} ($W = 1621.5$, $p = 0.07839$). C_{aug} resulted in a higher average of group edits ($M = 43.04$), though not significantly higher than C_{pub}, it is significantly higher than numbers of C_{fix} and C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.01$). These indicate the spatial configurations of C_{pub} and C_{aug} are more friendly to group edits.

(2) Sizes of personal territory (AA15 of Table 6.3) and personal edits (edits fallen into personal territory; AA10). Participants formed a significantly larger personal territory in C_{fix} ($M = 0.7475$ m^2, $SD = 0.1801$) compared with all the other three conditions (Wilcoxon Rank Sum Test, all $p < 0.001$), and had significantly more personal edits in C_{fix} compared with other conditions (Wilcoxon Rank Sum Test, all $p < 0.01$; AA10). Similarly, larger size of personal territory was formed in C_{mov} and more personal edits were done in C_{mov} compared with C_{pub} and C_{aug} (all $p < 0.05$). No significant differences were found between C_{pub} and C_{aug}, neither in the size of personal territory nor in the number personal edits. To summarise, C_{fix} results in the largest size of personal territory and the largest number of personal edits, the metrics of C_{mov} follows, and C_{pub} and C_{aug} have the least, indicating C_{fix} led to a much looser collaboration, in which participants worked more independently, whilst C_{pub} and C_{aug}, on the opposite, led to more interactivities in the group territory.

(3) Average distance (AA16 of Table 6.3). Participants had an average distance of 2.19 metres between themselves and their collaborators in C_{fix}, this is significantly bigger than other three conditions (Wilcoxon Rank Sum Test, all $p < 0.001$). Namely, in the other three sessions participants worked more closely.

(4) Regarding the use of personal space, as shown in AA17, AA18 and AA19 of Table 6.3, in C_{fix}, participants had an average of 2.40 entries of personal space, on average, each entry lasting 73.07 seconds with total duration 128.60 seconds. For C_{mov}, the participants did 2.85 entries, each lasting 49.54 seconds on average, with a total usage time of 112.19 seconds, no significant difference

was found in the number of entries or in the sum usage time. However, the average duration of each entry of C_{mov} is significantly shorter than that of C_{fix} (Wilcoxon Rank Sum Test, $W = 1512$, $p = 0.008019$), indicating that personal spaces of C_{fix} were possibly more used for a longer independent creation.

6.3.3 Interviews

Post-task interviews with participants revealed more reflective insights into the spatial configurations. Around 41,000 words of transcription were transcribed and a thematic analysis of the transcription was undertaken. For more information about the thematic analysis, see (Braun & Clarke, 2006; Yin, 2017). The starting point of the thematic analysis was a reading through of the transcript, then an inductive analysis of the was carried out, collapsing relevant patterns into codes. Next, these codes were combined into overarching themes, which were then reviewed and adjusted until they fit codes well. As shown in Figure 6.7, in total, 650 coded segments, 24 codes and 3 overarching themes emerged from the thematic analysis. Different from the metrics included in self-report and activity assessments, which can be predefined, the process of interviews is less structured. Therefore instead of reporting the results following the types of the measures (i.e. awareness, communication, performance, and space and territory), here the results will be reported according to the three themes emerged from the interview. The three themes emerged are learning effects, reporting the spatial configurations, and reporting the LeMo system. Here only the first two will be reported as the last theme is related to the implementation and technical limitations, which falls out of the scope of this chapter.

Learning Effects

Members of 18 groups mentioned the effect of the session sequence. Specifically, 40 coded segments contributed by 24 participants were related to learning effects. Participants reported the sequence is an "important factor" (Participant 15A, hereafter abbreviated to P_{15A}). The first session was felt to be hard as they were "just being introduced to [the system and they were] still adjusting" to it (P_{5A}), trying to "[figure] out how the system was working" (P_{16A}). When they "were progressing into latter sessions, [they] felt easier to communicate and use gestures to manipulate the sound, being able to collaborate more, more used to the system" (P_{5B}), these changes led to a higher level of satisfaction and more enjoyment in later conditions. It should also be noted that interestingly P_{11A} and P_{11B} reported the sequence effect adversely, they enjoyed the first session more because "the first one was an element of surprise, a total surprise" as that was "the first time they were using the system". That feeling of freshness made

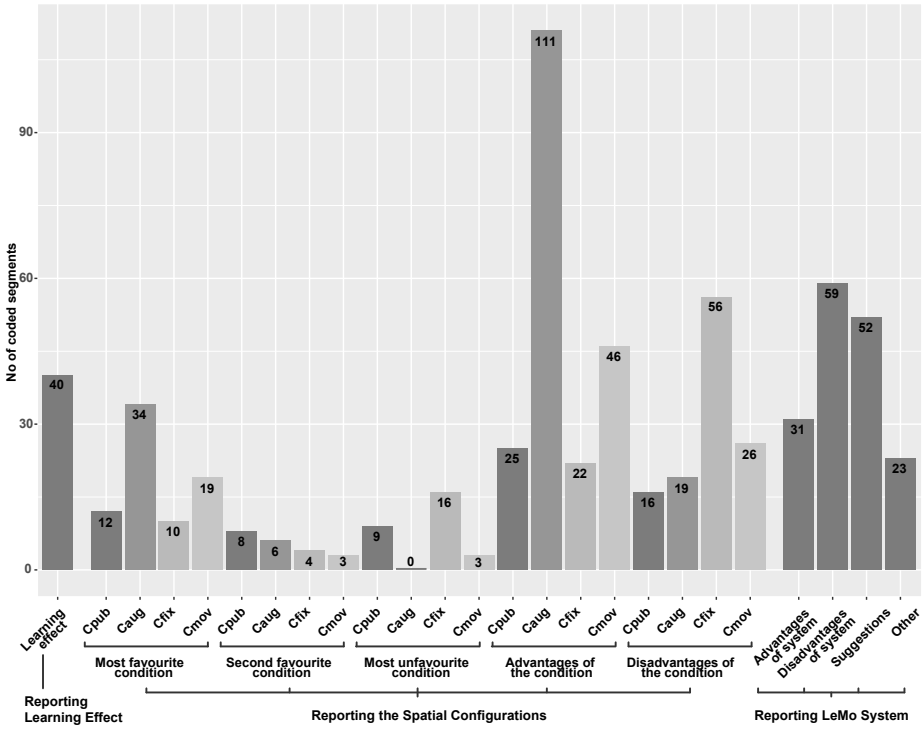


Figure 6.7: Ingredients of all the coded-segments of the interview, numbers of coded segments are shown along the bars.

that session more exploratory and more joyful to them. These learning effects might possibly affect the results of Post-Session Questionnaire and Comparison Questionnaire and thus should be well counter-balanced.

Reporting the Spatial Configurations

(1) **C_{pub} - Simple but can be chaotic.** Since there is no personal space, participants could, and had to hear all the interfaces all the time. In total, 16 coded-segments are about the disadvantage of this setting, some examples are: "a bit troubling" (P_{11B}), "music [was] always very loud" (P_{9A}), "it was global music, and there was someone annoying" (P_{2A}), "[they] not going to say anything" about disagreement because that possibly make them to be "rude" (P_{2A}). It was easier if there is something helpful "to perceive what I was doing, and not get confused with what [the collaborator] was doing" (P_{15B}), it was too "chaotic" (P_{20A}), "too confusing" (P_{22A} & P_{22B}), "annoying" (P_{25B}). They "cannot concentrate" (P_{25B}), "everything [was] open and quite noisy" (P_{26B}), they didn't "have the tranquillity to operating [the] sounds...everything [came] mixed, which [was] difficult to manage" (P_{22A}).

There were 25 coded segments from 14 participants reporting the positive side of the C_{pub}, some examples are: (i) pieces created in "personal space" might clash in a music way (P_{1A}), "better to work when knowing how it sounds all together" (P_{17B}), music pieces might match better; (ii) better for providing help to the other, according to P4A, they needed someone to lead them and thus the ability to hear all the work all the time was helpful; (iii) "space wise", namely, no space limitation, compared with having to work closer to "hear the sound well" (P_{12A}), C_{pub} does not have this constraint, they could choose to work "anywhere" (P_{24A}); (iv) "easier" to understand the condition (P_{6B}), fewer confusions when simply being able to hear all the things all the time (P_{13A}); (v) "collaborative wise" (P_{13A}), less separation, better collaboration compared conditions where "personal space" was provided (P_{3B}, P_{18A} and P_{18B}).

(2) **C_{aug} - Overwhelming preference.** There were 34 coded segments contributed by 24 participants favouring condition C_{aug}, higher than 12 segments contributed by 11 participants for C_{pub}, 10 segments contributed by 8 participants C_{fix}, and 19 segments contributed by 17 participants for C_{mov} (the sum of number of contributors here is greater than 52 as a few participant reported more than one favourite condition in the interview). The reason for the popularity can be concluded from the overwhelming 111 coded segments from 33 participants from 25 groups reporting the advantages of this condition, much higher than the number of segments reporting other conditions' advantages. C_{aug}'s advantages reported by participants can be grouped into 4 groups: **(i) Higher team cohesion and less sense of separation**. Participants reported that without the rigid personal space, they had to "work with the other person" (P_{6A}). With no rigid personal space, C_{aug} "forced [them] to collaborate more...because [they] had to stay very close" to compose music (P_{9B}). **(ii) An appropriate environment** for creativity, more consistency and convenience. As described by participants, it was "a middle point between personal space and no personal space" (P_{6A}), without even triggering something, "[they] could decide in a continuous way", "whether [they] were able to listen to the other sound sources or not, [and] to what extent [they] wanted to isolate [themselves]" (P_{16A}). Compared with having to hear all sounds in C_{pub}, this provided them with a "less stressing" (P_{4A}) context, and they could selectively move away to avoid "getting interrupted with the other" (P_{5B}) and overlapping music. Compared with C_{fix} and C_{mov}, being able to still "hear a bit of it in the background but not completely" (P_{20A}) was reported good as this kept them "up to date" (P_{9A}) and helped them to "tailor what [him/her] was making" (P_{22B}) to match the co-created music and to make something new and see if it "fit with" (P_{20A}) the old. C_{aug} provided them with "a little bit of personal space" although not a quite "defined thing" (P_{6A}), which provided the possibility "to work on some-

thing individually" and to "share work quite easily" (P_{20A}). **(iii) Easier to identify sounds**. Participants reported it was easier to "locate the source of the sound" (P_{16A}) and "perceive what [they were] doing" (P_{15B}), these factors then helped them "understand instruments better" (P_{7B}) and "not get confused" (P_{15B}); **(iv) More real**. Quite interestingly, instead of C_{pub}, which simulates the acoustic attenuation in the real world, C_{aug} was reported to be similar to the experience in the real world. Participants reported in C_{aug} "if [they] want to hear something, [they] just come closer, like in the real world" (P_{11B} & P_{11B}), "feeling like the real-time experience (P_{26B})".

It should also be noted that, along with these 111 coded segments reporting the advantages provided by C_{aug}, there are 19 segments reporting its limitations. These limitations fall into three groups: (i) a preference "to hear all the instruments all the time" in C_{pub} (P_{26B}); (ii) C_{aug} might lead to "another type of compositions" and "influence the piece" (P_{16B}); and (iii) not being able to hear all sounds led to a feeling of separation (P_{18A}).

(5) **C_{fix} and C_{mov} - Resemblance and differences.** Regardless of the mobility, the personal space provided by C_{fix} and C_{mov} share the same characteristics. Not surprisingly, the participants reported many common advantages and disadvantages shared by both conditions, including: The addition of rigid personal space was described as an "added advantage" (P_{7A}), it made it "easier to perceive what [themself] was doing and not get confused with what [their collaborator] was doing" (P_{15B}), provided them with a chance to "isolate themselves to create their piece" (P_{22A}), and to "think about something to add" (P_{9A}), helped them to "develop their own ideas" (P_{8A}). As a result, they used personal space "a lot [and] used [their] own creativity much more comparing with [other two sessions]" (P_{3A} & P_{3B}).

Common disadvantages reported include: The rigid form led to segmentation, and a feeling of being "forc[ed]" to work on something individually (P_{6A}), making them "forget" the collaboration and collaborator (P_{8A} & (P_{12A}), resulting in less collaboration, less "communication happening" (P_{7A}), "lost the idea of the joint music piece" (P_{16A} & P_{16B}), and as a possible result, each other's music pieces did not fit when brought up (P_{9A}). P_{4A} reported they were not familiar with music, and thus they "needed somebody to lead" them, so preferred to hear sounds all the time. Besides, P_{24B} reported that the visual personal space made the stages look "messy"

Differences between C_{fix} and C_{mov} - In total, 46 coded segments (from 26 participants) were reporting C_{mov}'s advantages and 26 segments (from 12 participants) reporting its disadvantages, compared with 22 segments (from 14 participants) and 56 segments (from 34 participants) for C_{fix}, indicating in general participants thought C_{mov} better than C_{fix}. Some example insights behind

the preference are: C_{mov} functioned like a "mute button" (P_{4B}), which could be used anywhere (P_{7B}), enabling them to "move around", work "closer...and see each other's things" and thus led to "more collaboration" between them (P_{1B}). Though C_{fix} had no advantages on these aspects, the location at the opposite corners provided a more "personal feeling" and a higher sense of belonging(P_{22A} & P_{22B}). Walking to the corner to access personal space was not a big issue for P_{7B} & P_{7A} as "the boundary is small". Besides, the relatively far distance also helped to "prevent [them] from clashing" (P_{7A}).

6.4 Discussion

Recall that Study II found that the addition of personal space located at the opposite side of the public space led to a shrunken size of group territory, fewer group note edits, a larger size of personal territory, more personal note edits, a larger average distance between collaborators, and fewer times of paying attention to collaborators. These negative impacts might probably due to that, in that study, the personal spaces distributed on the opposite side of the group space resulted in a larger distance between participants. As such personal space with different features (e.g. gradual boundary in C_{aug}, mobility in C_{mov}) might reduce, or even minimise, these negative effects.

One of the key research interests of this chapter is how to provide personal space to support creative collaboration, particularly CMM, in Shared Virtual Environments with minimised negative impacts on collaboration. Next, based on the results presented above, the following sections will firstly discuss the necessity and the impacts of introducing personal space in C_{aug}, C_{mov} and C_{fix}. Where appropriate, comparisons between conditions will be made.

6.4.1 Necessity of Adding Personal Space

Previous research has highlighted the necessity of providing personal space with fluid transition to public space during collaboration (Scott et al., 2004; Shen et al., 2003; Sugimoto et al., 2004) and suggested that people did construct public and personal space in collaboration in VEs (Men & Bryan-Kinns, 2019). Though it is undeniable that no personal space did bring some benefits – Specifically, when no personal space was available, C_{pub} was reported to provide the experience of the least difficulty of tracking the collaborator (CQ1 in Table 6.2), the strongest sense of collaborator (CQ2), best communication quality (CQ3), the least difficulty to cooperate (CQ6), making C_{pub} to be the simplest one among these four configurations for participants to learn and get used to. The issues of having no personal space are clear. Firstly, especially for the music making task

in this study, according to the interview results, participants reported the background can be messy to develop own ideas, their creativity requires a quieter and more controllable environment. Considering individual creativity forms an important part of the collaborative creativity, providing an appropriate environment became crucial. The personal space provided in C_{aug}, C_{fix}, and C_{mov} functioned like a "less stressing" context, within which, participants could better "understand instruments" and not "get confused". Secondly, participants need an opportunity to develop their own ideas. From the interview results, having personal space was reported to be "an added advantage", it helped to promote their own creativity, which was then combined and contributed to the joint piece. This echoes the findings in Study II, that providing personal spaces is helpful as it provides a chance to explore individual ideas freely, which then added an interesting dynamic to the collaborative work. Though some disadvantages of having personal space – less communication, higher isolation and being messy – were also reported, most of these limitations were possibly the result of introducing rigid visible personal space, and C_{aug} was founded to have addressed these limitations well (details will be discussed in the following section).

6.4.2 Impacts of Introducing Personal Space - C_{aug}

Measures in Table 6.3 show that C_{aug} significantly differs from C_{fix} and C_{mov} in many ways. When both significant differences (p <0.05) and marginal-significant differences ($p < 0.1$) are considered, compared with C_{fix} and C_{mov}, C_{aug} saw a smaller personal territory (AA15) and a bigger group territory (AA14), more mutual modifications (AA8), more group edits (AA9) and fewer personal edits (AA10), a larger distance between collaborators (AA16), more times of paying close attention (AA2) and a longer time of paying close attention (AA1). All these indicate that compared with the rigid personal space provided in C_{fix} and C_{mov}, the personal space served by augmented acoustic attenuation in C_{aug} enabled a closer collaboration, H3 is therefore supported. C_{aug}'s advantages are shown in four ways. Next, each will be discussed in detail.

Minimised Introduced Impacts

C_{aug} is similar to C_{pub} in many ways, e.g. both have no visual boundary for personal spaces, or visual triggers for personal space. As a possible result, similar territorial patterns were formed in these two conditions (Figure 6.5). Not surprisingly, no significant differences were found in most of the types of measures listed in Table 6.1 and 6.3. The only differences revealed in these two tables lie in the communication measures. AA1, AA2 in Table 6.3 show

significantly more occurrences and longer duration of close attention were paid to collaborator in C_{aug} than in Cpub).

Few mild differences indicate that C_{aug} is not very different from C_{pub}, i.e. they received similar results in many metrics. From another perspective, fewer differences between C_{pub} and C_{aug} indicate that the limitations of adding personal space identified in Study II have been successfully minimised. Specifically, the size of group territory and the number of group edits maintained similar numbers, the means of C_{aug} are even greater, though not significantly (AA14 and AA9 in Table 6.3). C_{pub} and C_{aug} saw a similar size of personal territory and personal edits, and similar average distance (respectively, AA15, AA10, AA16 in Table 6.3), and C_{aug} even saw more close-attention paid to each other (AA1 and AA2 of Table 6.3). All these similarities indicate that by introducing a personal space with gradual and invisible boundary, these identified disadvantages of introducing personal space have been successfully eliminated. Possible reasons can be that C_{aug} managed to provide an interaction experience that is similar to C_{pub}. Recall that in conditions C_{PI} and C_{PV} of Study II, to access the personal spaces located at the opposite side of the public space, participants had to drift apart, which might have influenced the their spatial locations, changed the formation of group/personal territory they formed and the average distance between collaborators changed, and territoriality-based interaction (group/personal edits) changed. Here in C_{aug}, by enabling participants to use personal space anywhere inside the stage with no specific triggers needed, C_{aug} managed to provide a user experience as similar as possible to C_{pub}. The second reason is more related to the impacts on subjective experience, in C_{aug}, by making the personal space invisible and gradual, the isolation and difficulty of coordinating that introduced by the additional rigid personal space was minimised. E.g. in the interview, participants reported C_{aug} provided a proper level of group work as a working context, making easier to create new that matches the old.

Enough Support for Creativity

PSQ11 (Table 6.1) questioned the support that each condition gave to individual creativity. Although no significant differences were found, C_{aug} has a higher mean rating, possibly indicating a higher level of support. It should be noted that all the questions in PSQ were phrased either positively (PSQ1, PSQ2, PSQ 9, PSQ 10, PSQ 11) or neutrally (the rest), with no negative statements, which might have affected participants' ratings positively. However, this imperfection has a limited influence on this study, because PSQ results are mostly used for comparisons between conditions, which are affected equally due to all of the

conditions using the same phrasing. More insights regarding C_{aug}'s helpfulness are revealed by the thematic analysis, according to which, C_{aug} provided both "an appropriate background" with which participants felt "less stressed" and were able to "tailor" the individual composing to match the co-work, and a space personal enough to "work on something individually". Not many differences were found between C_{pub} and C_{aug} from the results of PSQ, CQ and AA, indicating C_{aug} provides a very mild solution to introduce the personal space, with limited impacts on people's and collaboration and use of space and territory being introduced, whilst still providing sufficient support for individual creativity during collaboration. Thus H2 is validated.

Impacts on Collaboration and Use of Space and Territory

According to measures of attention (AA1, AA2 in Table 6.3), compared with other conditions, C_{aug} saw participants paying more close attention to their collaborator. Possible reasons for this can be found from the thematic analysis and other measures of Activity Assessments (Table 6.3). Compared with realistic acoustic attenuation in C_{pub}, C_{aug}'s augmented acoustic attenuation setting forced or prompted people to work closer in order to hear each other's work, as reported by some participants. Compared with adding personal space with visible rigid boundary, by enabling participants to "decide" whether to hear other's work or not "in a continuous way", an invisible gradual boundary in C_{aug} led to less separation, and higher consistency between personal and public space, which matches the finding that people would like to be able to smoothly shift their artifacts from personal to public with intermediate shades in-between (Greenberg et al., 1999). Compared with rigid personal spaces in C_{fix} and C_{mov}, C_{aug} saw more mutual note modifications, more group note edits, and larger group territory, a closer average distance between collaborators (respectively, AA8, AA9, AA14, and AA16 in Table 6.3), all of these indicate that C_{aug} saw a less separated collaboration than C_{fix} and C_{mov}. Compared with the three levels of privacy provided in UbiTable (Shen et al., 2003) and the binary levels of privacy provided in SharedNote (Greenberg et al., 1999), the step-less sonic privacy provided by C_{aug} in this study possibly managed to better echo the suggestion that a boundary between personal and public space should be provided with gradations in subtle and lightweight ways to enable a fluid shift (Greenberg et al., 1999), H3 is therefore supported.

Popularity Over C_{aug}

The code "advantage of C_{aug}" has 111 coded segments, which is far more than the segments other codes have. Thirty-five coded segments are "most favourite

- C_{aug}", higher than all other three conditions. These indicate C_{aug} is the most popular condition. This can also be partially verified by the preference measure. Specifically, C_{aug} has the highest preference rating in PSQ9 (Table 6.1), and more participants chose C_{aug} as the setting they most enjoyed in CQ7 (in Table 6.2). Reasons behind this popularity are probably due to its unique advantages, which as reported by participants, includes: (i) higher team cohesion and less sense of separation; (ii) an appropriate environment for creativity; (iii) easier to identify sounds and iv) more real (though in fact, C_{pub} is more real from the perspective of simulation). These features of C_{aug} made it provide better support for collaborative creativity and therefore led to its popularity.

6.4.3 Impacts of Introducing Personal Space - C_{mov} and C_{fix}

This subsection compares C_{fix} with C_{mov}, the former provides personal spaces at the corner of the stage, whilst the latter enabled participants to use personal spaces anywhere (check Figure 6.1). The clear, sole difference between these two conditions is the mobility of personal space. The following sections will firstly discuss the impacts of adding personal space in C_{mov} and C_{fix} and then compare the two.

C_{mov} - Impacts of Adding Movable Personal Space

In C_{mov}, participants could pop up the personal space anywhere in the stage. In this way, personal spaces was provided with mobility. By doing so, several aforementioned negative effects found in Study II were reduced. Specifically, these include the time spent paying close attention to collaborator, the times of paying close attention to collaborator, size of group territory, the average distance, (respectively, AA1, AA2, AA14 and AA16 in Table 6.3). However, some significant differences remained, participants still had significantly fewer mutual note modifications, marginal-significantly fewer group edits and significantly more personal edits after personal space being introduced in C_{fix} and C_{mov} when being compared with C_{pub} (respectively, AA8, AA9, and AA10 in Table 6.3). This can also be verified bey the interview results. Compared with C_{aug}, participants reported a higher sense of isolation in C_{mov} and C_{fix}, both of which provided rigid-form personal spaces. Namely, C_{mov}, by making the personal space available anywhere in the stage, managed to drag participants closer and saw a similar group territory when being compared with C_{pub}. However, participants' behaviour was still affected in many ways. Participants were still being separated by the visible wall to some extent, which can be seen as a disadvantage of adding visible, solid personal space. In other words, C_{mov} did

better than C_{fix} in minimising the negative impacts of adding personal space, but not as good as C_{aug} .

C_{fix} - Impacts of Adding More Rigid Personal Space

C_{fix} provided a much more inflexible personal space, which influenced participants' behaviour in many ways (see the significant differences between C_{pub} and C_{fix} in Table 6.1 and Table 6.3). Not to mention participants' polarised ratings on C_{pub} and C_{fix} in CQ1, CQ2, CQ3, CQ9 of Table 6.2: Significantly many participants reported the least difficulty of tracking collaborator's activities (CQ1), the strongest sense of their collaborator's presence (CQ2), the best communication quality (CQ3), the least difficulty of cooperating with collaborator (CQ9) happened in C_{pub}, whilst C_{fix} was considered oppositely by significantly many participants. Their dislike of C_{fix} can also be seen in the interviews, in which the number of coded segments favouring C_{fix} and the number of segments reporting its advantages are the lowest, whist the number of segments disfavouring it and the number of segments reporting its disadvantages are the highest among the four conditions.

C_{mov} vs C_{fix} - Mobility vs No Mobility

In C_{fix}, to access personal spaces at the corners, participants needed to physically walk away from the centre and head to the corner, which might be the reason that C_{mov} saw a closer average distance between collaborators than C_{fix} (AA16, Table 6.3). This greater distance in C_{fix} possibly resulted the significantly larger size of personal territories (AA15) and more personal edits (AA10) in C_{fix}. On the contrary, the closer distance in C_{mov} created more chances for participant to pay or draw attention between each other, as a result, significantly longer time was spent paying attention to collaborators (AA2, AA4 in Table 6.3). With a closer average distance and more attention paid to each other, participants reported they had a marginal-significantly better quality of communication in C_{mov} (PSQ3 of Table 6.1). On the other hand, with participants being far away from each other and less chances for contact in C_{fix}, significantly many participants reported that they had the worst communication quality in C_{fix} (CQ8, Table 6.2). C_{mov} was also rated by much fewer participants to be the least enjoyable condition than compared with C_{fix} (CQ7, Table 6.2). Besides, C_{fix} also led to a reduced sense of collaborator's contribution (CQ6 in Table 6.3). As a possible result, C_{mov} saw a significantly more satisfying work output (PSQ4, Table 6.1).

The thematic analysis results also echo these findings. More coded segments are reporting C_{mov}'s advantages compared with those reporting C_{fix}'s,

and more coded segments reporting C_{fix}'s disadvantages than those reporting C_{mov}'s. Also, more coded segments are favouring C_{mov} compared those favouring C_{fix}. Participants reported it was good to be able to use personal space anywhere in the stage as it resulted in a closer distance, which led to more collaboration and made it possible to see each other's work. To conclude, compared with C_{fix}, C_{mov} resulted in a better communication quality, produced better feeling of collaborator's contributions, and was rated more enjoyable, thus it saw a closer collaboration and produced a more satisfying result, H1 is therefore supported.

6.4.4 Key Findings

In summary, the following are key findings from the results discussion above:

- Having personal space seems to be necessary for SVEs supporting collaborative, creative activities as it supports individual creativity, which is an important element of the collaborative creativity.
- C_{aug} minimised the negative impacts introduced by adding personal space (previously identified in Study II) better than C_{fix} and C_{mov}.
- C_{aug} was found to have the most minimal impacts and even to influence the attention between collaborators positively. Both C_{fix} and C_{mov} produced a more alienated collaboration, indicators of which include significantly bigger personal territory and more personal edits, and significantly fewer mutual note modifications and fewer group edits, significantly lower sense of collaborator's activity. Additionally, C_{fix} saw significantly more note edits, and significantly less ordinary attention paid between collaborators.
- Providing personal space with a fluid boundary is preferable, it provides enough support for individual creativity with the minimal cost, and can even lead to a closer collaboration (specifically, greater close attention was paid between collaborators in C_{aug}).
- Compared with stationary personal space, personal space with mobility led to better communication, produced a better feeling of collaborator's contribution, had a higher rating in enjoyment, and produced a more satisfying output, and thus C_{mov} supported collaboration better than C_{fix}.

6.5 Design Implications

Based on the 4 key findings made above, we suggest 5 design implications for SVEs focusing on supporting CMM:

(1) SVEs supporting CMM should come with personal space, as it provides essential support for the development of individual creativity, which forms a key part of the collaborative creativity. This is especially needed when the output of the task is more disruptive and mutually exclusive(e.g. audio related tasks), co-workers need a space where they can think of and develop own mind and work.

(2) For audio-related tasks (e.g. collaborative music making), manipulating acoustic attenuation as personal space is an effective way to support both individual creativity and collaboration. It allows users to shift between personal and public working space continuously by adjusting their relative distance. It comes with light-weight form, functions as a personal space well, and can increase close attention paid between participants. Besides, based on the findings, it does not introduce significant amount of negative impacts whereas rigid personal space does.

(3) Manipulating the level of augmentation (e.g. the augmented acoustic attenuation in this study) may change the level of "personalness" – feeling of being personal. In the C_{aug} condition of this study, participants adjusted their distance between themselves and collaborators to obtain a different level of being personal, e.g. total isolation can be achieved if both participants are working with a distance greater than 1.2 metres, beyond which point, the volume of objects drops to 0. Similarly, when personal spaces are provided with gradual and adjustable boundary, manipulating the parameter of the boundary (e.g. the degree of augmented attenuation) can impact the level of "personalness" and therefore adjust the impact of introducing personal space. E.g. the augmented attenuation can be set to a very low level to ensure sounds drops more naturally if an extremely minimal impact is being pursued. As such, it might be useful to add an enabler that allows users to adjust the level, allowing them to shift between having a "very personal" space with total isolation where they could not hear nor see each other's work), and having no personal space when they have to work together. In this way, users are enabled to manipulate the level between "personalness" and togetherness continuously, which is useful to allow users to develop own ideas and work together to tailor own work into the collaborative piece. Compared with adjusting "personalness" by distance in C_{aug}, adjusting it by changing the parameter might also be useful as co-workers can then stay anywhere whilst still being able to adjust the "personalness" the personal space provides.

(4) When it is hard or impossible to fit in a gradual, light-weight personal space due to the type of the task, a rigid-form personal space can be considered. And it is better to provide rigid personal space with mobility, as the mobility feature gives users more freedom for accessing the personal spaces, and produces

a better user experience with fewer negative impacts on the collaboration compared with rigid personal space without mobility. This implication also echoes the proposal raised in Study II.

6.6 Chapter Summary

This chapter has detailed an experiment exploring how four different spatial configurations impact the collaboration differently. Both quantitative and qualitative data were demonstrated and analysed, comparisons between conditions were made where necessary, differences of the impacts posed by the four conditions were found, key findings have been concluded. Specifically, parts of the key findings are that the augmented attenuation can serve the necessity of individual activities well, with minimal negative impacts on collaboration and even coming with a bonus point (more close attention between participants). Results also show that a rigid personal space with mobility serves users' needs better and is preferable over a non-mobile one. Based on the findings, and five design implications were given for SVEs supporting CMM. The next chapter will compare the three studies accordingly.

Chapter 7

Overall Discussion

This chapter does an overall discussion based on the findings of the three studies, and tries to correlates the findings of the three studies. This chapter is structured as follows: Section 7.1, 7.2, 7.3 include further discussions based on the three studies. Specifically, Section 7.1 discusses the role and impact of embodiment – a consistent feature available in all the three studies; Section 7.2 reflects the way to construct privacy during collaboration, i.e. how to balance privacy and openness in collaboration; In Section 7.3, the visual approach in Study I and the auditory approach in Study III will be compared and reflected. After these three sections, Section 7.4 will conclude the relevant findings of the three studies into implications for designing SVEs with a focus of CMM. Finally, Section 7.5 discusses how the these findings might contribute to design of SVEs beyond music making.

7.1 The Role of Embodiment

It would be beneficial to discuss the role of embodiment further, as this is a feature shared by both versions of LeMo, and according to work related to embodiment reviewed in Section 2.2.3, it might have greatly impacted the three studies. The following subsections discuss how the embodiment might have influenced the results.

7.1.1 Embodiment's Impact on Work Identity

As aforementioned in Section 3.5.2 (modelling and animating avatars), during collaboration, awareness of both who is contributing, and what they are contribution is important, and cues of identity play a key role in supporting such an awareness. Providing cues of identity can significantly affect participants' con-

tributions to collaborative music making (Bryan-Kinns & Hamilton, 2012). Similarly, the importance of identity in collaboration has been stressed by Gutwin & Greenberg (2004) in "who" category of their framework for understanding workspace awareness. However, conversely, the findings of Study I failed to indicate that the visual feature *Work ID* had any significant impacts on any of the measurements. We argued the reason for this might be that avatars had provided efficient information for constructing these awareness. E.g. in terms of construction of workspace awareness, the *Work ID* in study I mainly provides support on authorship, which only covers one element of the category "who". By contrast, avatar provides support on all the three categories of workspace awareness (who, what and where). According to (Gutwin & Greenberg, 2004), elements of workspace awareness include who, what and where, each include some specific questions. According to this, Table 7.1 includes a comparison between Work ID and the avatars used in LeMos. As shown in Table 7.1, avatar can not only provide support for "who" (presence, identity, authorship) but also for what (action, artefact) and where (location, gaze, view, reach). By contrast, *Work ID* could only provide support in "who" category. Specifically, by seeing the colour difference of work, collaborators might be indirectly supported on sense of co-presence, the identity of the collaborator and the authorship of the work. All these elements had been fully supported by avatars. Instead of relying on the differences of work done to convey coordination information, avatars provided sufficient and direct support for many aspects of workspace awareness. As a results, participant might have a stronger feeling of co-presence by seeing their collaborator's avatar compared with seeing what changes their collaborator did.

To conclude, avatars could fully cover and provide even better support for the three elements in "who" category, and even provide support for "what" and "where", which are adequate support for building a proper level of workspace awareness. This might be the reason that *Work ID* played an insignificant role in the collaboration in Study I. In other words, identity of members and work are not necessarily unimportant in collaboration in SVEs, it is just the support from the *Work ID* design feature (the button hues changes) was surpassed by avatars to a large extent, and avatars were available in all the conditions of Study I.

7.1.2 Embodiment's Impact on Awareness

Embodiment was found to promote sense of presence, co-presence and sense of other's activities. For example, one of the findings of Study I is that much fewer annotations were used to convey presence compared with those used in Bryan-

Table 7.1: Fitting *Work ID* in Study I and avatars provided in all three studies into Gutwin & Greenberg's Table of Elements of Workspace Awareness.

Category	Elements	Specific Questions	Support from *Work ID* of Study I	Support from avatars of Study I, II & III
Who	Presence	Is anyone in the workspace?	Indirect support	Direct support
	Identity	Who is participating? Who is that?	Indirect support	Direct support
	Authorship	Who is doing that?	Direct support	Direct support
What	Action	What are they doing?	No support	Direct support
	Intention	What goal is that action part of?	No support	No support
	Artefact	What object are they working on?	No support	Direct support
Where	Location	Where are they working?	No support	Direct support
	Gaze	Where are they looking?	No support	Direct support
	View	How much can they see?	No support	Indirect support
	Reach	How far can they reach?	No support	Direct support

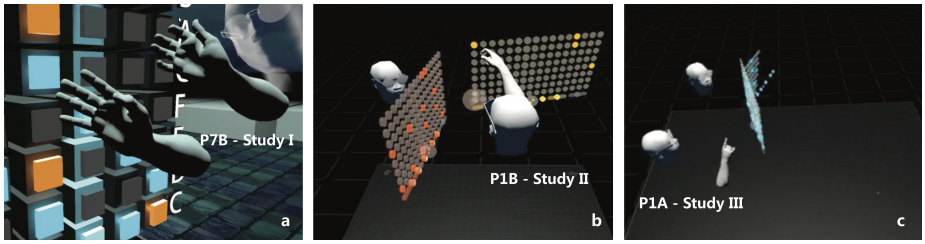


Figure 7.1: Participant 7B in Study I is waving hands to attract the collaborator's attention (a); Participant 1B of Study II is using hands to draw the collaborator's attention to a certain place (b); Participant 1A of Study III is using hands to draw their collaborator's attention to a specific part of the interface (c).

Kinns (2011). The reason might be that the avatars have provided sufficient cues on "who" is performing "what" at "where" (see Table 7.1), so participants did not need to rely on other cues to form a solid sense of each other's presence. Results of Study II show that C_{PI}, which breaks the continuity of avatars by setting opaque walls to separate personal space with public space, was more difficult for collaborators to track each other's activities and saw a significantly shorter duration of attention paid to each other compared with other two conditions (time length of attention measure in Table 5.3), and saw a reduced sense of collaborator's presence (CQ3 of Table 5.2 and thematic analysis). By contrast, other conditions C_P and C_{PV} of Study II, by either providing no personal space or transparent personal space, kept the avatars visible throughout the session, and provided better support on maintaining these senses. Hence it is suggested to keep collaborators' avatars always publicly visible for SVEs supporting collaboration.

Study III, with this finding bear in mind, made all the personal space either

invisible or transparent and kept avatars visible throughout the session. Though the avatars in Study III worked well, there is still room for improvement. P_{21B} (participant B in group 21) of Study III expressed the willing for a more detailed and holistic avatar that might also contain body and legs, which, according to them, might be helpful to better locate each other, and track the activities. Similarly, P_{10A} of Study II expressed that a full-body avatar might alleviate their concern of possible collisions. Apart from a more complete avatar, P_{15A} of Study II wanted to be able to personalise the avatar, e.g. having the ability to choose sexuality of avatars. Avatars with more body parts and greater potential to customise might be two future directions for improvements for the avatar feature of LeMo. For other SVEs supporting collaboration, these two ways to improve avatars can also be considered.

7.1.3 Embodiment Supports Referencing

Bryan-Kinns (2011) identified the frequent use of annotation as a localisation cue (mainly by drawing arrows), but Study I only saw one such usage. One reason might be that collaborators had more "natural arrows" – their avatar hands – to use. By reviewing the screen recordings, there are numerous examples in all the three studies. Figure 7.1 shows three cases where hands were used to draw each other's attention and to refer to a specific position. Compared with drawing arrows to inform localisation, avatar hands seem to be easier to use, more flexible (where to point can be changed easily) and even more extensible (hands have many gestures while arrows are static). As such, supporting avatars, especially avatar hands should be strongly encouraged for SVEs supporting collaboration.

7.2 Balancing Privacy and Openness

Privacy gives people the ability to seclude themselves and schedule their activities independently. As briefed in Section 2.4.3 (privacy in collaboration), in the design of any multi-user system, privacy has long been an issue, e.g. how to visually representing users and the privacy states of objects in SVEs (Butz et al., 1998). Here in Study I, II and III, the "privacy" can be divided to three main aspects: (i) visibility, (ii) hearability, (iii) accessibility of virtual artefacts. There are also other modalities apart from visibility and hearability, e.g. capability of being smelled, however LeMos used in the three studies did not involve other modalities and currently the majority VEs do not include other sensations either. Therefore, this section only focuses on a discussion of the privacy of visual and audio modality, plus the accessibility of virtual artefacts.

One traditional way to differentiate public and private things in daily lives

and some software is simple and direct, which is "visibility = public = accessibility", and "no visibility = private = restricted or no accessibility". However, this does not work so well when it comes to VEs, as blocking sight of unshared virtual objects breaks the illusion that all the collaborators are present in the same VE and manipulating the same set of objects. For example, if an object exists in one player's VE, it should also exist in other players' VE, since SVEs are usually simulating an illusion that all players are present in the same VE, interacting with the same set of virtual objects. In a natural sense, everything in the VE should be the same for all the users, including their appearance and properties. For instance, people should hear, see and sense the same set of virtual objects inside the virtual environment if the system is trying to convey the concept that they are collaborating in the same environment. Hereafter we refer this rule as "same perceptions from same perspective". One key finding of Study II is that the visibility of other's actions and the virtual artefacts is preferred and important to maintain team awareness. Hence, a way is needed to control the accessibility and privacy of virtual artefacts, whilst still keeping them visible. To do so, a natural way is preferred to reduce the possible negative impacts on the realness of VEs. Next, several possible ways will be discussed.

7.2.1 Ways to Construct Privacy in VEs

As aforementioned in Section 2.4.3 (Privacy in Collaborations), Butz et al. (1998) suggested two tools to manipulate the privacy condition of objects in augmented reality: *privacy lamps* and *vampire mirror*. However, though the interaction of the two methods (positioning the lamp and mirror) are natural, the way that the *privacy lamp* and *vampire mirror* work are very "magical" and artificial. Such items have never existed in human lives, i.e. there are no lamps that omit an object by shining brightness on it nor are there mirrors that can selectively reflect objects. Therefore, these two tools violate the rule *same perceptions from the same perspective*, making both tools very artificial and unreal.

Some insights about the negative impacts of using artificial tools have been collected during the process of the three studies. Recall that in condition C_{PI} of Study II, the walls separating the personal space from the public space are only transparent to the owners of the personal space, causing that the walls were selectively transparent to the owner of the personal space. In other words, wall transparent to owner user A became opaque to user B, which made the wall less real. As a possible result, it was harder for participants to understand this feature during the tutorial session, compared with explaining the wall of C_{PV}, which seems to be more strait-forward and natural by being consistently

Table 7.2: Implications for creating privacy and manipulate accessibility.

Channel	Rigid way	Gradient/step-less way
Visual privacy	Using opaque objects to block sights (e.g. walls)	Lower visibility space (e.g. fog)
Sonic privacy	Using soundproof objects to block sounds (e.g. walls)	Using sound attenuation to block sounds
Accessibility	Transplant real accessibility concepts to VEs (e.g. locks and keys)	Control the accessibility with distance, the nearer, the more accessible

semi-transparent for all users. Also some participants reported the wall in C_{PI} was hard to understand. Hence, instead of "artificial" tools, it seems better to use a more natural way, which imitates the reality to constrain the hearability, visibility, and accessibility of virtual objects.

To manage the hearability of virtual objects, as concluded in Table 7.2, two ways have been practised in Study II and Study III: **i**) Using rigidly-sound proof objects, like the virtual soundproof wall in Study II. The practice of Study II has shown the soundproof wall concept is easy for users to understand and get used to. The double-side soundproof concept in Study III seems to be even easier to understand compared with one-way soundproof wall in Study II, reasons might be that the former is more common in daily lives. **ii**) Using gradient spatial acoustic attenuation, i.e. the acoustic attenuation in C_{aug} in Study III, this way imitates the natural physical phenomenon of volume dropping and hence it might introduce less negative impacts on the realness of the VEs and is easier for users to grasp.

Similarly, to manage the visibility of objects, two directions can be considered (Note this suggestion is very preliminary, as it is not directly drawn from the findings of the three studies): (i) Rigid way – using opaque blocking objects, i.e. an opaque wall or simulated opaque gas; (ii) Gradient way – controlling the visibility with distance, e.g. imagine in a foggy morning, things getting more and more invisible with distance increasing. This phenomenon could possibly be simulated to control the visibility of objects.

Likewise, to manage accessibility of objects, two ways have been practised by the three studies: (i) Completely blocking certain users from a certain area or objects. For example, in Study II, participants were asked not to enter each other's personal space. In this way, the user with the access of the area has exclusive accessibility of that area and objects inside that area. There are many concepts of controlling accessibility in the real world that could possibly be transplanted to VEs, e.g. keys, padlocks, doors, voice-print, fingerprint and so on. (ii) Using distance as a tool to control the accessibility of objects (Note this

is a preliminary suggestion as it is not directly drawn from the three studies). By the very nature, without movements, a physical object is out of reach and accessibility if its distance to a user is longer than the user's arm. In VEs, if only natural interaction is provided (i.e. what people can perform in reality, like grasp by hands, no remote control), then people naturally have less access to objects that are more far away.

7.2.2 Balancing Privacy and Openness

Finding a balance between privacy and openness is important. A total openness was provided in all conditions of Study I, C_P of Study II and C_{pub} of Study III. In these conditions, users have accessibility to all music interfaces, and can hear and see all the objects. According to results of Study II, being able to see each other all the time can lead to a better feeling of working together, and a better awareness of the activities of collaborators. Having each other in the same space enabled them to provide feedback constantly and easily. These two features helped individuals to plan, orient and coordinate their own work to fit the group activities. However, the deficiency is also clear. It became totally a mess as collaborator's work might interfere with each other's in the same space. Results of Study II and III have shown the necessity of having personal space. This echos the argument that privacy is needed during collaborative work to obtain "the freedom to be left alone" (definition of privacy by Warren & Brandeis) 1890. The opposite extreme of no privacy is to make everything private, including visibility, hearability and accessibility of objects. C_{fix} of study II can be an example of such setting. The disadvantages are clear, according to findings related to C_{fix} in Study II, such setting might possibly lead to reduced sense of collaboration, co-presence and collaborators' activities, and increased loneliness and sense of isolation. Neither extreme works, as such, it is essential to balance between openness and privacy of virtual artefacts.

Based on the practices of three studies, there can be two suggestions for the balancing: (i) When providing privacy is needed, designers and creators of SVEs should firstly consider constructing privacy via the modality closed related to the output of the collaborative task, e.g. the audio channel should be firstly considered for CMM tasks because the output is a music piece and is related to audio. (ii) As many channels as possible should be left public to promote the construction of team-awareness. Reason behind these two suggestions are: in all the three studies, as CMM was the collaborative task, we mainly used the audio channel to provide sonic privacy to support users' creation. It turned out that the provided privacy was sufficient. Specifically, Study II found Additional personal spaces supported individual creativity, Study III reaffirmed this find-

ing and found C_{aug} could minimise the side-effect introduced. Namely, audio modality was used to provide sonic privacy, which was enough to support the creation and reduce the disruption between collaborators. Other channels (e.g. visibility) were left public and perceivable for all users, helping the construction of team-awareness and workspace awareness. C_{PV} of Study II is a good example, it provides sonic privacy but it still kept the visual channel public, which, according to the interview and thematic analysis, promoted their sense of collaborator's activity whist still provided enough privacy for individual creativity. C_{PI} of study II is a counterexample, it blocked user's sight of each other whilst being inside personal space. It was reported to be "isolated", harder to track collaborators' activities.

7.3 Visual vs Auditory Approaches

Study I explored how to support CMM in SVEs via two visual tools Work ID and 3D annotation, Study II and III explored it through an auditory approach, specifically, by testing different auditory configurations. It would be very beneficial to compare the two approaches, however, it is quite difficult and impractical to compare all the tools and configurations given the fact that some of the tools are so different from the some of the configurations, make them even incomparable.As such, this section specifically compares 3D annotation in Study I and augmented attenuation of Study III where comparable. The reason of choosing these two features is that both of them can be approached anywhere in the virtual space, and that commonality increases their comparability. The visual feature 3D annotation of Study I, by enabling players to write 3D lines, supported their communication. It can be seen as a visual cue. Study I explored the how *3D annotation* might support the collaboration. Differently, the augmented attenuation feature of C_{aug} in Study III explored an auditory approach. Both approaches have turned out to be effective. Next, these two approaches will be compared against each other, seeking the potential differences and possible usage scenarios.

7.3.1 Modalities

The augmented attenuation is an audio approach, whist 3D annotation is a visual approach. This fundamental difference results in their unique advantages and disadvantages, which then determine their scopes of usage. Specifically, the visual approach can fully avoid influences on the audio channel, leaving that modality purely for composers to hear the project they are working on. On the contrary, the audio approach used in C_{aug} in Study III may impose

unavoidable effects on how the audio sounds , because the sonic privacy being used is produced by augmenting the acoustic attenuation of the audio itself, i.e. what users can hear not only depends on the melodies from the instruments, but also depends on their locations.

7.3.2 Interaction Type

Explicit interaction is consciously deciding to interact, e.g. clicking the icon of an application to open it. It is what we normally think about when we are interacting with a computer (Serim & Jacucci, 2019). Compared with explicit interaction, implicit interaction does not require users to perform conscious actions. Implicit interactions usually rely on users' movements (e.g. head movement, eye movement), during the process, the user should not be consciously interacting with the system. Unlike the 3D annotation, which requires explicit interaction to make 3D lines, the augmented attenuation only relies on users' passive listening and active physical locating in the virtual space. As a result of applying explicit interaction, 3D annotation requires users to learn how to make the interaction happen and requires corresponding devices (gesture tracking devices in this context) to make it happen. As such, it has higher learning cost and higher requirements on tracking equipment. However, the difference between the two modalities does not necessarily lead to the difference in interaction type. This difference between visual 3D annotation and acoustic attenuation is more a result caused by the writing nature of the 3D annotation required a more explicit input, which caused the introduction additional interaction and increased the learning cost. In other words, other visual approaches might also use implicit interaction, and audio approaches might also require explicit interaction.

7.3.3 Key Support for Collaboration

Study I revealed that the 3D annotation can support the social aspects of the collaboration, help people to warm up at the beginning, support the non-vocal communication, and provide help for collaborators to understand each other's attention. So it is a good tool to intensify the links between collaborators. While the augmented attenuation gives collaborators the choice to be separated, and hence provides support for individual creativity. With this flexibility, users have the choice to develop their own work, and to switch fluidly between working on own and teamwork. Both approaches turned out to be effective, therefore when supporting creative collaboration in SVEs, there can be two directions to put efforts. One is to strengthen the link between the collaborators to increase the communication quality, the other is to provide a scheme to allow collaborators to isolate themselves to provide support for individual creativity.

Table 7.3: Comparison between the two approaches applied in Study I and Study III.

	3D annotation	Augmented Attenuation
Modal	Visual	Auditory
Interaction type[a]	Explicit interaction	Implicit interaction
Key support	Supporting communication between users	Supporting development of individual creativity
Advantages & disadvantages	(i) New interaction introduced, learning cost and requirements for supportive devices caused by newly introduced interaction	(i) No new interaction introduced, almost zero learning cost due to no explicit interaction introduced
	(ii) No influence on audio channel	(ii) Influence on audio and composition
	(iii) Users can hear the exactly same audio[b]	(iii) Users can not hear the exactly same audio
Applications	Wider range of application, not restricted to audio tasks, audio tasks requiring precise audio output, or users with hearing/speech impairment	Restricted to auditory tasks with no requirement for precise audio outputs

[a]See more about explicit interaction and implicit interaction in Serim & Jacucci (2019).
[b]Rigidly speaking, what users hear still slightly differ unless the spatialisation of audio is disabled.

7.3.4 Advantages and Disadvantages

Based on the points made above, here we compare these two approaches' difference in terms of their impacts on composition and learning cost:

Impacts on Composition - The 3D annotation only takes the visual channel, hence it avoids influences on audio, which is both the primary medium and final output of collaborative music making. Being able to hear exactly the same audio and audio only changes when manipulating the musical interface makes it easier for collaborators to gain a common recognition of the piece being produced. On the contrary, due to its nature, whilst using augmented acoustic attenuation, what users hear not only depends on the music piece itself but also greatly depends on the their positions relative to audio sources, which unavoidably affects what users hear. The bigger difference between users' locations, the bigger difference between what they hear, and the more difficult they get a consensus. Nonetheless, collaborators can choose to walk closer for a smaller discrepancy between what they hear.

Learning Cost - To enable users to make 3D annotations, gestures to add or drop lines were introduced. Though these gestures enable users to control how and when they write 3D annotations explicitly, they did bring learning cost

and other potential costs (e.g. requirements of gesture tracking devices, user experience issues caused by the interaction). On the contrary, the augmented attenuation only relies on users' head positions - an implicit input that does not introduce new interaction for users to learn and to master. However, we should note this difference (explicit/implicit interaction) is more a result of these two specific tools, rather than a difference between visual and auditory modality. As mentioned above visual modality does not necessarily require explicit interaction and introduce new interaction and audio modality does not inevitably avoid so, both really depends on the way how that approach functions. Though explicit interaction introduces learning cost, it is not necessarily a worse choice compared with implicit interaction, as explicit interaction enables more forms of input interaction and hence can serve a larger range of interactivity.

To summarise, compared with augmented attenuation in Study III, the tool 3D annotation has three major advantages: (i) Its application is not limited to sonic tasks because it provides support to communication, which is required by many collaborative tasks in SVEs; (ii) Due to exploiting the visual modality rather than audio, it might be able to support users with hearing or speech impairment; (iii) It completely avoids impacts on the auditory channel. This supportive measure suits where the output comes with stringent requirements, and users must be able to hear exactly the same audio output during the collaboration. By contrast, the augmented attenuation has a narrower application range, it provides better support for individual creativity, with still providing enough context of group work. The (slightly) differed audio output make it only appropriate to audio related-tasks with no rigid requirements, e.g. people are improvising music for fun.

These two supportive features do not contradict each other, and could be applied simultaneously. To manage the simultaneous use, a manipulation system might be needed. For example, the transparency of the visual 3D annotation and the degree of augmented attenuation can be adjusted to tailor their impacts (visibility/audibility), fitting collaborators' needs during different stages of the collaborative composing. When only one feature is needed, the other can be adjusted to zero, entirely wiping out its impacts.

7.3.5 Choosing Visual or Auditory Tools for CMM in SVEs

Next, following discussion above, we propose three implications for choosing visual or auditory tools supporting to CMM in SVEs.

(1) When there is no rigid requirements on audio outputs, auditory tools like augmented attenuation can be exploited to create audio privacy, which can then be used to promote individual creativity during the collaboration.

However, augmented attenuation introduces differences in what collaborators hear, making it only applicable to contexts with no rigid requirements on audio outputs. So when there is rigid requirements on audio outputs, using tools based on other modalities is suggested, such as the visual tools – 3D annotation.

(2) Visual tools and auditory tools do not necessarily contradict with each other. As such, they can be applied simultaneously, or may be served with a flexible switch, so users can chose based on their need during different stages of the collaborative composing.

7.4 Implications for CMMs in SVEs

SVEs, different from traditional media, has shown its special strengths in supporting collaboration in the three studies. Those SVEs supporting CMM, like LeMo, have their own characteristics and requirements, hence it would be beneficial to conclude some design principles to suit them. Based on the claims made previously, this section conclude 6 principles to inform the design of SVEs supporting CMMs. Next, these principles will be specified.

7.4.1 Consider Bare Hands Interaction

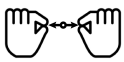
VR provides us with a space, although virtual, it comes with a real sense of depth. With one more dimension, more intuitive interactions with higher fidelity and DOFs could be designed, delivered and applied. E.g. using finger movements as cursor movements, and finger clicks as mouse clicks. Such interaction methods have higher fidelity and are able able to improve user performance in terms of searching for non-present targets (Pausch et al., 1997), manipulating (Ware & Jessome, 1988) and rotating 3D objects (Hinckley et al., 1997). Gestures can be designed and used as triggers, similar to the mouse gestures, cf. CrxMouse[1]. Different from mouse with 2 DOFs and 2 to 3 buttons, each avatar hand available in LeMo has five fingers, each finger has 2 to 3 bones, each bone is a piece of object that has 3 DOFs, hence there is a much greater richness for interaction design in VE. As examples of applying such affordance, in total, 8 gestures were designed and used in LeMos. Table 7.4 listed all these eight gestures and their possible usage scenario.

Next, based on the practice of applying gestures in the three studies, 3 design implications are proposed for gesture design in VR:

(1) Decide whether to use palm or finger movement as the trigger based on the interface size and required accuracy. Finger movements provide better

[1]CrxMouse: `https://chrome.google.com/webstore/detail/crxmouse-chrome-gestures/jlgkpaicikihijadgifklkbpdajbkhjo?hl=en`

Table 7.4: Gesture Applied in LeMos.

Name	Representation	Usage scenario	Analogues
Finger-single-click		High accuracy, intensive activities	Mouse single click
Finger-double-click		Low accuracy, moderately intensive activities	Mouse double click
Palm-single-tap		Low accuracy, intensive activities	Mouse single click, single tap a physical object
Palm-double-tap		Low accuracy, moderately intensive activities	Mouse double click, double tap a physical object
Palms-facing		For occasional activity	Both palms facing down
Finger-drag		High accuracy, for controlling sliders	Dragging visual/physical slider
Pinch-drag		High accuracy for re-positioning objects	Moving a physical object
Pinch & stretch		For enlarging objects, zooming in and so on	Stretching physical objects

accuracy and require smaller movement range compared with palm movements, and thus might better suit the situation where the size of interface is small, and higher accuracy is required. E.g. LeMo II has a larger size of matrix, resulting in smaller buttons compared with LeMo I. Thus *click-by-index* gesture was chosen rather than *tap-by-palm* gesture, which used in LeMo I, see Table 7.4.

(2) Choose gesture based on the intensity of uses. Using simpler gestures for more frequently-used interaction. Usually, the most frequent use in the activity is the interaction that achieves the basic function of the interface. E.g. the basic interaction with a music interface in LeMos is to add and remove a note, so one of simplest gestures – *finger-single-click* and *palm-single-tap* – were chosen to perform this function in LeMos, see Table 7.4.

(3) Mimic the interactions that people are already familiar with to ease the costs of learning and memory. E.g. in LeMo II, *finger-single-click* (a single click by finger) is similar to a single mouse click, *pinch & stretch* gesture is similar to grabbing a physical object and stretching it in the real world, see Table 7.4. This implication also partially echos the principle 4 "make use of existing skills" and principle 5 "consider natural interaction" proposed by Serafin et al. (2016), and the argument on applying gestures that people develop to manipulate real objects, such as picking up, positioning, altering, and arranging objects (Jacob

et al., 2008).

7.4.2 Binding Natural and Magical Interaction.

Simulating natural interaction existed in the real world is not enough, additional application of the unlimited possibility that VEs present is suggested. Using natural interaction is suggested for the basic and frequently-used manipulations, like the gestures people develop to manipulate real objects, such as picking up, positioning, altering, and arranging objects (Jacob et al., 2008), so users could grasp them quickly and easily (e.g. pressing the button, dragging the slider in LeMos, see Table 7.4), and leaving the magical interactions for users to explore (e.g. the double tap to pack up/pop up the interface). This echoes the argument in Serafin et al. (2016) – "consider both natural and 'magical' interaction", but the difference is that here we argue binding these two types of interaction seamlessly is also necessary and suggested, as it can help users to understand and remember these in-between interactions. For example, the gesture to generate music interface inherits the natural phenomenon that objects become bigger when being stretched, however it is still magical as people could not do this in real world. This type of interaction inherited the advantages of both natural interaction (easy to remember, understand and recall), and the magical interaction (making things impossible in reality possible).

7.4.3 Hiding Unavoidable Limitations

Design must be made in a way to hide or weaken rather than expose inevitable, technical limitations to the user. Different from the principle "consider display ergonomics" proposed by Serafin et al. (2016), here the limitations are expanded and not limited to display anymore, they now include but are not limited to display ergonomics, HMD ergonomics, controller ergonomics, tracking devices ergonomics, audio display ergonomics, space limitations. For example, LeMo II chose to use smaller music interfaces to enable users to see the whole grid interface considering the narrow FOV provided by the HTC Vive HMDs. Other consideration include: (i) Design the experience in a way that users can play in discrete sessions rather than a long continues experience to allow user to have a break to alleviate the tiredness caused by these limitations (e.g. the weight of HMDs), (ii) Avoid text reading, if have to, using large size (due to the 3D display of text is not good for read, and low resolution of displays), (iii) For wired HMDs, virtual workspace should be designed that people exchange position less. Because wires of HMDs might get entangled, especially when there are multiple players. A good example is the spatial configuration C_{PV} and C_{PI} of Study II, with each personal space positioned at the opposite side the public

space, saw much less entangled wires happening than the configuration of Study III, in which users could access personal space anywhere in the stage.

7.4.4 Use Embodiment to Strengthen Social Awareness.

The embodiment plays an important role in SVEs supporting CMMS. Comparing Study I with the 2D CMM system – Daisys (no avatar) applied in Bryan-Kinns (2011), the fewer usage of annotations to convey presence and localisation might be due to the fact that avatars already support this well. By seeing the avatars that synchronised with real body movements, participants were able to perceive enough information about where the collaborators were and what they were doing. In study II, participants' comments and preference on C_{PV} (public space + publicly visible personal space) over C_{PI} (public space + publicly invisible) indicate that keeping all players' avatar visible continuously is important to maintain a steady sense of collaborator's presence and activities, which is key factors for a good collaboration. Therefore, in SVEs supporting CMM, to ensure a proper level of sense of each other's presence and activities, an application of proper avatars is required. The followings are some preliminary suggestions based on the experience of carrying out the studies. Given that the avatars applied in LeMo did not come with realistic materials or models but performed well, a more important factor of the avatar seems to be the low latency of synchronisation, rather than high realness. Apart from this, if no eye tracking was applied, we suggest opaque glasses should be considered for the avatars, to avoid direct eye contacts. This is because the fake static eyes might reduce the realness of the VEs, which can then weakens users' sense of presence.

7.4.5 Manipulate the Space

Study II and Study III have shown the virtual space is the medium where collaborators communicate with each other and perceive information as well as the tool which can be used to impact the collaboration. As such, manipulating the space in a meaningful way can be powerful to support the CMM in SVEs. Below, we propose two suggestions based on the Study II and Study III.

(1) Augment the physics of the space. Augment how the dissemination of modalities (e.g. light/sound/odour) when travelling across the space. For example, in C_{aug} of Study III, the acoustic attenuation was found to be a powerful tool to enhance the music production. And according to the findings of Study II and III, this light-weight form tool has less negative impacts than some other solid form of tools (e.g. the solid soundproof walls in Study II). Similarly, other modalities, like visibility, can also be augmented within the space to enhance

the experience based on needs. For example, making the space more foggy will possibly result in less visibility and higher level of privacy.

(2) Divide the space. The virtual space can be universally shared like the C_P in Study II and C_{pub} in Study III, it can also be divided and more personally owned, like C_{PV} and C_{PI} in Study II and C_{mov} and C_{fix} in Study III. Dividing space creates several spaces and makes it possible to let people own space, which can then be used as personal space. Personal space, according to findings of Study II and III, can support individual creativity and increase efficiency with some side-effects (e.g. shrunken group territory and group edits). The side-effect became even greater when the space is more rigidly divided, like the C_{PI}, which uses invisible wall and blocked the vision, or C_{fix}, which provides less mobility compared with C_{mov}. Hence, the physics of the material and the way used to divide the space are important and need to be carefully considered to minimise the costs. Less rigid material and light-weight blocking method are suggested, for example, blocking material can come with with higher visibility and mobility.

7.4.6 Evaluate the Presence.

Instead of a principle, this is more a step to do after the delivery or during the iteration of developing a VE system. Presence is the magic that VR presents, it gives user the feeling they are there and are able to interact. Is is so fundamental to VR systems that an incompetence in delivering presence will fail the VR systems. As such, accessing its quality is important. Tools like questionnaires can be applied to measure how well the system does in producing the sense of presence. Different from Serafin et al. (2016)'s principle "create a sense of presence", which suggests to action should be taken to create the sense of presence. However, we believe presence is more an overall result, or in other words, an overall mark a VR system gets to reflects the immersion level. It is not something that could be directly improved, instead, its improvements are based on many other aspects (e.g. realness of VEs, timely feedback and so on) and it can be weakened by mistakes. So it is more suitable to be a reflective tool rather than a design principle to follow.

7.5 Contribution Beyond Music Making

Although all the three studies in this thesis were focusing on supporting Collaborative Music Making, the implications proposed previously in the three studies and this chapter might still have reference value to design of SVEs beyond music making. Though it should be noted the following discussion and any suggestions made are the personal opinions of the researcher as a practitioner in the field of

VR rather than directly drawn from the findings of the three studies reported earlier. The following parts of this section discusses the possible value of the these findings in a wider scenario.

Study I - 3D annotations has shown its potential in making signs and simple texts, and improving warm-up process, the former of which could not be easily delivered by vocal communication. These advantages of 3D annotation are not necessarily limited to audio related tasks. It can be used for supporting collaborative drawing as well, as a supportive communication tool. Hence supporting 3D annotation can be considered when it comes to other types of collaboration, aiding the usage of main communication channels (audio and visual communication). E.g. while performing a furniture arrangement task together, people also need to make marking signs to supplement vocal communication, and might also use 3D annotation to draw as a warm-up activity.

Study II - Two implications from this study are: providing personal space is suggested and personal space with public visibility is preferred. For tasks beyond music making, these implication are still of reference value. For example, in collaborative drawing, people might also need a way to develop own ideas and a way to get rid of interference at some points during the process of drawing. Also for most collaborative work, having a consistent sense of other's presence and activities is important, hence personal space with public visibility is still suggested.

Study III - One of the implications of Study III is that providing personal space with light-weight form is suggested as it introduced minimised side effect. This might also apply beyond audio-related tasks. Because the light-weight form is not limited to audio, it can be one of other modalities (e.g. visual). For example, similar to augmented acoustic attenuation in Study III, a visual augmentation might be used for vision related collaborative tasks (e.g. collaborative drawing) in SVEs. Multiple modalities can also be used simultaneously for tasks involving multiple modalities, an example task can be making a short animation and creating an accompanying music track for it. The other implication proposed in Study III is to manipulate the level of augmentation to adjust the level of privacy. E.g. for music making, the augmented attenuation can be set to a very low level if an extreme openness is being pursued. So adding a method allowing users to adjust the level can allow users to shift between having a "very personal" space with total isolation where they could not hear nor see each other's work), and having no personal space when they have to work together. This is not limited to music-related tasks, because visual privacy can be manipulated in a similar way (e.g. adjust the density of the fog in the VE).

Section 7.1 - The Role of Embodiment. The finding that the embodiment plays an important role in the three studies to support users' co-presence,

sense of other's activities are not limited music making. Most types of collaborative tasks can benefit from a high-level of co-presence, e.g. allowing collaborators to see each other's avatars during collaborative drawing may help them to coordinate.

Section 7.2 - Balancing Privacy & Openness. Finding a balance between privacy and openness is important, not only for CMM, but also for other tasks. And the implications given in Section 7.2 – using the main modality of the task to construct privacy and keeping other modalities public – might be applicable for other tasks as well. For example, for collaborative drawing, a visual privacy could be derived by a simulated foggy environment, with keeping other modalities (e.g. the audio) public. Some example ways to manage visual privacy, sonic privacy and accessibility have been detailed in Table 7.2, which are not limited to sonic-related collaborative tasks either.

Section 7.3 - Visual vs Auditory Approach. Section7.3 compares the visual and auditory approach, the 2 implications given at the end of the section are closely related to music making task and hence have limited reference value for tasks beyond music making.

Section 7.4 - Implications for SVEs and VRMIs. Though the principles proposed in Section 7.4 are for applicable Collaborative Music Making systems (CMMs) in SVEs, they are still of value for SVEs beyond music making, because none of these rule is limited to music making. For example, the implication for gesture design also has reference value for tasks beyond music making.

7.6 Chapter Summary

This chapter has compared and discussed the results and findings of the three studies where comparable. Specifically, this chapter has firstly discussed the role of embodiment – a possible important factor thorough the three studies. Secondly, this chapter has reflected how to manipulate privacy and openness by reviewing the relevant spatial conditions of Study II and III. Thirdly, the visual approach in Study I and the auditory approach in Study III have been compared. Fourthly, six principles for SVEs supporting CMM have been proposed. And finally, how the findings might contribute to design of SVEs beyond music making has been discussed. The following chapter is the final chapter, which will conclude this thesis.

Chapter 8

Conclusions and Future Perspectives

This chapter recapitulates the findings in respect to the research question proposed in Chapter 1, and the contributions of this thesis. Limitations will be discussed and potential future perspectives will be indicated.

8.1 Major Contributions

As listed below, in total, there are five groups of major findings. Group 1 provides a basis for the studies, whilst group 2, 3, 4, 5 directly address the research question: How to better support collaborative music making in shared virtual environments.

1. Developing a system that supports CMM in SVEs. Following design principles for VRMIs, LeMos were designed and built and successfully fulfilled the research aim. As far as the author is aware of, LeMos are one of the earliest systems supporting collaborative music making in SVEs. By building LeMos, the gap between CMM and SVEs has been filled.

2. Exploring how visual cues might support collaborative music making in SVEs. The first study explored how users use the two visual cues, the unique advantage and limitations of supporting 3D annotations as a communication tool were identified and explained. Implications were also given to inform future application of 3D annotation in VEs.

3. Exploring how to design virtual spaces to support CMM in VEs. Study II and Study III focused on testing the how different spatial configurations impact the collaboration in music making. The major findings are: people form territoriality in VEs and perform territorial behaviour, adding personal space is

beneficial but might introduce some negative side-effects, personal space with visibility is preferable, light-weight form personal space has minimised negative impacts. Multiple design implications were given accordingly in Section 5.5 and Section 6.5. These implications can inform the spatial design for future CMMs in SVEs, and some of them are applicable to inform the design of SVEs beyond music making.

4. The measurements and evaluation of CMM in SVEs. The three empirical studies presented in this thesis applied both quantitative and qualitative approaches to collect data and reflect the collaboration. Four themes of measures were developed and practised, three of which were for assessing the collaboration and one was for accessing space and territory usage. Interview, questionnaire and system-logged data were used to cover these measures and they turned out to be effective to detect the impacts of different experimental conditions on collaboration. These methods are of reference value for other future studies about collaboration in SVEs.

5. The embodiment usage, balancing privacy, using different modalities, and implications for SVEs. Chapter 7 reflects LeMos and findings of Study I, II and III. Specifically, it discusses the embodiment usage, ways to balance privacy and openness, visual and auditory approaches and proposes implications for CMMs in SVEs. All of these implications are of reference value for CMM in SVEs and some of them are informative for collaboration beyond music making.

8.2 Limitations and Future Perspectives

8.2.1 Limitations

As far as we know, the methodology and study design of this thesis have several limitations.

Limitation of Study Design. Study I prohibited vocal communication to encourage the use of 3D annotation and to explore to what extent 3D annotation might support the communication. However, it should be noted prohibiting voice communication is very unnatural. Namely, the prohibition might have created an artificial situation where participants experiment with alternative ways to communicate with their peer.

The three experimental conditions of Study II were not fully randomised with C_P (the condition provides only public space) always came the first. This was to avoid impact of introducing personal space on the observation of emergence of territories. As discussed in Chapter 5, the drawback of this solution is clear, participants were more experienced in latter sessions, so they might have more

positive feelings of later sessions. We were fully aware of this limitation prior to the study, the reason the author choose to do so is to avoid impacts of adding personal space on the formation of territories, which is a basic but fundamental starting point to understand how people uses virtual spaces in collaboration.

Limitation of Methodology. Questionnaire Design. The questions of the Post-Session Questionnaire used in all the three studies were either phrased positively or neutrally, with no negative statements. Though for most questions, there are negative options on the scale for participants to choose, this limitation might still impact the rating. For example, in Study I, the description of PSQ1 in the Post-Session Questionnaire is: *In the virtual world, I had a sense of "being there"*. Though there are both positive (fully agree) and negative options (fully disagree) on the Likert-scale, the positive description might positively impact participants rating on this question. However, this imperfection might only have had a limited influence on this study, because PSQ results are mostly used for comparison between conditions, which are affected equally due to that PSQ of all of the conditions of all the studies used the same phrasing.

Data-Log-System. The first study was based on LeMo I, which did not come with a proper data-log-system, hence the logged data is very limited. The transformation information of heads, hands and interactivity were not logged. These data might be very powerful in revealing how people coordinate with each other whilst using the two visual cues, and in cross-verifying other findings.

The thematic analysis. The thematic analysis was carried out by the author himself. Having only one manipulator means more bias might have been introduced into the process. However this is a compromise given the large amount of time performing thematic analysis takes, finding another professional to perform the thematic analysis without payment is almost impossible. Hiring someone with payment is also quite impractical considering the large amount of money needed.

8.2.2 Future Perspectives

In the future, it would be interesting as well as beneficial to dig deeper into 3D annotation, e.g. more freedom for making 3D annotation, such as different strokes, colours, and its application in a wider fields. In terms of space design in SVEs, exploring how to design and apply personal spaces with fluid boundaries in s wider range of creative scenarios in SVEs might be a good perspective. For example, for collaborative drawing in SVEs, personal space (visual privacy) might be provided by creating a foggy environment, the more far away from the drawing objects are, the more blurry the the objects are. We are also interested in how the personal space's boundary and the augmentation level

can be manipulated and whether these manipulations can result in different impacts on the collaborative behaviour.

8.3 Closing Remarks

In this thesis, a system named LeMo that supporting CMM in SVEs has been presented and three studies based on it have been demonstrated. The first study found that people used annotations to support collaboration, three key types of uses were identified, which partially follows the aME classify scheme. The second study confirmed the emergence of territories in collaboration in VE, and two types of territory were found, echoing the findings made in tabletop research. We also found the personal space is essential to collaboration and it affects the formation of territory, thus requires careful consideration. Study III identified the effectiveness of using augmented acoustic attenuation as a tool to provide sonic privacy, and its advantages as an invisible light-weight form personal space. Based on results and findings of the three studies, the embodiment usage, ways to balance privacy and openness, comparison between visual and auditory approaches, and implications for CMMs in SVEs have been discussed, and the applicability of the implications beyond music making have been discussed. The author wishes this thesis provides useful insights for future CMMs and SVEs.

Appendix A

Material of Study I

A.1 Ethical Approval

See the ethical Approval in Figure A.1.

A.2 Questionnaire

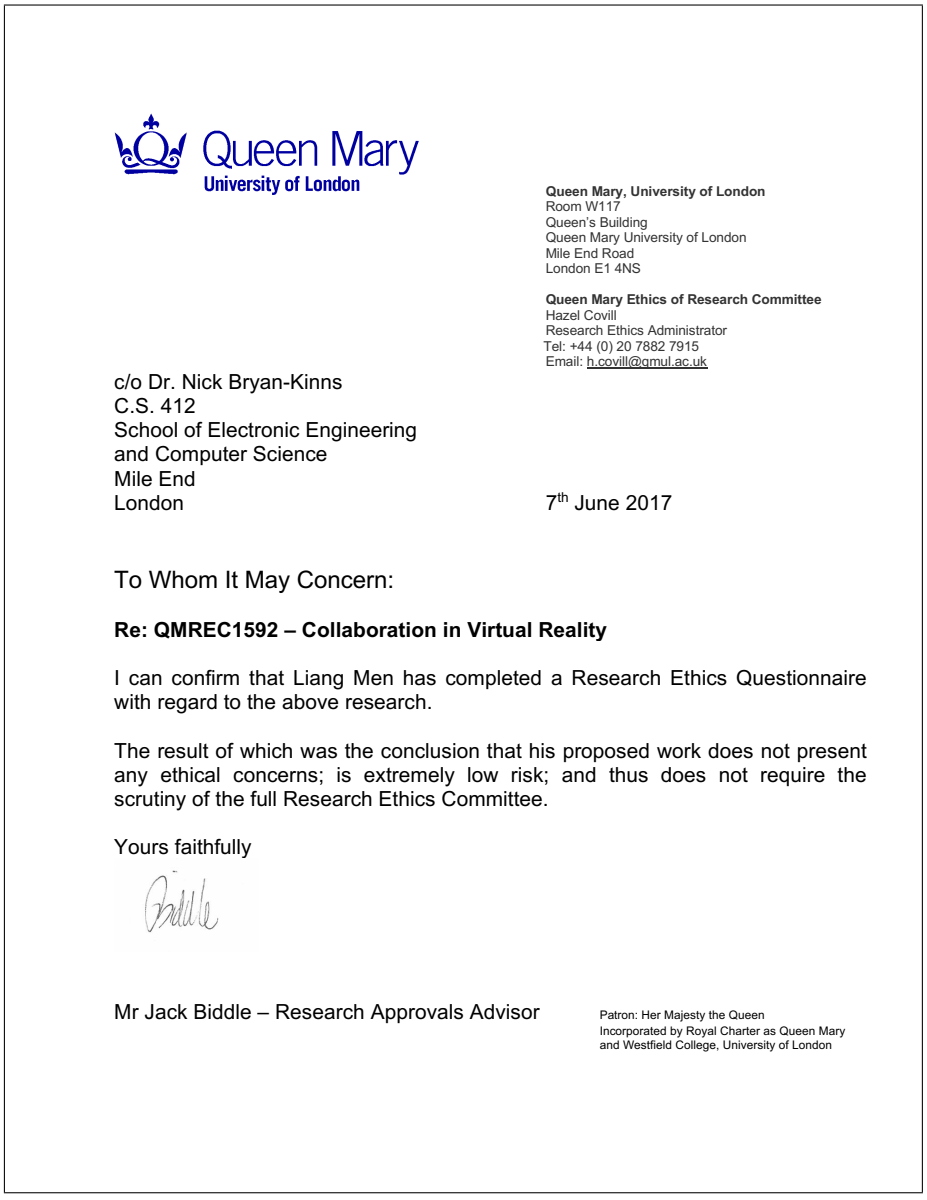
A.2.1 Background Information

- Your participant ID
- Your gender (female/male)
- Your age (18-29/30-39/40-49/50+)
- Have you ever experienced virtual reality? (I have not tried it, this is my first time; I have tried them, but only once; I have tried them 2-5 times; I played VR frequently)

A.2.2 Seven Post-Session Statements

These statements come with a 7-point Likert rating.

- In the virtual world, I could strongly feel someone was there collaborating with me together. (1-fully disagree; 7-fully agree)
- I had a clear sense what he/she was trying to do. (1-fully disagree; 7-fully agree)
- I had a feeling, at some points, my interaction partner created notes according to my notes. (1-fully disagree; 7-fully agree)



Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Ethics of Research Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr. Nick Bryan-Kinns
C.S. 412
School of Electronic Engineering
and Computer Science
Mile End
London

7th June 2017

To Whom It May Concern:

Re: QMREC1592 – Collaboration in Virtual Reality

I can confirm that Liang Men has completed a Research Ethics Questionnaire with regard to the above research.

The result of which was the conclusion that his proposed work does not present any ethical concerns; is extremely low risk; and thus does not require the scrutiny of the full Research Ethics Committee.

Yours faithfully

Mr Jack Biddle – Research Approvals Advisor

Figure A.1: Ethics Approval for Study I

- We had a high-quality non-verbal communication. (1-fully disagree; 7-fully agree)
- We had a high-quality non-verbal communication. (1-fully disagree; 7-fully agree)
- How satisfied are you with the piece of loop music you two finally created? (1-not satisfied at all; 7-fully satisfied)

A.2.3 Open-Ended Questions

- Regarding the collaboration experience, do you have anything to say?
- How do you think of the interaction with the sequencer in VR, e.g. is it good/bad/intuitive/hard?
- How do you feel about the VR experience?

A.3 Thematic Analysis Results

Parent Code(N)	Code (N)	Participant	Segments
3D Annotations's	Advantages (15)	P1A	[The] documenting communication would be better as well as in this, for example, it was really nice you can write.
		P1A	Vocal communication would be better, very quickly make my [ideas] clear to my partner by using musical terminologies. But if he is not music theory-based, the vocal could be harder as with the lines, I could just circle the notes to say that was G and going back to C, from that perspective, drawing was more effective.
		P1A	I think the relationship you had would be different. If all you have is the verbal communication, I think you will be using your words more sensibly, and you will be given more prominent cues, if you could write as well, you could talk about what you were doing and then write the prominent noun. It would be good if you have equivalent to white board, you could write precisely, so I could write smaller.
		P3B	The next step is to listen and adjust. At this stage, the line can be useful. We can do the adjustments according to the line...For example, there was a line drawn by her, and I deactivate a button according to the line she drew.
		P3A	As long as we don't write texts, [3D Annotation] is OK to support communication.
		P3B	But you can use this line to influence each other. In the middle point, because she was on the left, I wanted to make the other note lower. I saw that she had adjusted [according to the line drawn by me].
		P5B	I could draw something, like if I felt satisfied, I can draw a tick.
		P6B	I thought the writing part was interesting because you could both do it...
		P6B	We'd like make little marks in the air, and exactly like you know things you want to drop and go. I do like the drawing, interesting.
		P8A	I wrote "OK"to confirm the finish of the work. Did you see that? (P8B: Yes.)
		P9A	I think drawing is quite useful, at the beginning two people started to communicate.

		P9A	I feel that if I can draw lines during the cooperation, I can still feel like I have a language.
		P9B	It is helpful to draw lines, but the recognition is not very good, there will be breaks during drawing.
		P16B	In the beginning, it was nice to create drawing together, feel like more you have your partner in the virtual world as well. And it was nice, like a physical warm up.
		P16B	It is more useful at the beginning as communicator, warm up and enjoyment, make people feel comfortable in VR, but...
	Disadvantages (13)	P1A	You need to relate to very logical semiology.
		P1A	Because of the nature of the writing, it is quite big, and there is a limit of how much you can write.
		P1B	And you need to stand where the person wrote it stood. (P1A: Yes, you have to read it.)
		P3A	It's hard to recognise and it's hard to write. You don't get used to writing with your hands...
		P4B	The line is not clear, you could not write clearly [accurately].
		P4A	It's hard to breaks strokes, too different from the normal writing experience.
		P6B	It was also strange because if wrote something that make sense to me from my perspective, because it was 3D, it was just rubbish to you (the interaction partner).
		P6A	If you start writing the message, it takes so long, when you finish it, it does not make sense. Another kind of talking would be easier.
		P9B	It is helpful to draw lines, but the recognition is not very good, there will be breaks during drawing.
		P10B	I always drew lines involuntarily, which is not what I wanted to draw.
		P10B	It is difficult to judge whether this line is for communication purpose or not.
		P12A	What you just pointed out is what I felt about, when you trying to write stuff, it is quite difficult to make the other person understand what you are trying to write.
		P14B	Cause we feel it is really limited and also sometimes it was 3D, it was more tricky to read. A: I can not face her and write.
	Suggestions (13)	P6B	Yes, a solution maybe you have an area around the board, was like a 2D plane that you could write on, like a white board, so that we could see it from many angles. Although drawing in air is kind of interesting, just interesting VR experience, just that as collaboration between people. You know draw characters, could be nice.
		P10B	It is difficult to judge whether this line is for communication purpose or not. It might be useful to use different hues to differentiate.
Work ID's	Advantages (10)	P1B	It is helpful to know that you are not the one who intended something.
		P3A	I can better understand which key is on, where [which button] the sound comes from? Because you don't know which button is the effect after it is on. It's all the same, I don't know how to change it.
		P4B	There are differences, two colours, there are my special colours, I know what I pressed, I would feel uncomfortable to cancel button belong to someone else will.
		P5B	Two hues [Work ID] are also very good, you can distinguish which one is made by which person.
		P5B	I think if a note belongs to her, I would try not to change it. But if the note had any influence on the work, I would still change it without many concerns.
		P7B	It's good [to have Work ID], but in fact, when I do it for the second time, there is no colour in time. I know that he pressed it, because I remember what I pressed. I think he may know which one is mine. You remember your own intention.
		P10A&B	For strangers, it might [introduce concerns]. For us, nope.

		P12A	So it helped us to distinguish the action. We might pay little respect to the other one. You had tendency not to change what the other people do.
		P14B	I did sometimes, but it was nice to see if these were actually her button and not mine. It would make me brave in changing things as when everything was the same, sometimes you could not remember what was your button, you clicked and change the button, I did not want to intervene or something.
		P15B	I probably like to have more parameters. The temple, the note, or maybe volume, more satisfaction, more freedom, constrained, colour as kind of guiding tool you interact with somebody, it is more process based.
	Disadvantages (9)	P1B	It is a lot better to see their notes than interpreting what they are doing and where they see the harmony going.
		P1B	We already know the progression, we already know the melody, structure, it doesn't matter which to put the notes down, we know it is the right notes to put in the right place, because of the plan we had.
		P1B	Make sure we could see where the other person's idea is going.
		P3B	For example, I can see that this part is what she did. This part is what I did.
		P4A	Well, it depends on people. If the other person is very professional, I will have concerns [when deactivating other's button when there is Work ID].
		P5A	I don't think [work ID] is useful. I can basically remember what is hers and what is not. There are only a few buttons. If there are more, [work ID] may be more useful.
		P6B	I kind of like it when the buttons are in the same colour, because you don't feel any ownership, so I like it better, it just like a bunch of things happening. It did not matter if I turned off, it could be his, it could be mine. No big deal, I just destroyed everything you did.
		P9A	I know that the button is his. I changed it with a bit concern.
		P10A&B	Have a sense of who made what.
Reporting LeMo System	Limitations (3)	P1A	I was very much kept on the first four steps because I knew there was very low chance of collision.
		P3B	There are restrictions, the loop is short, too short, I can't make a long song.
		P3A	I waved my hands to her. I thought she could see it, but she didn't. I had to reach out to her oblique front.

Table A.1: Thematic Analysis Results of Study I

Appendix B

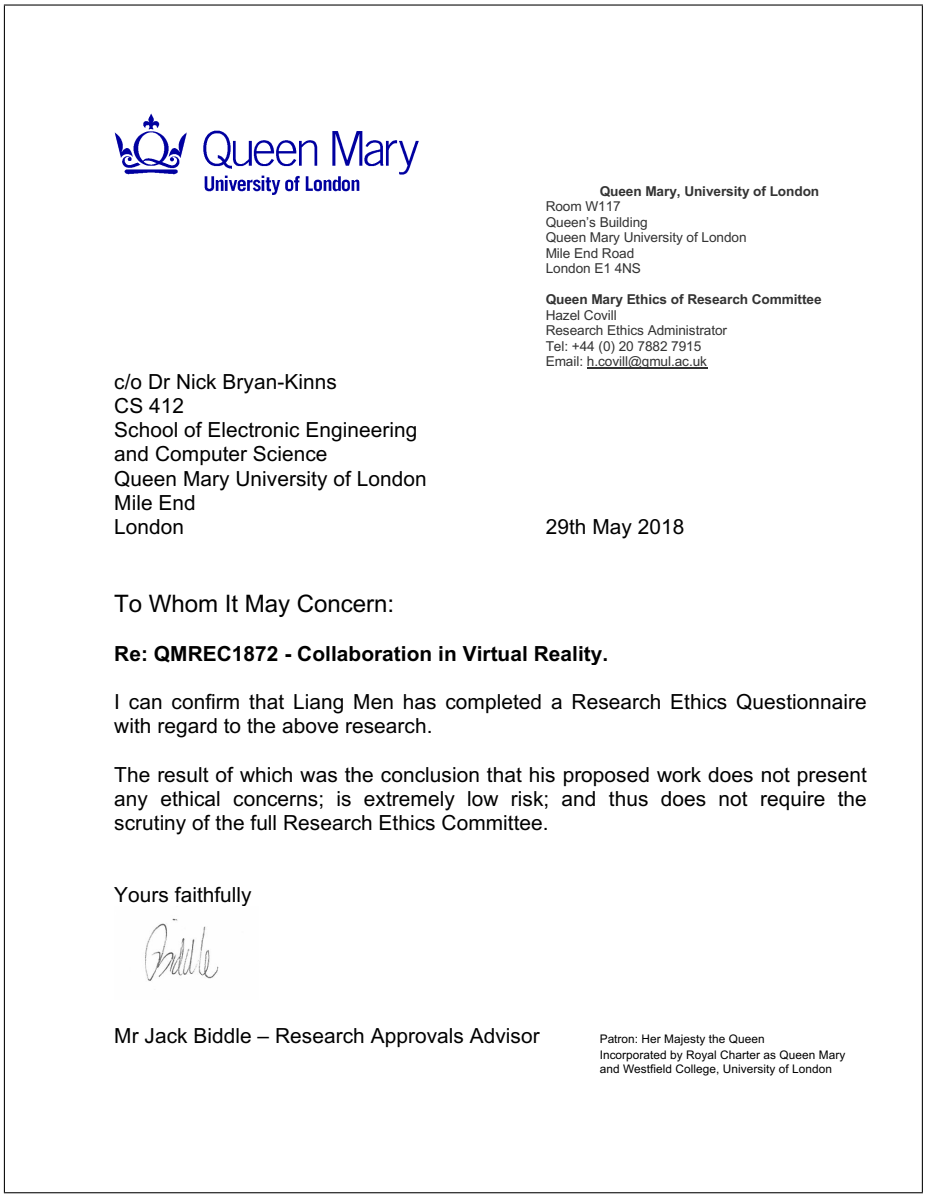
Material of Study II

B.1 Ethical Approval

See the ethical Approval in Figure B.1.

B.2 Demographics Questionnaire

- Your participant ID.
- Your gender. (female/male)
- Your age.
- How well do you know your collaborator in this experiment before the experiment. (I know him/her very well; I met him/her several times before, not know well; I met him/her only once before; I don't know him/her before the experiment)
- How would you evaluate your musical theory knowledge? (1-no theory at all;10-theory expert)
- Do you play a musical instrument? (yes/no)
- If the above is yes, which instrument(s) and how long have you been playing it/them?
- How would you describe your experience of composing music together? (1-no experience at all; 10-highly extensive)
- Have you ever written a song or made a piece of music? (yes/no)
- If the above is yes, how many?



Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Ethics of Research Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr Nick Bryan-Kinns
CS 412
School of Electronic Engineering
and Computer Science
Queen Mary University of London
Mile End
London

29th May 2018

To Whom It May Concern:

Re: QMREC1872 - Collaboration in Virtual Reality.

I can confirm that Liang Men has completed a Research Ethics Questionnaire with regard to the above research.

The result of which was the conclusion that his proposed work does not present any ethical concerns; is extremely low risk; and thus does not require the scrutiny of the full Research Ethics Committee.

Yours faithfully

Mr Jack Biddle – Research Approvals Advisor

Figure B.1: Ethics Approval for Study II

- How familiar are you with computers? (beginner/intermediate/expert)
- Have you ever experienced virtual reality? (I have not tried it, this is my first time; I have tried it, but only once; I have tried it 2-5 times; I played VR frequently)
- Have you used collaborative software in real time before? (yes, I have; no I haven't)
- If the above is yes, could you please give details?

B.3 Post Session Questionnaire

- In the virtual world, I had a sense of being there. (1-not at all; 10-very much)
- I always had strong feeling that my collaborator was there, collaborating with me together, all the time (1-not at all; 10-extremely so)
- How satisfied are you with the final piece of loop music you two created in this session? (1-not satisfied at all; 10-extremely satisfied)
- How would you rate the quality of communication between you and your collaborator during the session? (1-very low quality; 10-very good quality)
- I had a clear sense what my collaborator was doing. (1-not at all; 10-extremely so)
- I feel like the addition of private spaces in this session is very helpful to the task. (Only appeared in C_{PI} and C_{PV}; 1-not at all; 10-extremely so)
- Any reasons behind for the above rating? (Only appeared in C_{PI} and C_{PV})
- The amount of *your* contribution to the joint piece of music is (1-insubstantial; 10-substantial).
- The amount of *your collaborator's* contribution to the joint piece of music is (1-insubstantial; 10-substantial).
- What do you think of the quality of *your contribution to the joint piece of music is (1-very poor quality; 10-very good quality)
- What do you think of the quality of your *collaborator's* contribution to the joint piece of music is (1-very poor; 10-very good)
- Any reasons behind your above ratings on the contributions?

B.4 Comparison Questions

- In which session, you made the music you were most satisfied with? Any reasons? (most satisfied music/second satisfied music/least satisfied music)
- Which session did you find most difficult to track your collaborator's activities? Any reasons? (hardest/second hardest/least hardest)
- Which session did you have the strongest sense that your collaborator was there in the same virtual world, working with you together? Any reasons? (strongest/second strongest/least strongest)
- Which session did you have the best quality of communication between yourself and your collaborator?(best quality/second best quality/lest best quality)
- Which session had the best setting for creating a good piece of music collaboratively? Any reasons? (best setting/second best setting/least best setting)
- Which session did you find most difficult to cooperate with your collaborator? Any reasons? (most difficult/second most difficult/least difficult)
- Out of the three sessions, which session do you you feel you made the most contribution to the joint piece? Any reasons? (the most contribution/the second most contribution/the least most contribution)
- Out of the three sessions, which session do you you feel your collaborator made the most contribution to the joint piece? Any reasons? (the most contribution/the second most contribution/the least most contribution)
- Were there any differences between these sessions? If yes, please describe what was different?

B.5 Overall Questions Regarding Experiences

- Regarding the collaboration experience, do you have anything more to say?
- Any ideas or comments about the interaction experience with LeMo? E.g. in what way you feel it good/bad/intuitive/hard?
- How would you like it to be changed?

- If you want to hear more about this research in the future, please leave your email, thanks.

B.6 Thematic Analysis Results

Parent Code(N)	Code (N)	Participant	Segments
Learning Effect	Learning effect (17)	P1B	I can't think because I didn't pay too much attention there was something different, I found out that while I was using it I was getting better at it, that's it basically.
		P1B	Yeah, yeah and how to use it basically, and what to do.
		P1A	(LM[1]: Do you think it's also because you have more experience.) That is also one point yes.
		P3A	I think it counted for a lot (P3B: Yes) because you were, we were quite confident with what we wanted we had also come up with a few tricks on how we wanted the music to sound (P3B: Yes).
		P3A	You're working together, it is the one as well) so you're, already very excited.
		P4A	Yes yeah. I think it's partly for me, it was because we got used to the instruments by that time.
		P8B	Yes, we were experienced in the third session, and we did not know much in the former two sessions.
		P9B	Because initially when you the first, the, very first time when you're asking people to do experiment so you put it put, put them in the same space and OK, there's [there are] things [for you to do] so it makes quite difficult but then if the same thing was moved to the third one, maybe.
		P9A	I think being the first session being, like the session where you just try things together maybe made it make the second and the third one more efficient, because we made mistakes. Yeah, and we tried things together and because we were both learning we didn't feel like we're making delays or you know messing up the creative work, and then we learned and we felt more confident in the second session and the third session. Whereas if the order is different then we have to learn in our own private space and then when we come to the public space, then we need to you know adjust to each other.
		P9A	So I think if the first session was last the second and the third session, we would probably need more time for us to adjust but because we did the first session and we, you know make mistakes, and we learned how to do it more or less, the second and the third turned out to be much better.
		P10A	I have gained experience in how to make it.
		P12A	Yeah first one was hard.
		P12A&B	It was hard because everything is all together and this began just now (P12B). I guess, I didn't understand much in the first one, and then you you know (P12A).
		P13A&B	Because in the third session, we were quite unrestrained (P13A). Yes, we just made something casually (P13B).
		P17A	Yes, I think so, I feel more, much more confident in the system, there's a few things I would change about the drum interface, but, yeah I think I was more confident and while I was doing in the third one I could like muck around a bit more. There's a few weird things I was doing in the room, on my own.

[1] "LM" refers to Liang Men, the experimenter of the study.

		P18B	I think the biggest thing for me was just having lots of time to explore the system and by the third session I think I was more comfortable with using it so I think I preferred it more for that reason.
		P19A	Because more familiar, with the environment.
Most satisfied music in	C_P (0)	-	-
	C_{PI} (6)	P4B	I prefer two, yes.
		P8A	I think it is the second.
		P10A	The second is better, in the way I made the best piece...
		P17A	Yeah.
		P19B	Yes.
		P19A	Yes.
	C_{PV}(5)	P4A	I think it was, I think it was session 3. The last one.
		P8B	I think it is the third session.
		P10B	I think the third setting is the best.
		P10A	In the third session, because I have learnt something from the second, so I created a piece that sounds even better, though I did not know what to draw [draw means making patterns formed by notes].
		P15B	I think it was the second session.
Most unsatisfied music in	C_P (3)	P8B	I think it was in the first session.
		P10A	Anyway, we made the worst piece in the first session.
		P10B	Yes, the first piece is the worst.
	C_{PI} (0)	-	-
	C_{PV} (1)	P8A	In the third session, perhaps.
Most favourite condition	C_P(7)	P3B	Yes, yes, first, first session is that's the best setting.
		P5B	Um, the best is the first one.
		P8B	I prefer the first.
		P10A	The first setting, because we could discuss immediately.
		P16A	I would choose the first one the first one, because I am a beginner and if my partner is also beginner, then I will choose the first one.
		P21A	I preferred the first setting.
		P21B	I don't, I don't agree with that, I thought the first one was [the best].
	C_{PI}(10)	P4B	Not transparent, I prefer not transparent, because.
		P8A	I like the second.
		P9A	I like the second one.
		P10A	I think the second one.
		P11A	I chose the second is best [laugh].
		P12B	Yeah, yeah, yeah.
		P12A	I think for me, session II was the best as well.
		P18B	I think the third one for me.
		P19B	Probably the third.
		P19A	The third one.
	C_{PV}(20)	P2B	I like the transparent one.
		P3A	Yes I think so. I think I would still prefer the transparent one where you can see your partner.
		P3A	I think I would prefer the third one.
		P3A	I love the third session as well. Not the second one.
		P5A	I personally think that the soundproof area with transparent wall is better, because I could see my collaborator was editing things over that side.
		P5A	Yes.
		P6B	I think I would choose three because that way if I wanted to look I could but I wasn't looking, but if I wanted to I could I guess it's something an option.
		P7B	The last one, I think,
		P7A	Yes, the last one, visible one, I can see which colour she is using, which instrument, and I can avoid repeating it.
		P9B	Yeah.
		P11B	I think it is the third one.
		P13A	(LM: Why do you think the third is your favourite?) It satisfied the...

		P14A	Second one for me.
		P14B	Second one, me too.
		P15A	For me, I like the second one.
		P15B	Me, too.
		P16B	Well, I prefer the second one is the best second.
		P18A	I think for me that second session work[ed] best, because I mentioned in the questionnaire as well, but I found it easier to respond to what xname (P18B's name) was doing, because even though I couldn't hear that I could see which instrument he was working on, so it made it easier for me to, to kind of respond to that or to kind of know what we were both doing, whereas in the third session I felt there was a higher risk of that we might both work on the same thing and that I do something that is completely off compared (P18B: Good) to what he did.
		P18A	I think I like the second.
		P20B	I would favour the transparent.
Most unfavourite condition	$C_P(2)$	P9A	I think so.
		P11B	The first session.
	$C_{PI}(3)$	P10B	I think the second is the worst.
		P8B	(LM: The second setting is your most disfavoured?) Yes.
		P11B	Worse than the first one.
	$C_{PV}(1)$	P5B	The worst is the session in the middle.
Advantages of	$C_P(16)$	P2A	Because you then have to kind of constantly like you have that constant feedback.
		P3A	But I really enjoyed session one (P3B: Yes) as well, even though we weren't collaborating that much, but I really enjoyed it.
		P3A	You're working together, it is the one as well) so you're, already very excited.
		P5A	If you want to work with someone who knows music, you might think that the first setting is the best, and they have a high level of understanding [music and each other].
		P5B	[First] setting is also easier for people to negotiate.
		P8B	Only public space, everyone is together.
		P8B	In the third session, I was looking [at my collaborator. However, even with visibility, I still felt [the third session] was not as good as the first one, in which I could both see and hear.
		P10A	The first setting is the best because we could discuss...
		P11B	Not quite, because what I prefer is, like what he said, two people in the same space will interfere with each other in the same space, but I feel after all we need to collaborate to create something, so what I care about is getting a quick feedback from my collaborator, I care more about this. I don't like being alone, without knowing what the other person is doing.
		P16A	For example, if we're creating something together, something exciting and we know, understand each other, then the first one would be better session because then we are both listening to each other and working together that would be faster.
		P16A	(LM: So do you think the private spaces are essential?) It's sensual for those who understand the music so they can work on their own and then collaborate but those who don't understand the music and they are creating it for the, maybe amateurs, then they should do the first one.
		P21A	Because it's better to be able to hear everything is going on. I can see it being useful for experts who like who can hear the whole piece, and go there in private space and make something without interfering with the other additional music, but I love what do I think it was, actually changed anything for our session everything I hoped for our session.

	P21B	I think our communication was better when we were only working in the public space, and the first session I felt like we communicated much more than that one. Although we did not heard each other as well in the third one we can see each other.
	P21A	Yep, in terms of joy, in terms of just musical music making, yeah.
	P21B	But I do like the idea of the private spaces, I just I don't think it was very good in terms of our, what we've produced (P21A: Yeah). I think it was better just to work in the public space. My prediction was better, but it looked cool really cool, the third one (P21A: Yeah),
	P10A	However, I felt, the least stressful in the first session as I did not have anything to be responsible for. (P10B & LM: You had anything to be responsible [in other sessions]?) In the second session, I need to create something on my own.
$C_{PI}(30)$	P2A	I felt more comfortable taking risks and also I could because I have no, yeah I have no idea.
	P4B	Yes yes, because I, I felt good because I thought that I didn't but bother my collaborator about his work.
	P4B	The walls, I felt that that had indeed my private place, and I could work dependently.
	P4B	Opaque one, because I could understand, if my collaborator's instrument was located in public space, our in his, OK, otherwise I could not understand only if I heard it.
	P4B	(So you think it's more intuitive if you can hear it at the same time you should see it, if you cannot see it you shouldn't hear it) Yes.
	P5A	I think it's quite different. At least, although you were leading the process, after I had my personal space, I could just complete some of my ideas in my space.
	P5A	Yes, in this experiment, for example, you were the dominant. While you were there doing the creation, I can only passively participate. (P5B oh, from that perspective, it makes sense. The first session was quite messy. There were all kinds of sounds.)
	P6B	My only comment is I thought the private space was cool because that's something that you don't get in reality, like if two people got together and made music you can't do private space. I mean I guess you could if you plug your headphones in and out.
	P7A	[If there is no personal space, she may be there trying her ideas, which might influence what you hear.] oh, yeah, definitely.
	P8A	(LM: For testing?) Yes.
	P9A	Because I think when you have your private space then you automatically feel the need to let your collaborator know what you're doing, and also like you don't want to be rude and say OK I'm just gonna go into my private space and experiment on things, so you let them know what the plan is, how you're going to do it, how you kind of bring stuff into the public space and then you try to balance your time between the amount of time you spend in the private space and in the public space to make sure like, the, the final piece satisfies both partners I think.

P9B Yeah I think I mean so in the first session, basically, like you were in a common space so like, and that was the start of creating a collaborative music you know and things doesn't work out quite well, but then when we were given a separate space so you go there in a separate space and you walk around and you have your own privacy, you play around with music and you're satisfied, then you bring it back to public space and you both agree and release the music, so I think that was the best part. And, yeah, I think the more, it more interesting was like, so as I asked you before also the difference within season in season 2 and 3 being just the transparency on that wall, so I think in the session 3 like, you could see OK your partner is there somewhere though you cannot hear the voice, but then that gives kind of, because you have already passed the two stages, so in the third stage, so that kind of gives you more confidence and then, as result so I think we created better music at the last session.

P9A ...Actually because for me like the third one was a slightly distracting, because I found myself like trying to see like, the dots that you were making and trying to like, like unconsciously trying to adjust (P9B copy music), yeah, copy you and adjust myself to you, whereas in the second session because the only cue you have is audio so you're, just talking you can't see anything. (P9B: Yeah) like you finalise your piece and then you bring it in the public space and then you edit it there together.

P9B ...Without affecting your partner or let's say.

P9B So you, you first run your pots on your own space and, OK, OK, fine this is OK, OK, let's, let, let's give it to my partner.

P10A In the second one, because I was creating independently, so I had the best feeling.

P10A Second one, perhaps. Because I could not see you, and you could not see me, so I can focus on creation, and then do the discussion later.

P11A So I don't need to distract to see what colour she is using.

P11A That is I could be very concentrated and finished my own part first, and also...

P11A Yes, I like it. And I think the second and the third are better than the first because while making music together, we could create some messy sounds, so it is better the collaborator not being able to hear these sounds, avoiding affecting the collaborator. For the second setting, I was able to create a piece within a quiet background. I could drag it to public space once I feel satisfied with it, and then the other person could hear it. Based on this, you could continue to create your own. If she doesn't like it, you can drag it back to private space to modify it, and then drag it to public space. This is what I feel. So I think the second and the third are better.

P12B What basic and then done just play and what is, yeah so if you're on same same level, it's easier, but if it's one of us in one. We only had one space, and it's really hard so it's easier to when there are musicians they'd like to solve what they'd like to do their own stuff, and, and if you work together be a musician, maybe in the same space it can be problematic as well, so it's better if they separate, but it's, it was better when we didn't see each other.

P12B Because we, it made us talk more, so (LM: You mean) that's that's weird because, because we would, I was thinking that if you see each other it's easier, but it's not because we then didn't feel the need to talk. I didn't feel the need that much to talk and I was like oh I can see her I'm gonna put it in here, but it didn't for the work I didn't see that much, so actually.

P12B Session two was actually something like OK came up (P12A: Yeah yeah) like a base, yeah, for eight minutes this is like creating the base.

	P16A	(LM: You mean the opaque wall [helped] you to be more focused?) Yeah.
	P18A	So this isolation for me was probably I mean, I don't know, if it was, was counterproductive because I made more in the third, because I thought, if I make more I can then get rid of one again if that doesn't fit see what he's doing.
	P20A	And I can think of from my own practice writing music, sometimes writing in the same room as somebody is really important because you talk about what you do and you gain and generate new ideas based on what you're doing, and the conversation pushes forward,
	P21B	Yeah I think the fact that it was solid like the walls was the wall wasn't transparent meant that we had to communicate more.
	P21B	No, think the opaque wall was easier, it forced us to collaborate more, so...
	P21A	It is also better because it doesn't give you any visual distractions, because when you're looking into the public space when it's semi-transparent, you can see what's happening and get also in their pot in their private space, so you can tell what's in the public and with some private necessarily. If you have a quick look, if you have a quick look you can tell obviously if you have a quick look at necessary at all what's happening. When it is opaque, it is simpler.
	P21B	Yeah, minimised distraction as well.
$C_{PI}(59)$	P1B	I think yes the transparent is better than the opaque one because in the transparent I can it's likely you want to go a bit silent and create something and then bring it, but you still can see what the other person that the other person is working, for example in tempo, guitar. (LM: OK) If you memorise the colours.
	P1B	I think I can get prepared, no I don't know.
	P1A	Anyway, the most difference is, the most different thing for me is the wall and I prefer the transparent one because that the second one that the wall is not transparent and it makes me a bit weird that some at some point, he's, he's gone, he's behind the wall. Yes, and I don't know what he is actually doing behind the wall. And the third one is better, because when we see each other into private space we know that we may need some time to cool down and then create something new to replace the one that is in a public place or, or just to add more elements into it and we know each other's intention if the wall is transparent.
	P2B	I felt the space are larger.
	P2A	But I think it was easier to go into the personal space in the third session because you know you kind of don't want to like, sort of hide-away completely, yeah it makes you a bit like apprehensive about it's just weird you know if you're playing a game with somebody.
	P2A	Yeah so, if there's like, the, if it's transparent then it's like it's easier to go in there - (P2B: Yeah)I don't know to do it not some reason.
	P2A	I felt more comfortable taking risks and also I could because I have no, yeah I have no idea.
	P3A	But with session three even when I was standing at my own place, I could see what he was doing, that kind of gave me an idea of what I wanted to do.
	P3B	Yeah definitely (P3B agreed with P3A's saying above).
	P3A	Yes (P3B yes) so, I, I, for me I could do it and hear it and sure sort of I knew that this was in this was going to go into that the (P3B: Public piece) public peace as well.
	P3A	It was nicer because I was doing my own thing where I could keep an eye on him as well, that was really nice, notice someone I like.
	P4B	Yes, yes, because I, I felt good because I thought that I didn't but bother my collaborator about his work.

P4B The walls, I felt that that had indeed my private place, and I could work dependently.

P4A I think it's session 3, yeah. Because just by looking at the colours, I kind of know what you're working on, (P4B OK) but I don't want to overlap, I don't do two things you know.

P4A (LM: So you mean you have a better sense of what your collaborator was doing.) Yeah, and I could see the rhythms you know.

P4A ...And the colours really help. They're very bright colours so you can see from the far distance.

P4B Yes, yes, if you want to see him, what, how he works, OK, otherwise I may confuse [myself] if the wall is transparent.

P5A I personally think that the soundproof area with transparent wall is better, because I could see my collaborator was editing things over that side.

P5A Correct. Everyone is working together, not that I seem to be working alone, and then suddenly there is a person out there.

P5A Yes, without seeing what my collaborator was doing at the opposite side, I did not know what she is doing, and then I felt less enthusiastic in creation.

P5A I think it's quite different. At least, although you were leading the process, after I had my personal space, I could just complete some of my ideas in my space.

P5A Yes, in this experiment, for example, you were the dominant. While you were there doing the creation, I can only passively participate. (P5B oh, from that perspective, it makes sense. The first session was quite messy. There were all kinds of sounds.)

P5B (LM: In the last session, without his sounds, it became less messy, is this what you mean?) Yes, yes, yes.

P5A In the third session, for example, from the perspective of creation, or team creation, the third is better, that is, everyone can have a development space to complete their own ideas fairly, and then put the ideas in the centre to compare whose piece is better or worse.

P6B My only comment is I thought the private space was cool because that's something that you don't get in reality, like if two people got together and made music you can't do private space. I mean I guess you could if you plug your headphones in and out.

P7A (LM: If there is no personal space, she may be there trying her ideas, which might influence what you hear?) Oh, yeah, definitely.

P7A Yes, the last one, visible one, I can see which colour she is using, which instrument, and I can avoid repeating it.

P7B Yes, I agree with what she said. Because I later on found out that she seemed to have a drum over there, because I also had drums here. I forgot to look at it, then I found that the drum I set up basically does not sounds good, so if I had seen [her personal space], I might not set up the drum, [which can save some time and effort].

P8B The third one you could see what the other was doing, which made it possible to know how and what instruments were being used by the other person, and the other person could also know what I was using.

P8A (LM: For testing?) Yes.

P9A Because I think when you have your private space then you automatically feel the need to let your collaborator know what you're doing, and also like you don't want to be rude and say OK I'm just gonna go into my private space and experiment on things, so you let them know what the plan is, how you're going to do it, how you kind of bring stuff into the public space and then you try to balance your time between the amount of time you spend in the private space and in the public space to make sure like, the, the final piece satisfies both partners I think.

P9B Yeah I think I mean so in the first session, basically, like you were in a common space so like, and that was the start of creating a collaborative music you know and things doesn't work out quite well, but then when we were given a separate space so you go there in a separate space and you walk around and you have your own privacy, you play around with music and you're satisfied, then you bring it back to public space and you both agree and release the music, so I think that was the best part. And, yeah, I think the more, it more interesting was like, so as I asked you before also the difference within season in season 2 and 3 being just the transparency on that wall, so I think in the session 3 like, you could see OK your partner is there somewhere though you cannot hear the voice, but then that gives kind of, because you have already passed the two stages, so in the third stage, so that kind of gives you more confidence and then, as result so I think we created better music at the last session.

P9B Without affecting your partner or let's say.

P9B So you, you first run your pots on your own space and, OK, OK, fine this is OK, OK, let's, let, let's give it to my partner.

P10B Because in the third one, no one would disturb you, so you can focus on creating, right? And at the same time, I could still have a rough sense of what he wanted to do. For example, what he wanted to draw or whatsoever. Like in the third session, by seeing what he was doing, I am pretty sure, his piece is worse than mine.

P11B Because I could see what instruments he was suing and what melody he was doing, so I could have a rough sense of the music, and have a better control about what instruments. I should choose.

P11B I think it is better to see what he was doing instantly and do the discussion directly. (LM: So it was really bad not being able to see his activity) Yes.

P11A Yes, I like it. And I think the second and the third are better than the first because while making music together, we could create some messy sounds, so it is better the collaborator not being able to hear these sounds, avoiding affecting the collaborator. For the second setting, I was able to create a piece within a quiet background. I could drag it to public space once I feel satisfied with it, and then the other person could hear it. Based on this, you could continue to create your own. If she doesn't like it, you can drag it back to private space to modify it, and then drag it to public space. This is what I feel. So I think the second and the third are better.

P12B Like, but not seeing actually what what she's doing.

P12B But like that you actually we do the piano, for example, I would see that you are doing the piano, but nothing else, that would be nice.

P13A It is like, I could look over there, and see what he was doing.

P13B (LM: Did it make it more convenient for you to coordinate?) Yes, it did.

P13A (LM: Does it also increase your feeling that somebody is there working with you together, was this sense the sense enhanced.) Yes.

P13B Yes, the feeling of togetherness, creation, we could discuss.

P13A (LM: You had a stronger feeling that the other person was there.)yes

P15A The first one is too messy, hers mixed with mine, too chaotic, and then the third one feels too closed, I didn't know what she was doing. (P15B: Yes)

P15A And the second one is probably somewhere between the first and the third, like I could see her, I made mine and, and she did hers.

		P15B	No, don't have much to say, I felt almost the same. I think I still like the atmosphere in session 2, you have both private and public space, you can be independent from each other, and still being aware each other's progress, and seeing each other. Session 3 is too closed, and session 1 is too messy, neither is a good session. One is too messy and one is too closed, there was no sense of cooperation.
		P15B	Because perhaps both people are creating, like using the same instruments, which might not end with good results.
		P15A	I feels, while I was creating [in the third session], I cared the other person a bit more, but now knowing what she was doing.
		P15A&B	Being able to see is more intuitive(P15B). Feel more steady while doing the creation, there wont be a feeling of hanging(P15A).
		P16A	But if we have some other expertise, for example there's collaborator's expertise in drums and I have expertise in something else. (LM: Yeah?) Then we can do for the session II, I think, because then we have some space he will work on drums and I work on some others and then go to public and see what is going on.
		P17B	I felt like for some reason, we balanced working publicly, collaboratively and privately the right amount or better at least than in the other situations.
		P17A	I definitely felt more allowing within those sessions while I think the music we produced in those session was slightly better.
		P17B	I felt the other way, I felt the music in the second session is better and I wasn't in the, at the time I didn't feel like wholly aware of the transparency or not, but that doesn't mean to say that I wasn't, I just didn't consciously you know, observe it, but maybe, but obviously what you consciously make a note of them, what what you subconsciously are aware of it, it's a different thing.
		P17A	Well, so I wrote a lot music on my own, so I can of used to be in a small dark space, um but I felt I, I think we communicated better in the second one definitely, because we had like a clear, you could feel what was going on a little bit more, though in the third one we definitely felt more isolated.
		P18A	Like this the private space takes away that audio communication of the piece, but it doesn't take away the visual cue which I like, so I could still see the colour.
		P20A	...And I can think of from my own practice writing music, sometimes writing in the same room as somebody is really important because you talk about what you do and you gain and generate new ideas based on what you're doing, and the conversation pushes forward,
		P20B	I think if you don't have opaque wall we can actually just get the feedback of what the person is working on, instrument is working on, in case somebody forgets and deviates from what it has been decided that we shall, there should be a visual feedback, but there should be a space and we cannot hear what other person is composing or stuff like that, you can block the music but you should have a visual feedback. So I think this transparent private space should be, I would favour the transparent.
Disadvantages of	$C_P(9)$	P5A&B	Yes, in this experiment, for example, you were the dominant. While you were there doing the creation, I can only passively participate (P5A). Oh, from that perspective, it makes sense. The first session was quite messy. There were all kinds of sounds (P5B).
		P5A	If it is only the setting in the first session, it will be hard to come up. It may be that some people are more powerful in personality, the final work is just a manifestation of his personal ability. Others have no words or contributions in the final output.
		P9A	Um, the first session I think it was the least collaborative for me, the one without any private or semi-private spaces.

	P9B	Yeah I think I mean so in the first session, basically, like you were in a common space so like, and that was the start of creating a collaborative music you know and things doesn't work out quite well, but then when we were given a separate space so you go there in a separate space and you walk around and you have your own privacy, you play around with music and you're satisfied, then you bring it back to public space and you both agree and release the music, so I think that was the best part. And, yeah, I think the more, it more interesting was like, so as I asked you before also the difference within season in season 2 and 3 being just the transparency on that wall, so I think in the session 3 like, you could see OK your partner is there somewhere though you cannot hear the voice, but then that gives kind of, because you have already passed the two stages, so in the third stage, so that kind of gives you more confidence and then, as result so I think we created better music at the last session.
	P9A	But then at the same time if you want to try things in the public space, it will get a bit messy I guess. (P9B & LM: Yes) Because your partner is trying things at the same time.
	P10B	Yes, that's the case, isn't it? In the first session, we interfered with each other, which affects the efficiency.
	P15A	The first one is too messy.
	P17A	The first one I felt was too messy.
	P17A	I felt we would plan over each other, I can't really like get a hand, I was doing and I was distracted by you being like there, because I did like that kind of taking stuff to a private space. Then kind of bring it back, so I don't know, so I said I prefer the third one, I think I've heard the sound of the third one, but I don't know if I preferred the experience. So make sense?
$C_{PI}(18)$	P1A	Anyway, the most difference is, the most different thing for me is the wall and I prefer the transparent one because that the second one that the wall is not transparent and it makes me a bit weird that some at some point, he's, he's gone, he's behind the wall. Yes, and I don't know what he is actually doing behind the wall. And the third one is better, because when we see each other into private space we know that we may need some time to cool down and then create something new to replace the one that is in a public place or, or just to add more elements into it and we know each other's intention if the wall is transparent.
	P2B	That moment is she hiding in the personal space I have no idea what's going on.
	P3A	Number two was the hardest because I couldn't see what he was doing (P3B: Yeah).
	P4A	I think it's session 3, yeah. Because just by looking at the colours, I kind of know what you're working on, (P4B OK) but I don't want to overlap, I don't do two things you know.
	P4A	(LM: So you mean you have a better sense of what your collaborator was doing.) Yeah, and I could see the rhythms you know.
	P4A	And the colours really help. They're very bright colours so you can see from the far distance.
	P4B	Yes, yes, if you want to see him, what, how he works, OK, otherwise I may confuse [myself] if the wall is transparent.
	P5A	I personally think that the soundproof area with transparent wall is better, because I could see my collaborator was editing things over that side.
	P5A	Correct. Everyone is working together, not that I seem to be working alone, and then suddenly there is a person out there.
	P5A	The second feeling I had was I felt quite awkward.
	P5B	There was a time during which I couldn't see you.

		P5A	Yes, without seeing what my collaborator was doing at the opposite side, I did not know what she is doing, and then I felt less enthusiastic in creation.
		P8B	Regarding the privacy, I just did not know how many music pieces the other was doing and what types of instruments he is using.
		P10B	In the second, we could not see each other at all, it is like...
		P10A	Making something carelessly, with no cooperation.
		P11B	Because actually, at the beginning of the second session, we agreed what we wanted to make, and I roughly know it, but I felt if we had not, I would not know what he wanted to make, and that does not feel good.
		P15A	The first one is too messy, hers mixed with mine, too chaotic, and then the third one feels too closed, I didn't know what she was doing. (P15B: Yes.)
		P15A	It influenced, and I had a sense of unknown, not knowing what instruments she used and what she chose.
	$C_{PV}(8)$	P4B	(LM: So you think it's more intuitive if you can hear it at the same time you should see it, if you cannot see it you shouldn't hear it?) Yes.
		P5B	Because for me, the third session increased my workload, which is, I needed to drag them out of the private space, why should I do that? It is absolutely unnecessary.
		P8B	Um, in the third one, although I knew what he was doing, I did not know hear that.
		P8B	In the third session, I was looking [at my collaborator. However, even with visibility, I still felt [the third session] was not as good as the first one, in which I could both see and hear.
		P9A	Actually because for me like the third one was a slightly distracting, because I found myself like trying to see like, the dots that you were making and trying to like, like unconsciously trying to adjust (P9B copy music), yeah, copy you and adjust myself to you, whereas in the second session because the only cue you have is audio so you're, just talking you can't see anything. (P9B: Yeah) Like you finalise your piece and then you bring it in the public space and then you edit it there together.
		P9A	Yeah, because unconsciously you can't help to look at what they're doing and try to you know adjust yourself, so it makes sense with their music.
		P12B	Yeah, and you cannot really. So to be honest, and I try to see and it's it's makes it harder because and I think that yeah it's like a loop, oh, I don't need to communicate. Because people don't communicate and they don't need to and like, that, yeah, so. Something that that there should be not, they shouldn't discuss information so much that be donating communicate after all, so it is really important.
		P19B	The second is a bit messy. (LM: You mean you felt messy when you were able to see what the other person was doing?) Yes.
Adding personal space helpful	Helpful (21)	P2A	I did use it in the last one just because you know I wanted to kind of mess around but I didn't want to introduce all this noise into the stuff that already sounded like really nice, em so actually I did find that it, it was helpful but then whilst I was in that space I kind of stopped quite quickly because even though I didn't really feel like I was finished, I wanted to reintroduce it to the back into the room, so that like you know it was like you could kind of hear what was going on and be able to like incorporate that with you rather than just like could I hear the magic thing so but it was not it was nice to just...but I do think that it is useful for that purpose.
		P2A	(LM: You think it [was] useful?) Yeah, yeah it was.
		P4B	I felt that that had indeed my private place, and I could work dependently.

	P4B	I felt good because I thought that I didn't but bother my collaborator about his work.
	P5A	I think having the soundproof area is still very essential.
	P2A	Yeah it was nice that it was an option.
	P5A	I find it useful.
	P5A	I think it's quite different. At least, although you were leading the process, after I had my personal space, I could just complete some of my ideas in my space.
	P5A	If the focus is cooperation, it is helpful, I felt it still very helpful, because it can better inspire everyone's (P5B: Participation) Yeah, participation.
	P8B	Adding personal space is not unnecessary, it is necessary, but I still prefer the first condition.
	P8A	Yes, I think so, it is helpful.
	P9A&B	(LM: So both of you agree that the addition of the persona space is helpful.) Yes (P9A). Yeah, I think so (P9B).
	P10B	I don't think so, I think the first one is quite bad.
	P11B	(LM Do you think adding personal space is helpful? I think to some extent, yes.
	P14A&B	(LM: But you both of you feel like the addition of the private space is very helpful.) Yeah yeah yeah (P14A&B).
	P14B	I believe it there must be a private space. (P14A: Yeah.)
	P16B	Oh, it helped because in the first one, we don't have any private space, and we were listening to each other's sound but in the second one, we have our own private space and we [were] listening to overall.
	P18A	It's helpful but maybe with the, that you still have some cool of, of communication either visual or audio so you.
	P19B	(LM: do you think adding the personal space helped?) Yes.
	P20B	(LM: I see, so do you think the addition the primary space is helpful?) Yes
	P21A	(LM: So you think the additional private space is helpful) I think it's helpful, yeah, I just mean you know in our case, it's not, but for people making real music, I think it would be very useful.
Depends (6)	P2B	I think this depends on your music background but I don't know her maybe if she saw the virtual stuff She [doesn't] know what peach it is but for me I know exactly what peach going to be so when I press eight I know exactly it's going to be.
	P2B	You are so gentle, that depends on the person, for me, I'm not a shy person.
	P4A	If you don't have that much experience then of course, private spaces is good, good, good.
	P5A	Well, it depends on the purpose of the tasks. You said the task was to make a song?
	P10A	Adding personal space, because we are already very familiar with each other, so, private space does not really matter, I feel.
	P21A	But I think our level the private space is not that helpful...
Not helpful (10)	P1B	No no no no, it did not make any difference to me. I tried to use it at some point but I didn't think it was.
	P2B	Yeah, for the first and second one especially when I know the melody I just the tap in the melody I know and for the third one actually I'm just a trick I bring her sphere and I see what's in her score and I just do something according to her score so I don't need a testing space between.
	P2B	(LM: OK and so you don't think these the difference of the walls affect your composing, do you think so?) I don't think so.
	P2A	Yeah, me neither, because I didn't really use the wall.

		P4A	No no, so for me, because we had such a short time, eight minutes I think I was too short to go away and come back and but maybe that was because I played instruments before I'm not sure whether you've had musical experience.
		P4A	No it's something it's very important but not for this short amount of time.
		P5B	I don't see much difference, for me. Probably because I was the dominant composer in the experiment. Which is...
		P5B	(LM: Then you think that private space didn't help you?) No.
		P7B	Well, but I think there may be someone who wants to use it, but for me, adding that one [personal space] is useless.
		P18B	(LM: So you don't think the addition of private space is essential.) Um, I don't think so.
Whether the three conditions are different	Different (17)	P2B	The space are different I feel the first session we didn't have the, the our space and I think the public space is smaller, I have that kind of feeling don't know maybe because...
		P2B	Yeah maybe, virtually, it felt a bit smaller. Fortunately our theories are.
		P4B	Yes I felt when you at the boundary, the walls, I felt that that had indeed my private place, and I could work dependently.
		P3A	Yes, I did at least so for me such number two was the hardest because.
		P5B	(LM: Do you think there is any difference between these three sessions?) Yes.
		P5A	(LM: Do you think there is any difference between these three sessions?) Yes.
		P6B	For me I feel like the first session the second, session, (LM yeah), like the first session was the most different because it didn't have private space.
		P8A	In the first session, I could hear his notes.
		P8B	The first session is the most complete session, the second one shields things in the private spaces,
		P9A	Yes definitely.
		P11A&B	Yes.
		P13B	Yes.
		P13A	Yes.
		P14A&B	Yeah yeah, I thinks so.
		P16B	Yes, like...
		P21A	Yes.
		P10A	Yes, in the first session, we could discuss with each other, in the second [we did] quite independently, we could not see nor hear, in the third, I could still see. These differences might have some impacts.
	Not very different (3)	P1B	For me I didn't decide [notice] that they the walls were actually obstacle or anything.
		P7A	There is no big difference.
		P18B	Um, honestly, I'm not sure how much the difference yeah.
Reporting LeMo system	Advantages of system (16)	P4B	Yes, yes it was a great experience, because it was my first time using (P4A Oh really?) reality.
		P4A	Thanks, I really enjoyed that.
		P6B	Yeah I thought that was cool, it was quite...
		P6A	It's really fun.
		P8B	Nice, nice, the experience is pretty good.
		P9B	I think I find, I mean all the setup everything it was cool, I mean I really enjoyed creating a music.

	P12B	This is good because, because of, because if, if somebody had doesn't have big expertise in music, they still can create something because if you put all the harsh notes, it's really hard to work with them you need to know them more or less. While we like but you have half notes you have two two times when when you have half notes. Between the dot Mi Fa and Ti Do is half notes, those are half notes so do all the other ones are one one whole notes from each other, so, and, if you play two half notes together, like next to each other, that's not gonna sound good. That's what about your ear maybe. Sorry.
	P14A	But only the serving thing was, the global track I mean the public tracks well like I could hear it so them didn't make sense to me so much.
	P14A	Experience was fun I mean.
	P14B	Yeah, it was fun...
	P17B	I was, I was really impressed with the design.
	P18B	Yeah, I really really enjoyed, it, it was really...
	P19B	Give the possibility to choose, like to choose what type of personal space.
	P21A	Yeah, I felt I thought really like if a lot like you were pressing something, even you didn't have the tactile sensation. I felt really like you were pressing something which was nice.
	P21A	Yeah exactly that's I think yeah I like that a lot, that's nice. (P21B: like interactive or something.)
	P21B	I really enjoyed the how aesthetically pleasing it was, it was so pretty I didn't want to leave.
Disadvantages & bugs of system (13)	P5B	I think that there are a lot of faults on my side, that is, the tracking [recognising gestures] not sensitive enough.
	P5B	Oh, my hands are often flipped.
	P6A	Yeah I tried to look and then those like things doesn't appear on the screen until you're close enough to them.
	P8A	I feel the manipulation [interaction] of the interface is still not very flexible.
	P9B	And sometimes the button does not click. You have to repeat the same button. (P9A: Also technical problems.)
	P10A	I see, so apart from the hardware problem, I felt the experience is a bit laggy.
	P10A	Another thing is my eyes felt quite stressful when wearing it.
	P10B	It is too virtual...
	P14B	Another thing is quite often I find it difficult to push the buttons (P14A: Yes) because the fingers it was my finger was straight, but the leap motion was reading, you know I was pushing the one at the bottom or I wasn't working one of them I couldn't turn it off at some points. Yeah exactly, so many by mistakes and then they turn it over and you turn on the one below.
	P14A	Yours, really keep it and your very more accessible places right I would change that yeah and and like non-global pitch didn't make any sense to me, I mean if we are collaborating might as well have the same global pitch right? Yeah I mean why would I why would I be in C and he would be in c-sharp, I mean we want to be the same the, the pitch, it should have been global is my point.
	P15A	That is the sliders, even I withdrew my hands already, it is still moving.
	P19B	But I think while inside the personal space, being able to hear the public space feels a bit.
	P20B	Just the thing, that where the system gets overloaded, when we try to compose more than, the frame rate.
	P20B	Just the thing, that where the system gets overloaded, when we try to compose more than, the frame rate.

Suggestions for improvements (31)	P1B	No, just a suggestion (LM: Yeah) because I found that it's missing something to turn on and off really fast the patterns, so you create a pattern and I wanted to alter between two different drum patterns in the end. (LM: Yeah) but every time I had to design like um...And design meaning or just change it really fast (LM: Yeah) so that means that, because I was pressing the wrong buttons and I didn't have so much control. It would be better to have some I think something to you can tap it to l it or tap it to disable it. You know what I mean.
	P1A	Or other functions like go back, redo.
	P3A	No, but you should give the tasks as well so that so that people can, can try the tasks, so that they're already they are already going with the flow where they know they have to make the music and so the task.
	P3A&B	Not 8 minutes, probably 5 minutes(P3B). Yes, probably 5 minutes are enough so, because you need to get used to the environment and environment and play around with the spheres and then sort of figure out how you're gonna collaborate (P3A).
	P3A	...More time...For the sessions for the training.
	P6A	I don't know, I think sequencer for me comes from a very hardware related perspective and it's something that is on the table and... (LM: But that would require much more musical background.) P6A yeah just saying like every instrument you choose, it comes with a like all of the context that comes with this right, and sequences are really need not collaborative things, in my experience of sequencers, this is what I mean.
	P6B	I miss the ability to improvise, so I was trying to figure out ways to improvise like either change it all the time, before it repeats but it wasn't fast enough to do isn't like that, so maybe like mute at every fourth time.
	P6A	And, I think if you're looking, if you want to do like research about hand movement it make sense to use Leap Motion. But what you are doing, I think it is much better to use the controllers. This is my feel.
	P6A	Yeah, because the, the nice thing that I like in with the controller of the vive controllers is that they track really, really accurately (LM yeah yeah), so it's really give you the perception that this is your hand (LM: Yeah) every small failure in tracking breaks this perception so if your hand doesn't track perfectly, you think that it's not your hand anymore, it feels like something else, when it's suddenly.
	P9B	Don't have any background on music or stuff.
	P9A	Yeah I agree it was very immersive like, I, I really felt being present in the virtual environment and like you see the face and hands of your partner which is nice (P9B yes) because you really like you're actually something physically touching things you know.
	P9A	Would be just to make the space between the last row of notes and the controls like the eraser switch slightly bigger because I, I had the difficulty that if I wanted to hit I was afraid I would hit erase or switch and like I had to be really careful and not did so that actually made things worse it's probably better if you just go for it.
	P11A	I think it is better to add 1, 2, 3, 4, 5 on the left of the interface.
	P11B	I think the scene could be modified.
	P12B	Eight is not like enough because you have half notes as well.
	P12A	I think yeah I think the time is bit. It's a bit stressful, cause you know timing and then until you realise like how we were doing.
	P13A	No, no, I mean, another session with some improvements. Like, the one in my private space, I could be able to not hear the public space, this way, I can hide myself in side personal space, doing the creation and then put it into public space. Otherwise, if there is something in the public space, I can not work inside my private space.

P13B	And the pitch, the first slider, the one controls all the interfaces, it is better to be independent, control them separately.
P13B	Another thing is could the space be a bit bigger, the main point is because you want to have four [interfaces], you always feel there are collisions.
P13A	Allowing people in private space hear the public space.
P14B	What instead I think would be really useful is when I'm in my private and I'm working on that to have a mute button to get rid of what's in the public space (LM: Yeah) because sometimes when I am focused on what I'm doing without hearing the other things, sometimes there were overlapping sounds, and it was difficult to hear especially with the drum what is this doing, I don't know. The mute, if I could mute the public space, while I'm in the private so I can focus on there and then bring it out then that would be...
P14B	Still very limited, you only have those you don't you can't do very quick notes especially if you're doing the drumbeat those. (P14A: Yeah.) P14B and you have only four sounds, that would be good to have a bit more, like synthesisers.
P14B	Yeah, cannot make long notes also it's just basically one be just four section, it's very short, you cannot create the melody.
P15A	Can the avatars be a bit more detailed, like you have to option to choose a girl or boy?
P16A	If there's something like you can add a feature like to mute the sound of the bubble, easy, because we have to turn one, open it and then mute and that.
P17A	Yeah, I agree, I think the thing I like [to add] is on like the drum-kit descriptors of what the things were, things I couldn't keep remembering what like different...
P18B	Um, I think this, this sounds could be a bit better, that's all.
P20A	Yeah, it's also like if I wanted to make a long note, (LM: Yeah), I can't or join notes together, I can't change the individual amplitude or velocity of each of the notes, which was quite limiting, only having one octave is really limiting to think about reversing all of your chords within an octave, I played the imaginary piano a few times, the track workout where all my chords were in, in a single octave if that was, I think a limitation. So just I think but that's the thing I guess yeah it's a create an entire production.
P21B	I had a few other suggestions that I added on this I don't know if I can remember them all but maybe, being able to lock some of the bubbles, from like they look, like if you're sure that you really like it, and you don't want your collaborator to change it. Maybe like lock it.
P21A	I had two suggestions, I can't remember one of them, but the other was, it's just is more of a developer thing than an actual issue, but the ability to have two different scales, there's a second seven note scale but if you have, like a major in a minor scale for all the keys, it will hold on for dimension, but it's not really a, it's not a criticism.
P19A	Can the rhythm be independent?

Table B.1: Thematic Analysis Results of Study II

Appendix C

Material of Study III

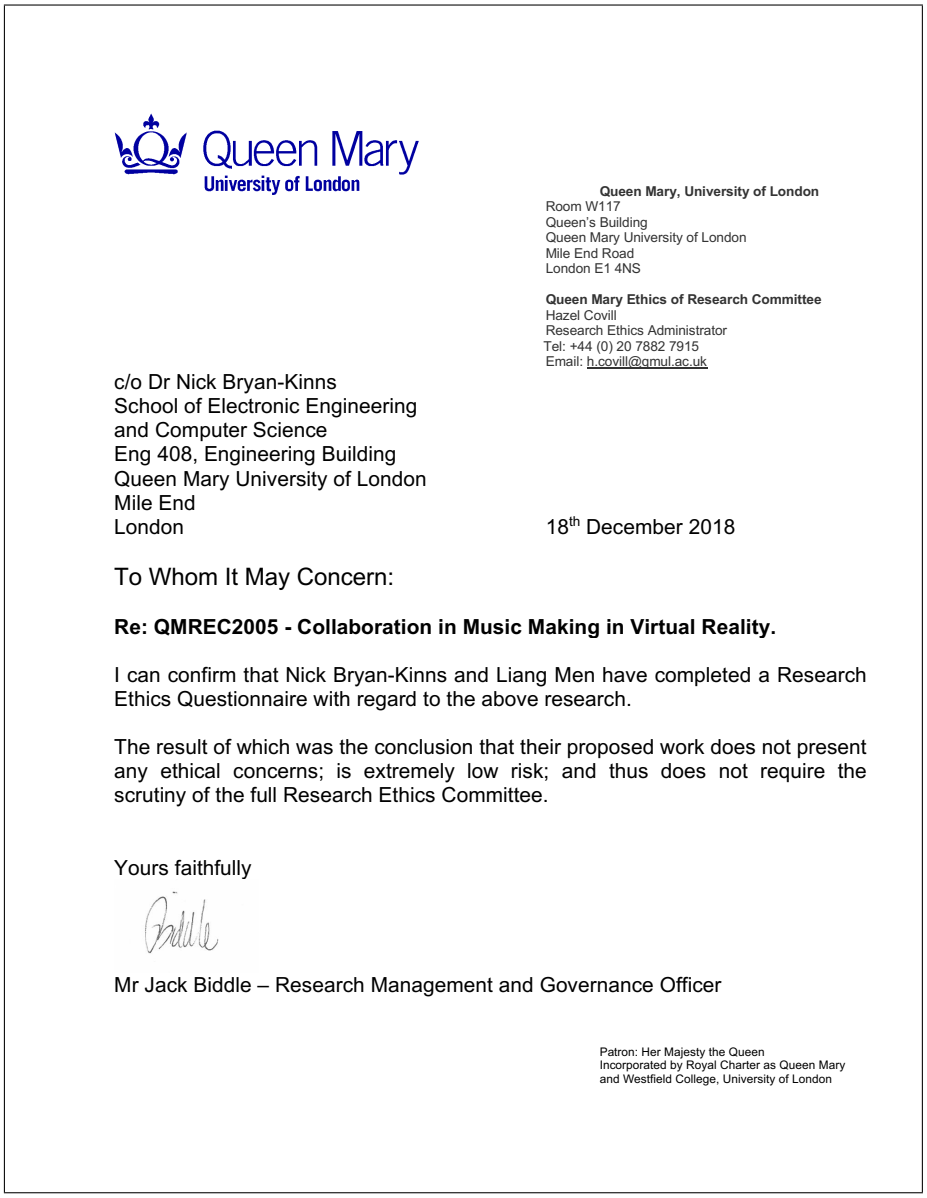
C.1 Ethical Approval

See the ethical Approval in Figure C.1.

C.2 Questionnaire

C.2.1 Demographics Questionnaire

- Your gender. (male/female)
- Your age. (text input required)
- How well do you know your collaborator in this experiment before the experiment. (Four options: 1. I know him/her very well; 2. I met him/her several times before, not know well; 3. I met him/her only once before; 4. I don't know him/her before the experiment)
- How would you evaluate your musical theory knowledge? (1-no theory at all; 10-theory expert)
- Do you play a musical instrument? (yes/no)
- If the above is yes, which instrument(s) and how long have you been playing it/them? (open-ended question)
- How would you describe your experience of composing music together? (1-no experience at all; 10-highly extensive)
- Have you ever written a song or made a piece of music? (yes/no)
- If the above is yes, how many? (open-ended question)



Queen Mary
University of London

Queen Mary, University of London
Room W117
Queen's Building
Queen Mary University of London
Mile End Road
London E1 4NS

Queen Mary Ethics of Research Committee
Hazel Covill
Research Ethics Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Dr Nick Bryan-Kinns
School of Electronic Engineering
and Computer Science
Eng 408, Engineering Building
Queen Mary University of London
Mile End
London

18th December 2018

To Whom It May Concern:

Re: QMREC2005 - Collaboration in Music Making in Virtual Reality.

I can confirm that Nick Bryan-Kinns and Liang Men have completed a Research Ethics Questionnaire with regard to the above research.

The result of which was the conclusion that their proposed work does not present any ethical concerns; is extremely low risk; and thus does not require the scrutiny of the full Research Ethics Committee.

Yours faithfully

Mr Jack Biddle – Research Management and Governance Officer

Figure C.1: Ethics Approval for Study III

- How familiar are you with computers? (Three options: Beginner; Intermediate; Expert)
- Have you ever experienced virtual reality? (Four options: 1. I have not tried it, this is my first time; 2. I have tried it, but only once; 3. I have tried it 2-5 times; 4. I played VR frequently.
- Have you used collaborative software in real time before? (Yes, I have/No, I have not)
- If the above is yes, could you please give details? (open-ended question)

C.2.2 Post-Session Questionnaire

The question items below apply a 10-point-Likert scale.

- I think the spatial configuration in this session was extremely helpful for creativity. (1-not helpful at all; 10-extremely helpful)
- I feel like the spatial configuration in this session was extremely helpful to support the development of my own ideas. (1-not helpful at all; 10-extremely helpful)
- I enjoyed the spatial configuration of this virtual world very much. (1-did not enjoy at all; 10-extremely enjoyed)
- I always had a strong feeling that my collaborator was there, collaborating with me together, all the time. (1-not at all; 10-extremely so)
- How satisfied are you with the final piece of loop music you two created in this session. (1-not satisfied at all; 10-extremely satisfied)
- How would you rate the quality of communication between you and your collaborator during the session. (1-very low quality; 10-very good quality)
- I had a clear sense of what my collaborator was doing. (1-not at all; 10-extremely so)
- The amount of your contribution to the joint piece of music is. (1-insubstantial; 10-substantial)
- The amount of your collaborator's contribution to the joint piece of music is. (1-insubstantial; 10-substantial)
- What do you think of the quality of your contribution to the joint piece of music is. (1-very poor quality; 10-very good quality)

- What do you think of the quality of your collaborator's contribution to the joint piece of music is. (1-very poor quality; 10-very good quality)
- Any reasons behind your above ratings on the contributions? (open-ended question)

C.2.3 Comparison Questionnaire

In this questionnaire, participants were required to choose the session that matches the description.

- In which session, you enjoyed the spatial configuration the most? (most enjoyed/least enjoyed)
- In which session, you made the music you were most satisfied with? (most satisfied/least satisfied)
- Which session you found most difficult to track collaborator's activities? (most difficult/least difficult)
- Which session did you have the strongest sense that your collaborator was there working with you together? (strongest/least strongest)
- Which session did you have the best quality of communication between yourself and your collaborator? (best quality/worst quality)
- Which session had the best setting for creating a good piece of music collaboratively? (best setting/worst setting)
- Which session did you find most difficult to cooperate with collaborator? (most difficult/least difficult)
- Which session do you feel you made the most contribution to the joint piece? (most contribution/least contribution)
- Which session do you feel your collaborator made the most contribution to the joint piece? (most contribution/least contribution)

C.3 Thematic Analysis Results

Parent Code(N)	Code (N)	Participant	Segments

Learning effect	Learning effect(40)	P1B	I'd rather say like the later ones were simply for the reason that we get more used to it.
		P1B	...get used to it after 1 or 2 session.
		P2A	Have never collaborate music via a screen.
		P3A&P3B	(LM[1]: Do you think that's more because of the sequence effect or because you are getting more familiar and more master) P3A & P3B: Yes.
		P4B	Um, we got better as we went on it.
		P4B	Especially in the last session we enjoyed it.
		P4A	(LM: So you think it is more because of the sequence effect) Yeah, it is an important factor too.
		P4A	Towards the end you are getting better.
		P5A	First one was a bit difficult, only because we were being introduced to this space setting, so was quite different in terms of adjusting to it, you know being able to work with each other in the space setting
		P5B	As we were progressing into the further sessions. It was easier to communicate and use the hand gestures to create the sound.
		P6B	I think it is just because at that point I understood what was going better than before.
		P7B	As sessions went by, we were able to collaborate more.
		P7B	So first when we did it for the first time, we didn't know what he is doing, what our interests are, but as sessions progressed [things getting better].
		P8A	I think like I was getting in used to to the to the task, so I don't think in the first setting we communicate a lot.
		P10B	Must be because you've been practising for like three sessions so the last session is like better. We knew hot to do it...
		P11B	We were best used to it and we, we thought we were good at it late
		P11B	The first one was an element of surprise, a total surprise like (P11A: It was the first time) the first time we were using [the system]
		P11B	And in the fourth one, we were actually getting good at it...
		P14B	...Because by the time we got to the fourth session.
		P14B	We already knew what we are exciting, like personal space this is how we did use this personal space and
		P14B	Because in the first session we did collaborative work, then we went to our personal space but that didn't allow us the correct time and everything.
		P15B	Cause we knew what to do
		P15B	I mean I was more comfortable moving around in the fourth one,
		P15B	Like you're more used to it and you were more comfortable...
		P15A&B	(LM: So it's not due to the settings, just due to the sequence you took sessions?) P15B: Yes, possibly. P15A: Possibly yes. That's definitely a factor.
		P16A	So in the first session, I think most of the, most of the time was spent at least in my case at still figuring out how the system was working.
		P16A	And also practising with the feedback between my action, and what's coming back from the system, so I can't say that in the first session I was really focused on the task.
		P18A	Would have made a difference but we were more used to it.
		P19A&B	P19A: I think the collaboration improved with session and session. P19B: Yeah acquaintance with the environment
		P24A	For me, it was session one just because it was our first session, it did not go very well...
		P24B	I guess we were still kind of getting used to it.
		P24B	So session 1 as well.
		P24A	I couldn't keep track of where xname(P24B's name) was creating her personal space and where I was creating my personal space.

[1] "LM" refers to Liang Men, the experimenter of the study.

		P25B	I think it was quite difficult in the first section.
		P25A	Yeah, because we did not get used to these equipment.
		P25A&B	(LM:...Due to that it was the first session and you are not very experienced with it?) P25B&P25A: Yeah.
		P25B	I think around 50 per cent. For two or three minutes we're just trying to press the buttons and hear to hear the sound that given.
		P25A	Yeah, feel the same (P25A agreed with P25B's point above).
		P26A	Maybe the familiarity was the main reason as the sessions progressed
		P26A	I became more familiar with the system that might be an issue so I will not deny that point.
Most favourite condition	$C_{pub}(12)$	P6B	The last one
		P6B	I kind of get it was the fourth one...
		P10B	We enjoyed the last session the most.
		P11A	I think in the third one, we did not have personal space but we could hear each other's sound where I was standing.
		P12A	So my preferable is session 3...
		P12B	Mine was between 3 and 4
		P16A	In the second session I thought, I feel that I enjoyed it
		P17A	I think the first one is the most realistic,
		P17B	Yeah (P17B agreed with P17A's comment above)
		P18A	Yeah, the final one.
		P19B	I like session three.
		P21B	I think the first setting
	$C_{aug}(34)$	P1A	Probably the second one for me.
		P3B	I prefer session 2.
		P6A	The first session.
		P6A	The one was in session 1, that interesting one so no personal space and I really like the really strong sound attenuation thing, that was my favourite.
		P7A	Space sitting I would say on demand was best.
		P8B	I like the session one the most because.
		P9B	The last one.
		P9A	The last one was the one I enjoyed the most...
		P9A&B	(LM: So you prefer the fourth session the most) P9A&P9B: Yeah.
		P10A	I think the fourth one
		P10B	Me, too. (P10A agreed with P10A's comment above).
		P11B	Sound attenuation it was really good.
		P11A&B	(LM: So you prefer the final one the best) P11A&B: Yeah.
		P11A&B	P11B: The final one the best. P11A: The fourth one.
		P11A&B	(LM: So you prefer the final one the best [second time of asking this]) P11A&P11B: Yeah.
		P12B	Mine was between 3 [C_{pub}] and 4 [C_{aug}].
		P12B	The fourth was good though
		P15B	(LM: So you prefer the setting [in] the third session the best?) The third one, there was sound attenuation, I mean I don't...
		P16A	(LM: You mean the boundary is too solid) Yeah exactly. (LM: So you prefer like the attenuation one?) Definitely.
		P18B	I definitely thought the first one.
		P18B	I'd go for the first, first setting.
		P18B	Yeah, I think I like yeah, number, number one.
		P19A	I like session four better,
		P19B	Setting, I guess it is four if, if the performance was the same, four would be better because you can
		P19A	(LM: So you do think the sound attenuation helps) Yes.
		P20A	Yeah, I prefer the third one.
		P20B	(LM: Third session is the one you, you[you two] favour [the most]) Yeah.
		P22B	OK I like that...

		P24A	Yeah in terms of like the personal space the use of with special session 3 because there was none, like you have to just move away from the sounds so we were working better together I think.
		P24B	Mine is the third one...
		P26B	I really enjoy the fourth session.
		P26A	Fourth one, the last one.
		P26A	Yeah, I agree with that.
		P26B	Yeah.
	$C_{fix}(10)$	P5B	Session two.
		P5A	Sessions two.
		P14B	We like the last one.
		P14B	May be stationary is better because you kind of like, it gives you
		P14A	(LM: So you prefer the stationary one at the opposite corner) Yeah.
		P15A	Yeah definitely the last one,
		P22A	I prefer the second one I prefer the second one with their isolated space.
		P22A	The second one is my favourite one.
		P24A	Space setting, I think session 4 worked out the best.
		P24B	Separated spaces at the end
	$C_{mov}(19)$	P3A	I enjoyed a lot in session four
		P4A&B	The fourth one.
		P4A	Yes, I think its because on the fourth, maybe like, we worked better like communication.
		P5A	Like session 2.
		P5B	Session 2.
		P7B	Session three.
		P7B	(LM: You prefer the third setting the best?) Yeah.
		P8A	Also my favourite one was the third one.
		P8A	But I liked it [session 3] very much.
		P2A	Fourth one, fourth one.
		P13A&B	P13B: I think we both felt like the session 1 was the best. P13A: Yeah.
		P18B	The third one.
		P18A	Yeah yeah that one, yeah.
		P18A	The third one when the personal piece comes over.
		P21A	I think I like the last one the most, where we could go into the private space anywhere.
		P22B	But I like both of them.
		P22B	The drop one.
		P25B	The fourth one for me.
		P25A	Yeah (P25A agreed with P25B's comment above).
Second favourite condition	$C_{pub}(8)$	P8B	The one without personal spaces...
		P8A	Yes, me too.
		P9A	[Session] one.
		P9B	Session one.
		P10A	The third session.
		P10B	No, I think my second favourite would be the third one as well.
		P11A	(LM: So the third one is that you prefer that as the second prefer)ah, yeah the second preferable...
		P13A&B	(LM: And which is your second favourite one, the third?) P13B&P13A: Yes.
	$C_{aug}(6)$	P12B	(LM: Yeah so that's the second favourite session for you?) Yes.
		P19B	I couldn't tell much. difference between three and four, but in four it was much more obvious that you could push the sound away and attenuate.
		P22A	The first one.
		P24A	(LM: The sound attenuate one is your second favourite one?) Um [P24A means yes].
		P25A	For me, [session] 2 or [session] 3.

		P25B	I liked the second one because I was close to her.
	$C_{fix}(4)$	P15A&B	P15B: The fourth one, not the first one. [laugh] P15A: Fourth one, yeah...
		P24B	And I think the fourth one is the second best because it's a bit more structured than the first, like if you want to go into your own personal space you actually have a whole like section four, and it went like have the middle section for the music and then go into your personal space then bring it back over, it just needs a bit more organised.
		P25A	[Session] 2 or [session] 3.
		P25A	For me is the third one.
	$C_{mov}(3)$	P12A	Session 2 I think,
		P15B	Like in the first one, I was like should I move out, where do I go.
		P26A	OK, so I would rank from one three - two -four- really - yeah I see.
Most unfavourite condition	$C_{pub}(9)$	P7A	No space at all.
		P15A&B	P15A: The one we could not use individual, like (LM: The handle?) P15B: Yeah the handle or the sound attenuation, There was one we didn't have that.
		P15A&B	P15A: (LM: It's the session 2 you don't have the handle) P15A&B Yeah.
		P20B	(LM: Which do you dislike the most?) Second I think.
		P20A	Yes, maybe the second cause it was too chaotic.
		P22A	(LM: Which to dislike the most, the final one?) The final one.
		P22B	(LM: Which to dislike the most, the final one?) The final one.
		P25A&B	(LM: OK, so both of you dislike the first session the most?) Yes.
		P26A	Three, one, two, and four...
	$C_{aug}(0)$	-	-
	$C_{fix}(16)$	P7B	(LM: So which session do you dislike the most? No, no the session the space setting) the corner ones.
		P8B	The one that is in the corners with...
		P9B	The one with the when we were popping out to the insulate boxes at the corners.
		P9A	Yeah, me too.
		P10A	I think the one with the (LM: Opposite?) Yeah.
		P10B	Corner, yeah.
		P13A	Yeah. I think the fourth session definitely, wasn't the last, because I feel like we both stuck to our corners in the personal space, and then only when we brought our separate work together, we just didn't collaborate quite well.
		P13A	The fourth one.
		P16A	The fourth session is the one that I enjoyed the least.
		P17A	I would say the third one (P17B: Yeah) because it is...
		P17B	(LM: That's also your the worst one you think?) Yeah.
		P18B	Session number two.
		P18A	Um, probably yeah, probably two, the separated thing, yeah,
		P26B	OK, I dislike, maybe [session] three.
		P26A	With first one because I was not able to concentrate on one bubbles and simultaneously I could not hear the effect of all bubbles together and work it on together.
		P26A	OK, so I would rank from one, three,two, four. Really.
	$C_{mov}(3)$	P19A	I think [session] two.
		P19B	[Session] two I think.
		P24A	For me, it was session one.
Advantages of the condition	$C_{pub}(25)$	P3B	I think once we are in our own space, we didn't collaborate that much. So in session 1 and 2, I think we collaborated much more than session 3 and 4.
		P6A	It was much better in the one without personal spaces...

P6A	I thought in the personal spaces, we were kind of like I was forcing myself to like to work on myself, and finish it and share it, whereas when it was like all in one [P6A means C_{pub}] you kind of have to work with the other person.
P6B	You had to cooperate with the other...
P7A	In the session that without personal space.
P6B	I can hear more...the first one, when I cannot hear your stuff, I got lost
P6B	Yeah, when I can go from that, it was easier to act.
P12A	Actually. There was no personal spaces but I thought that we could collaborate a lot more together.
P12A	I only had to listen to each other's own sounds OK to agree on something using our own personal spaces where we could feel like OK the son might sound nice to me when we had to bring you out so you know let you offer your collaborator you might have to change it a bit to suit them whereas if you're working into one environment without the personal spaces it was a lot easier to, like, see our different ideas and put it into together.
P13B	In session 2, it was like we're at the end we were just like working very closely side side (P13B: Yeah, I think that one was the one like we were most collaborative (P13A: Yeah).
P13A	We were like affecting each other's results.
P13A	The second one I feel like we collaborated the most is because we were constantly hearing each other's so we could constantly tweak our own as well as realise oh in yours it's that the 1 dot off so then like I went over and like removed one of your drums I think (P13B yeah) so the second one even though it was harder to like work on your own personal one, collaborative wise, it was the best.
P17B	I think we only collaborated in session 1.
P17A&B	(LM: OK and so you feel you should always be able to hear everything that would be better). P17A: Including voice, real world voice. P17B: Yeah.
P17A	I think the first one is the most realistic.
P17B	F the intention is to collaborate and you, you can use that to have your own ideas and then put those ideas together.
P17B	To be honest, I think it'll just be easier if no matter where you are you could still hear what the other person's done and then you if you didn't want to hear it you could just have that quick (P17A: Yeah) like personal space selection (P17A: Yeah) and you can just.
P19A	(LM: So you mean you would like to hear all of them) Yeah.
P19A	All of them together that makes much more sense to me.
P20B	I thought we were more collaborative when we didn't have the bubbles.
P21B	I think the first setting I found most easiest to come up with ideas.
P21B	But for making maybe it's I think it's just better to be able to hear what's happening [all the interfaces] or maybe there's an option to turn it on and off I don't know.
P23B	Cause I think the piece of music we created was like more together I felt rather than being something, like one person do something and the other person just put together, it felt more when we could hear each other it was a lot easier to make something float better together rather than.
P23B	Prefer being able to hear everything that sometimes I can hear like Money [her collaborators name] speaking and so then that member I mean it's like remover headphone, something like that. But being that it made it harder to create something from scratch, but don't fit, but at the same time I prefer hearing everything, so I when I was using a personal space I'll be in the person space making idea step out a bit, trying to hear everything, step back in, step out, step back in and kind of like it would be.

	P24B	In other session, we don't have personal space, I think we had discussed music better.
$C_{aug}(111)$	P1B	Closer together was the better ones because we collaborated more.
	P1B	Because we stood like this and then the one we stood like this we could see each other's things.
	P1A	We could hear them quite clearly without moving back and forth.
	P1B	One with the sound attenuation was quite nice because if we wanted to like not hear other person's for a sec, take a step back and take a step forward again.
	P3B	I think once we are in our own space, we didn't collaborate that much. So in session 1 and 2, I think we collaborated much more than session 3 and 4.
	P3B	It helped for collaboration but it was more difficult to make [music]
	P3B	No, it does [did] not help for creativity, but it helps for collaboration because we got closer to each other.
	P4A	The volume drops maybe it was less stressing.
	P4B	I like I liked it for the same reason though, that you can, um, you can kind of spread out your instruments and the less important ones you like I put that over there and I'll take it back later when I want to join it in again.
	P5B	Like if you want to work on your ideas without getting interrupted with the other one [FE yeah] maybe that session four was good for that purpose.
	P6A	It was much better in the one without personal spaces,
	P6A	We were kind of like I was forcing myself to like to work on myself, and finish it and share it, whereas when it was like all in one (P6A indicates the Pub only condition) you kind of have to work with the other person.
	P6A	Because remember if I wanted to like like move away slightly from something my partner was doing I could just step back a little bit and not have to hear I've been having to go all the way to the other side of the room and even then I could still kind of hear some of the things.
	P6A	It's like I could have a little bit of personal space, but it wasn't like a defined thing I had to keep moving in and out of, I could still share my work quite easily and we could also be close together so we could work on the same things rather than having to be like inside our personal spaces.
	P6A	(LM: So it's kind of like a middle point) Yeah, between personal space and no personal space.
	P7A	I would say the sound attenuation instead of completely blocking the sound of what the other is doing, the attenuation is actual.
	P7A	Yeah, if I want to actually listen to what he is doing, only then I can make my own music, so what I can do is I can reduce my volume and listen to what he is doing, from that far end itself (LM: Aha) without going into his space.
	P7B	Yeah like for material instruments I was able to understand the instrument better.
	P7A	Yeah, that's what I also like because there is no personal space with the collaboration was much more better much much better also if at all we need some personal space we will take our what we call the widget[interfaces] to our corner and then because of attenuation you won't be able to clearly listen to what the other person is doing, the music doesn't overlap so, then if we want to listen to what the other person did, we just walk closer to.
	P7A	(LM: So [you mean] the first one [first session] kind of like give you a sub-personal section it's not totally personal but kind of [a] lower degree personal space) Yeah.
	P8B	(LM: You can have some personal space) Yeah, yeah.

P8B	(LM: Without even trigger something?) Yeah yeah, exactly, this was the most [convenient].
P9B	The last one was the session that forces you to collaborate more the most because you had to stay very close to compose music, another [session] tends to be a little more, um.. So each of us on his own.
P9A	Because it was easier if, for some seconds I just wanted to step back trying to think of something a melody, or something else, I could just take a step back and I couldn't hear anything...
P9A	The last one was more immediate and easy to handle, in my opinion, and I enjoyed it more.
P9B	Yes session 4 for the way because I mean it's easier to shut all the outside noise and it forces you to stay close to each other.
P9A	But if I step back I can still hear it a little bit, so I can focus on my own instrument without having the background too noisy [still hear other interfaces a little bit]
P9A	In the first one, you can always keep up to date because you can hear it in the background and also if you want to hear any specific instruments [you can step forward]...
P9A	I like the fourth one because if you really if you just wanted to listen to one instrument as louder than the others you could just put it closer, in the others you just hear the result everything in every second, but if you want to hear, [e.g.] OK, let me hear the drum what it's doing if I can improve it, what is going on there, I can just go closer (P9B: Yeah) and I can hear it but I even if I still listen I still hear the background.
P9A	Session 4 has already good [attenuate settings].
P10A	We can both like create on our own and right after see our change.
P10B	Because we could hear in real time what was happening.
P10B	My reason is the same that we could hear in real time what was happening so that it was easier to collaborate but I also think I'm biased towards the last session because the video was like a bit more like it was it evoked bit more feelings than the other ones.
P10B	Yeah so we enjoyed creating it together, so that's why the setting also played good...
P10A	Yeah, I think also another reason is that you could try to take some music off and then see how we change the [music], take it off and take off the distance.
P10B	You could change the distance to it so that you know you can hear a bit of it so that you have an idea what's happening, but it doesn't bother you.
P10A	Yeah but at the same time, create something, yeah.
P11B	We could like we could bring it away and no one would be able to hear it and as soon as I came close to his sound bubble I could hear everything, right.
P11A&B	P11B: You just come closer. P11A: Like in the real world.
P12B	Part of the fourth one was how we were still able to listen to the music that we made together but we were able to like walk away and create music while being able to hear the other music, which was very useful, because in the personal space we could not do that
P13A	The increase attenuation helped, because you could step back a tiny bit (P13B: Yeah) and it would like mute a lot but you could still hear it softly in the background, I feel like you wouldn't be as influenced.
P14B	It provides subtle help but I think you'll be more recognised by someone who's more familiar with sound like audio files.
P14A	Yeah, the usefulness would be for someone who uses it a lot, someone who knows it better, specialises in the field...
P14B	Someone who's like a very sound experts that would be able to recognise as well. (LM: You mean recognise the difference or recognise the helpfulness) recognise the helpfulness as well.

P15A&B	Cause like sometimes when they're all close together it's hard to see what you're...P15B: What I'm doing. P15A: Doing and like to hear what, what I'm doing, this sounds good, this sounds bad, yeah so that's why the last one like I am I
P15B	Um, when I was creating my own music, it was easier for me to perceive what I was doing, and not get confused with what she was doing, so that was a good thing.
P15B	We could tell the difference
P15A	When I was making one of the pieces it was like I literally can tell if that was like coming from me or from someone's music (LM: Ah, I see, so you mean you can locate the source of the sound better)
P16A	I think having the personal space, I mean I most enjoyed when I could decide in a continuous way if I was able to listen to the other sound sources or not, well in the personal spaces it was more like an on/off feature.
P16A	Definitely because in this way I can decide (LM: Like what?) to what extent I want to isolate myself.
P16A&B	Because when you engage in such a short time task, like seven minutes you have to create a musical pattern. P16B: Yeah, um. P16A: And you really need to listen to the other participant.
P16B	The time constraint we have actually made that we actually don't want to just isolate and...
P16A	I would prefer just to have the possibility of stepping back and make my..
P16B	I think the third was really nice but done really for the reasons of different ways to make music rather than of it, as I said earlier, of something to distinguish my stuff from the other's.
P16B	Yeah.
P16A&B	P16A: Yep...Yeah. P16B: Yeah.
P16A	Where the, the sessions that I enjoyed the most, and probably the third one was the one when I would choose.
P17A	It could be useful if its, maybe if you're trying to do with many instruments and like an orchestra so you could create one workspace on one corner and like leave it there and then go to the other person and create other forms and then something like that may be useful, but when two people collaboration, it is quite confused I think.
P18A&B	P18A: What, you can hear it. P18B: Well you be kind of hear it and like bring it over to me and then you can hear me. P18A: That was really cool setting to have that. P18B: Yeah, that one.
P18B	Yeah I think because you can kind of get it from the volume they're at and I was I'm doing the drums or whatever it's like went to the corner I actually couldn't really hear her, it was effectively the same. (LM: Ok)it was like and I think I wrote this already to you again, like when you're playing with your, your guitars unplugged from the amp and you can still kind of hear the other people (LM: Yeah) and then you can plug it back in we're still there.
P19B	I think being able to attenuate sounds is quite good.
P19B	We could complete something and then put it away, in the same, but be in the same environment and then work on something new without having to use our personal space.
P19B	I couldn't tell much difference between three and four, but in four it was much more obvious that you could push the sound away and attenuate.
P19A	There's much things like that we could do but you get a sense that he would do a lot of things with that, the distance pushing away certain...
P19B	Helps mixing.
P19A	Yeah I mean you could although we couldn't but you could create a co-creative delay, I don't...
P19B	I guess you would move towards and then back again.

P19A	...And then back again you could create that effect, like, like there are options that you could explore, because to be right flank with the sound effects, there are only three or four there's piano, there's two other and the drum so I mean, there were more of a happy mood this thing so and we were just playing around but with moving our heads that could create some other effects which I liked a lot.
P19A	(LM: So you means you can localise which instruments which) Yeah yeah.
P20B	I thought we were more collaborative when we didn't have the bubbles.
P20A	We did not have the personal space thing in session three.
P20A	Yeah, but session 3, we can step back and then not hear it so that was good because sometimes you don't have to hear all the things together, and if you want to work on something individually, you just step back.
P20A	(LM: So you men...you still want to hear a little bit) Yes (LM: But not very clearly just a little bit of it) Yeah.
P20B	Also, it is easier, if I want to go for a few seconds to see what she is doing and then go back to not hearing anything it's easier than doing the handle, I'm going, and telling her remove your handle, oh, so I can hear what you're doing and then so it's easier, I go for a few seconds hear what she's doing and then I go back to mine where I don't hear anything
P20B	Sometimes I just want to focus on the exact thing I'm doing (LM: Yeah) without the blur of other thing.
P20A	Sometimes it's good to hear a bit of it in the background but not completely
P20A	(LM: In order to match it. To make something new and matching the old. like no it's better to match) Yeah. So like, if I want to see if it goes with the other one, I don't want to hear the other one completely but just a tiny bit so, just like lean a bit closer and then hear both of them together and then lean back and then not hear that other one, then continue working on them
P21A	I liked it because you could go away and still hear a little bit in the background and then bring it back, but in the end when we listen to the final piece it was hard to find the right spot where I can hear them all
P21B	...I think I was thinking about the whole thing in terms of making a composition of music (LM: Yeah) I think if I think about the experience spatially that it also includes the visuals then I actually like the idea that you can get closer, far away from the sounds origin.
P21B	Yeah I guess when, it's kind of set up in a way that and you think about making the music in relation to what you see, (LM I see) I think the, the sounds increasing or decreasing based on how you walk to it is interesting if you think about the environment, like that you're building an environment (LM: Yeah) and but my head was kind of thinking about create a piece of music (LM: Yes) yeah as opposed to kind of the environment.
P22B	Because like if I want to see, I could I could vaguely hear what xname [P4A's name] was playing and what he was doing.
P22B	I feel like I could like hearing what you were going with it would just help me like understand and tailor what I was.
P22B	Yeah like you could see you could hear it a little bit what what was going on, but and that could help you then amend...
P23B	Cause I think the piece of music we created was like more together I felt rather than being something, like one person do something and the other person just put together, it felt more when we could hear each other it was a lot easier to make something float better together rather than...

P23B	I think that helped in some cases and then other times it didn't because it meant that as soon as I moved like one step away I couldn't hear the rest of the sounds or less I was stood right in it which I think I don't know wrong.
P23B	...whereas if I could just hear slightly, then it would be OK but that's not that much of an inconvenience. It doesn't mean that it works quite well that you can get on with your own idea and still hear other things by just stepping out, so yeah.
P23A	Because it allows me to know what she was doing so I can adapt to what she was doing no matter where I was standing, and it also helps get more consistent volume and if I was standing closer I would love to hear something but then the moment I walk away sounds slightly different so it's more consistent like the sound that we're.
P24A	Yeah in terms of like the personal space the use of with special session 3 because there was none, like you have to just move away from the sounds so we were working better together I think.
P24B	In other session, we don't have personal space, I think we had discussed music better.
P24B	I thought it was quite realistic and made it easier to kind of like move around in the sounds and work I think.
P24B	Not really sure, I think I enjoyed it because when I was creating the sounds I wouldn't have to go into my own space, I could just move away a bit and then move towards another piece of music which just seemed a bit more natural to me, so it made it easier for me to work with the sounds.
P24B	(LM: Without like triggering the handles) Yeah.
P24A	Because with the third session you have to work collaboratively a lot more, and it's kind of cool to like it was fine just get quieter and quieter which is.
P24A	Yeah, like you work with your collaborator really well in that one because you don't have the use of personal space, so they can actually isolate your instrument and you can isolate their instrument. I think it was better working with your partner in that one.
P25A	Not so useful because we can just keep away from each other and we cannot, we don't hear anything.
P25B	Like once you step away from the panel, then the music like the freedom decrease quite significantly.
P25B	I liked the second one because I was close to her.
P25B	(LM: You mean it kind of pushes you to stay closer?) Yeah.
P25B	(LM: And [you think] that's better for the collaboration?) Yes, I think so.
P26B	The fourth session is like a real time experience a real time experience or just moving here and just listening to music yeah it was good and the private area was a good concept I really loved it OK we can just say choose our own music and work on it yeah it was a good context.
P26A	Yeah I agree with him about the experience from all sessions indeed the fourth one was the most productive session.
P26A	Fourth one the last one. The most productive session, but we had optimised use of personal space and deteriorating sound with respect to distance it would.
P26A	Yeah the distance stuff and as well as personal space that would work best.
P26A&B	P26A: I think I will go with volume drops. P26B: Yeah, the fourth session yeah it was good like we were feeling like the real-time experience just going out and, and we obviously if we need to don't need the sound that we can just drop the volume off and work on it but I mean drop it off yeah drop it off some volume.

	P26A	One more thing in fourth season it was possible for us to bring our bubbles together and see the collaborative effect of all bubbles together simultaneously I could change two bubble.
	P26A	But in the fourth one, I mean photon I mean I could able to configure a bubble while maintaining some distance from other bubbles.
	P26A	I mean later on I can bring those bubbles together And test the collaborative effect.
	P26A	Then again I can pull them off and work it on separately.
	P26B	It was in fourth season that distance was the good one because we can always just leave one corner and just move outside and you shouldn't listen to that you'll just get to give a different sound from how it's like for example you will use sound and if you listen from far it seems to be the better one sometimes just from listening to from near you don't know it is good one...
	P26B	(LM: You mean originally it does not sound nice, but when you move far away with a volume drop, you hear and form a different feeling?) Yeah.
	P26B	And fourth one is the one that work on real life.
$C_{fix}(22)$	P3A	In the session 3 and fourth, we used our own creativity much more comparing with session 1 and 2.
	P3A	As in session 3 and 4, we do [did] our own music and then we combined them, yeah.
	P3B	It is because everyone is, they are on creativity. [with personal space], I can do according to my mind, and he can do according to his. Which is better for creativity, but for collaboration, it is not [good].
	P3A	In session three and fourth, we used a lot the personal space.
	P7A	Personal space is actually an added advantage we can't deny that.
	P7A	Yeah that is one thing but it doesn't matter since the boundary is small I don't mind.
	P8A	Having a self-space is good because you can develop your own idea...
	P9A	I want to develop something new, but just because maybe I wanted to think about a possible something to add.
	P14A&B	P14A: Yeah, organised. P14B: A bit more separate. P14A: You feel like OK, this is your personal space and then when you want to show your work you go to your public space so then that gives you like a place to think of where you want to put your work I think.
	P14A&B	P14A: More than that, it prevents you from clashing. You know session. P14B: At some points, we were very close to each other.
	P15B	Um, when I was creating my own music, it was easier for me to perceive what I was doing, and not get confused with what she was doing, so that was a good thing.
	P19B	Be able to walk to your personal space as much easier.
	P19A	Yeah, you can walk there create your own thing but when you do this like if two people are collaborating, like for a moment his sound went off when I created that.
	P19A	The personal space, so I think the one with the corners is better, you have a clearer idea that you have to go there.
	P22A	Think it's very useful because if you want you can join this space and create your piece without, for example in the first session, I was all. confusing so you couldn't hear anything it was all a mess, with the public space you can isolate yourself and create your pieces.
	P22A	In session two we did great because we were with our personal spaces.
	P22A	Because I did like the personal spaces but I liked them when I could move into it instead of it just coming around me.
	P22A	Yeah to be quiet that's the correct expression yeah so if I know that are in the corner, that feel like more personal.

	P22A	Exactly and especially if I am close to him or his sounds is if I deactivate my personal space anymore if he activated his so it doesn't feel very personal instead, in the opposite I have my space to create at the end then we put together in the middle that's perfect.
	P24A	Do work on like a single instrument we could go to personal space but otherwise you could just like go into the middle zone, warn hear all the, the instruments together. I think it was a good mix of the two, it was a good mix of session one and session three.
	P24B	And I think the fourth one is the second best because it's a bit more structured than the first, like if you want to go into your own personal space you actually have a whole like section four, and it went like have the middle section for the music and then go into your personal space then bring it back over, it just needs a bit more organised.
	P25A	F we move from our personal space, we still can hear each other. But for the second one, we need to be really close to the bubble I made.
$C_{mov}(46)$	P1B	The ones where we stood closer together was the better ones because we collaborated more.
	P1B	Because we stood like this and then the one we stood like this we could see each other's things.
	P1A	We could hear them quite clearly without moving back and forth.
	P2B	I like that one but yeah but I prefer the way you have your own personal space cause as soon as you've done, you can remove it instead of like taking the bubbles out kind of thing.
	P3A	In the session three and four, we used our own creativity much more comparing with session 1 and 2.
	P3A	As in session three and four, we do [did] our own music and then we combined them, yeah.
	P3B	It is because everyone is, they are on creativity. [with personal space], I can do according to my mind, and he can do according to his. Which is better for creativity, but for collaboration, it is not [good].
	P3A	I did my own music so it was fun.
	P3A	In session three and fourth, we used a lot the personal space.
	P4B	I was much better to have it come down immediately when you, in the fourth one wherever you are yeah. It's basically like a mute button that you can just press and that activates only for you.
	P4B	Like just moving around with you.
	P4B	If once the cylinder comes down that it moves with you.
	P4B	So essentially it's like having a mute button just for your ears.
	P4B	You can activate that no matter where you are, activated.
	P5A	Could get rid of the handles as well when we wanted to, that's where we produce the best music, the communication [gets better].
	P5B	Because you have a choice to move around.
	P6A	I thought in the personal spaces, we were kind of like I was forcing myself to like to work on myself, and finish it and share it, whereas when it was like all in one, (P6A indicates the Pub only condition) you kind of have to work with the other person.
	P6B	We behaved completely different, Cause I know we are the only person in the personal space, bring out.
	P7B	Personal space is actually an added advantage we can't deny that.
	P7B	Yeah whenever personal space are to demand, especially session 3 like I can put it the personal space anywhere, that was the best advantage.
	P7B	I think dropping the volume is more easier than attenuation, I don't have to walk and listen to it and come back, I can just reduce, listen and then go back.
	P7B	It's easier um, easier to do the volume.
	P7B	We did not had to walk from one corner, in the third session. We actually discussed our ideas and immediately I went to my personal space.
	P7B	Yeah more convenient.

P7B	It's not only helping me to create my own music but also to understand the instrument better.
P8A	Because you have like these hands available whenever and wherever you want.
P8A	So I think like having it in the corner well having.
P8A	Having a self-space is good because you can develop your own idea.
P8A	Because you have like space and, and you can create your own (LM: OK but also like keep an eye to the other.
P9A	I want to develop something new, but just because maybe I, I wanted to think about a possible something to add.
P12A	That one we could actually use the personal space and you're pretty fine when I felt like, um, maybe to have something black on your own, you might sometimes it's good as well so that you feel, I don't know, you can understand what you, you're trying to think of, rather than listen to someone else's one as well. So you could listen to your own one in your own personal space and then bring it out, tried it sounded the same or sounded well with um, with the, the other pieces of music.
P13A	I think cause the fact you could move the portable, the personal space made it easier to just like like pull on it if you want to work on something quickly and then just get rid of it.
P15B	Um, when I was creating my own music, it was easier for me to perceive what I was doing, and not get confused with what she was doing, so that was a good thing.
P18B	It's the one where it comes around, you, you pull it down and it's like you are inside.
P18A	(LM: So you do think the personal space helps.) Um, hum.
P21A	That, you, like if they didn't like it all when the private space was the corners because you can't, you, I wouldn't go to the other side into this space, but like this you can just go into the other one's space and you would know more.
P21A	It is convenient so I can make my own thing, but then you remove it and then you don't have to drag the bubble around.
P22A	Think it's very useful because if you want you can join this space and create your piece without, for example in the first session, I was all confusing so you couldn't hear anything it was all a mess, with the public space you can isolate yourself and create your pieces.
P22B	So I like the option to if you're, it, like if you just need to focus on one sound then you can just put the space [personal space] up and then you're fine, but I put of the two of them. I prefer they're just moving away.
P22A&B	P22A: Because in the second one. P22B: I feel that everything can just be too close to you if it's just wherever you are.
P22A	(LM: Ah, you mean you are creating something in a place where it does not belong to you, so you kind of feel less personal.) Yeah, exactly.
P23B	I preferable to hear everything that sometimes I can hear like Money [her collaborators name] speaking and so then that member I mean it's like remover headphone, something like that. But being that it made it harder to create something from scratch, but don't fit, but at the same time I prefer hearing everything, so I when I was using a personal space I'll be in the person space making idea step out a bit, trying to hear everything, step back in, step out, step back in and kind of like it would be.
P23A	So I prefer the second one because the personal space was a lot like I was able to move so if I wanted to work on an idea really quickly when I was standing next to xname [P23B's name] could quickly work on that idea, and then immediately I could I share that with xname [P23B's name], she could also make her own changes.
P23A	(LM: Like more efficiency) Yeah.

		P25B	You can still choose wherever you want because in the third one that by side by standing in the diagonal, that's too far away from the other.
		P26B	OK more like you just have to go to one place and conscious not like you can't get anywhere you want in second one you just can get a private place anywhere you want and you don't have to move like OK if you create two music and if you want private space then you have to go to your own place but in second it was that quite easy more convenient I think.
Disadvantages of the condition	$C_{pub}(16)$	P1A	It wasn't very creative stimulating for me but then the reality is rendered forward.
		P2B	I feel the first session to be quite difficult. He was trying to work on your own. music you going to be really those it wouldn't be, because they wouldn't much and couldn't do it by itself beforehand, we wouldn't be able to hear the music.
		P2A	For example, if I was a music professional and I am with someone who has a different level of expertise...If you have different equivalent, it may get a bit annoying. Depends on the one who you are with. For example, if you had a safe zone, he didn't have a safe zone, and it was global music, and there was someone annoying, da, da, da, stupid noise, you are going to be rude, you are not going to say anything.
		P9A	Or the others [other sessions] I had to like any where I was going the music was always very loud, so sometimes I just wanted to like take the volume lower.
		P11B	Was a bit troubling because we are too.
		P15B	Um, when I was creating my own music, it was easier for me to perceive what I was doing, and not get confused with what she was doing, so that was a good thing.
		P20A	Maybe the second cause it was too chaotic.
		P20A	So the second is the most chaotic one.
		P22A	It's definitely too confusing.
		P22A	I mean it's you don't have the, the tranquillity to operating your sounds or the everything's come mixed, which is difficult to manage.
		P25B	That's was quite annoying.
		P25A&B	P25A: Because is just like the public space. P25B: It was also annoying every every one was making the same thing, we can not concentrate.
		P25B	Because it was just annoying to hear all the noises (LM: Why would you describe that as noises rather than music) Especially we made a lot of, like the note (P25B: yeah) we made four bubbles.
		P25B	We just had a lots of [music notes/interfaces] together (P25A: For fun).
		P26B	Yeah, I felt quite hard with everything open and quite noisy And you can't just go far like in season four. And it was not kind of practical one while working.
		P26B	But the tall ones seems like, um, three [he meant session 3] the tall one is like practical reality, but in virtual world, I don't want like that kind of stuff.
	$C_{aug}(19)$	P1A	(LM: Do you think that is also helpful to create own ideas, create own music, and then feed it to the joint piece.) I think that would be quite hard, because then the pieces might clash in a music way. (LM: You mean they are not in tune) Yeah, exactly.
		P3B	It helped for collaboration but it was more difficult to make [music]
		P3A	We didn't use it too much this [kind of thing, refer to the sound.
		P4A	Also it makes it difficult and to hear all the instruments at the same time.
		P5A	For collaboration I don't it's convenient or practical.
		P8A	I was kind of confused what was like my like my music or if it was like our joint music.

	P12A	The only thing I did not like about the final one is that space wise [she meant the sound attenuation], it was wasn't too convenient because you had to literally be next to each other to hear the sound well, and when you actually make the music, does, it takes up a lot of space as well so you're really crowded in that sense. Um, so you couldn't really be much spaced out to listen to each other's.
	P13A	Pretty much like it wasn't like I was like you know it was an major increase in a way. Just slightly, that would be the harder probably I found.
	P16B	But then it also influences the piece because (P16A: Oh, yeah), then we might have another type of compositions.
	P17B	(LM: Do you think that helps?) Um, no.
	P17A	It kind of confuses.
	P17B	So you cannot work out how it sounds all together.
	P17B	But then you're further away from the one you've just made, so then you can't hear that one as well and then it kind of, [form a problem] yeah.
	P17B	To be honest, I think it'll just be easier if no matter where you are you could still hear what the other person's done and then you if you didn't want to hear it you could just have that quick, (P17A: Yeah) like personal space selection (P17A: Yeah) and you can just...
	P18A	While, not necessarily because I felt like when I hear it all I needed to hear it so I because every time I went off and to do that my own space or I couldn't hear and it was like a bit separate, I wasn't listening as much, so I think the collaboration wasn't as good, well I think it was when I could hear everything.
	P22A	I think it's very useful because if you want you can join this space and create your piece without, for example in the first session, I was all confusing so you couldn't hear anything it was all a mess, with the public space you can isolate yourself and create your pieces.
	P23B	That helped in some cases and then other times it didn't because it meant that as soon as I moved like one step away I couldn't hear the rest of the sounds or less I was stood right in it which I think I don't know wrong.
	P24A	Yeah, but it wasn't like certain space, like she could choose anywhere and I felt like sometimes it was just difficult to manage with the other instruments, like putting them in your personal space is a bit difficult to manage compared to just having that one certain space...
	P25A	If we move from our personal space, we still can hear each other. But for the second one, we need to be really close to the bubble I made.
$C_{fix}(56)$	P1A	I have to take my interface move it to the corner then move it back to hear what he was doing and this made it harder I think it was in session 3 or 2 session 3.
	P1B	I think session 3 is a good set up but in practice it was like quite likely to move the things it was quite hard to grab them.
	P1A	I like doing over the spaces were like hmm opposite but as I said in practice it didn't work out that well.
	P1A	Otherwise, it would be, if it was easy to just grab it and move to in to a thing.
	P2A	Yeah because I don't know why would you put yourself in the box and have to be walking over you know I mean...
	P2A	You have to make extra effort, you want to focus on the music instead of having to be walking, yeah just be everything's where you're standing.It is not a game.
	P2B	Possibly, because you could work, like to work by yourself for too long and would not want to come out, it would be more difficult to collaborate because you are working on yourself.

P3B	I think once we are in our own space, we didn't collaborate that much. So in session 1 and 2, I think we collaborated much more than session 3 and 4.
P3B	It is because everyone is, they are on creativity. [with personal space], I can do according to my mind, and he can do according to his. Which is better for creativity, but for collaboration, it is not [good].
P4A	...The personal space that for me is not very useful because I need somebody to led [lead] me.
P4B	Especially in the third round, because I didn't know why I had to go to a particular side the part of the room in order to get the personal space.
P4B	I don't like it that it's that it stays in one part of the room and I move towards it.
P6A	I thought in the personal spaces, we were kind of like I was forcing myself to like to work on myself, and finish it and share it, whereas when it was like all in one (P6A indicates the Pub only condition) you kind of have to work with the other person.
P6B	We behaved completely different, Cause I know we are the only person in the personal space, bring out.
P7B	When we got into a personal space there wasn't much communication happening between us, because we couldn't hear what he is doing and I could not hear what I am doing.
P7A	In the personal space sessions so every time we stepped in to our own personal space we completely had no idea what the other person was doing, so what we were doing was completely out of our one creative idea, so when we brought up our music together it sounds very hard [bad].
P8B	So you don't need to go very far to get to your personal space.
P8A	But at the same time I think like having it on the corner, you can forget about your partner because you are just like concentrating and focusing on yourself instead of, so at some point, at some point of time you may forget about him.
P9A	It takes a little bit, some seconds were to take these take it down and create your own environment, go there.
P9A	Also because I didn't need particularly to like create my own environment to build something, because in the meanwhile he was going on so as soon as they get out.
P9A	In the other one if I create my own environment I have no idea of what he's changing so when I go out maybe what I developed, (was, is not) does not fit anymore because he did other changes and.
P10B	Yeah, but then it was like a bit more distracting for me like personal space this was like easier simple to understand that if you wanted to do something that, that we wanted less of the music we could just move a little bit more away, if you wanted complete quite we could just move like completely away, but personal spaces like again we have to create it and then remove it again and again.
P10B	Because the process, I didn't, I don't know but I didn't enjoy creating music in the personal space.
P12A	I only had to listen to each other's own sounds OK to agree on something using our own personal spaces where we could feel like OK the son might sound nice to me when we had to bring you out so you know let you offer your collaborator you might have to change it a bit to suit them whereas if you're working into one environment without the personal spaces it was a lot easier to, like, see our different ideas and put it into together.
P12B	I think it is because we were also on either side of the public space, sometimes I forgot that xname [P12A's name] was there.

P12A	I thought like we couldn't collaborate (P12B: Yeah) as much in the first one because um, yeah it's that feeling that you feel like Oh someone else's, (P12B: Yeah) isn't there even though they are there.
P12&B	P12B: You feel like you are on your own in your own corner. P12A: You forget, oh, I should collaborate with her, let's come together in the middle, yeah.
P13A	Yeah. I think the fourth session definitely, wasn't the last, because I feel like we both stuck to our corners in the personal space, and then only when we brought our separate work together, we just didn't collaborate quite well.
P13A	Then we had to make a last-minute changes to make it fixed
P13A&B	P13A: And we lost in the last one, like do not know.. Do not know what is my drum and what is yours. P13B: Yeah so it was just all confusing yeah we didn't work together as much.
P13A	LM: So you mean the distant personal space) Yes. (LM: Also pushed you away and less collaboration) Yeah.
P16A	As soon as we isolated I think we completely lost connection on what was happening and that was the session in which I less enjoyed the musical result.
P16B	Have the same conclusion, but we were first there, into like I'll try it out and heavier isolation
P16B	We had we were forced to be isolated and I was creating actually a big difference I am cutting our collaboration.
P16B	It's more that we were forced to use the isolation thing, I think a little bit more.
P16B	Because I at a certain point I completely lost track of what's happening [with my partner?]
P17A	I think when personal space is involved, you get lost making your own music and you like by the time you realise that you should probably collaborate with the other person, it's already too late and just, there is [are] opportunities to create like unlimited sort of interfaces, so when you done with one interface, and you're happy with something then you always have the chances to create more interfaces, so instead of collaborating with someone else you end up wanting to create your own music.
P17A&B	, P17A: I think it is useful but. P17B: For collaborating with someone, it's not.
P17B	Yeah I think it's good if you actually have the intention to collaborate, but if you're if you're just if you just making these like just end up making stuff on your own then it's pointless.
P17B	Which is the [yeah] then it's helpful but if if you're not gonna collaborate, then it's not helpful, if you are not too tempted to make your own.
P17A	And I'm creating something in my own space so her space is locked my space is locked so even if I open my space and I can't hear her until I go into her space.
P18A	You have your own space in the end, that was the last one, that was like, we did not like separate, but um, I don't know, when we had that, we also coming back together after separate and I think it wasn't actually (P18B: Yeah) as connected anyway.
P18B	I wasn't sure whether that really helped that much, I mean in the beginning but we should have like using it to put the music in, once was made a bubble and put it over there, that would be maybe more useful than being in it and I felt like in session two but we both had our own spaces.

	P18B	That felt weird because for example I couldn't tell whose space is what, there's the points like oh is that your space can I go in your space, can we swap spaces and you realises well there's only two of us in here, why do we need two of these you know we just need one space and then the outside.
	P18B	That's much more like suddenly like I guess that's my room over there, but I didn't, it was harder to follow (P18A: Where in the room), where the one in the corner was.
	P18B	You got no sense of where the corner is.
	P18B	(LM: OK so it's not very convenient to use?) Um hum.
	P18A	That was a bit confusing because that was like, I didn't know which one will be, normally be my one, I was like, is that my one? I have opened it.
	P19A&B	P19B: I think personal space wasn't that useful. P19A: Yeah.
	P20A&B	P20B: Yeah, because with the handle, you can't really know what the other person is doing. P20A: Yeah, so you can't really.
	P21A	That, you, like if they didn't like it all when the private space was the corners because you can't, you, I wouldn't go to the other side into this space, but like this you can just go into the other one's space and you would know more.
	P21A	Because we went there and we did not come back.
	P22B	I thought that you know the isolated space thing the personal space (LM yeah) I thought it's useful as an idea, but like I feel for collaboration it can sometimes be a bit, not a, it can be a bit not very useful at the same time because like if I want to see, I could I could vaguely hear what xname [P4A's name] was playing and what he was doing.
	P23A	So I prefer the second one because the personal space was a lot like I was able to move so if I wanted to work on an idea really quickly when I was standing next to xname [P23B's name]. I could quickly work on that idea, and then immediately I could I share that with xname [P23B's name] she could also make her own changes.
	P24B	In other session, we don't have personal space, I think we had discussed music better.
	P26B	OK more like you just have to go to one place and conscious not like you can't get anywhere you want in second one you just can get a private place anywhere you want and you don't have to move like OK if you create two music and if you want private space then you have to go to your own place but in second it was that quite easy more convenient I think.
$C_{mov}(26)$	P2B	Possibly, because you could work, like to work by yourself for too long and would not want to come out, it would be more difficult to collaborate because you are working on yourself.
	P3B	I think once we are in our own space, we didn't collaborate that much. So in session 1 and 2, I think we collaborated much more than session 3 and 4.
	P3B	It is because everyone is, they are on creativity. [with personal space], I can do according to my mind, and he can do according to his. Which is better for creativity, but for collaboration, it is not [good].
	P4A	The personal space that for me is not very useful because I need somebody to led [lead] me.
	P7A	When we got into a personal space there wasn't much communication happening between us, because we couldn't hear what he is doing and I could not hear what I am doing.
	P7A	In the personal space sessions so every time we stepped in to our own personal space we completely had no idea what the other person was doing, so what we were doing was completely out of our one creative idea, so when we brought up our music together it sounds very hard [bad].

P8B	I got a little trouble in session three with the surrounding and personal space being track and I think that one was worse than the nothing.
P9A	It takes a little bit, some seconds were to take these take it down and create your own environment, go there.
P9A	Also because I didn't need particularly to like create my own environment to build something, because in the meanwhile he was going on so as soon as they get out.
P9A	In the other one if I create my own environment I have no idea of what he's changing so when I go out maybe what I developed, (was, is not) does not fit anymore because he did other changes and.
P10B	Yeah, but then it was like a bit more distracting for me like personal space this was like easier simple to understand that if you wanted to do something that, that we wanted less of the music we could just move a little bit more away, if you wanted complete quite we could just move like completely away, but personal spaces like again we have to create it and then remove it again and again.
P10B	Because the process, I didn't, I don't know but I didn't enjoy creating music in the personal space,
P11B	The second one was sort of like confused and like what, what's he doing and like.
P12A	I only had to listen to each other's own sounds OK to agree on something using our own personal spaces where we could feel like OK the son might sound nice to me when we had to bring you out so you know let you offer your collaborator you might have to change it a bit to suit them whereas if you're working into one environment without the personal spaces it was a lot easier to, like, see our different ideas and put it into together.
P17A	I think when personal space is involved, you get lost making your own music and you like by the time you realise that you should probably collaborate with the other person, it's already too late and just, there is [are] opportunities to create like unlimited sort of interfaces, so when you done with one interface, and you're happy with something then you always have the chances to create more interfaces, so instead of collaborating with someone else you end up wanting to create your own music.
P17A&B	P17A: I think it is useful but. P17B: For collaborating with someone, it's not.
P17B	Yeah I think it's good if you actually have the intention to collaborate, but if you're if you're just if you just making these like just end up making stuff on your own then it's pointless.
P17B	Which is the [yeah] then it's helpful but if if you're not gonna collaborate, then it's not helpful, if you are not too tempted to make your own.
P17A	And I'm creating something in my own space so her space is locked my space is locked so even if I open my space and I can't hear her until I go into her space.
P19A&B	P19B: I think personal space wasn't that useful. P19A: Yeah.
P19A	Yeah, you can walk there create your own thing but when you do this like if two people are collaborating, like for a moment his sound went off when I created that.
P19A	The personal space, so I think the one with the corners is better, you have a clearer idea that you have to go there.
P20A&B	P20B: Yeah, because with the handle, you can't really know what the other person is doing. P20A: Yeah, so you can't really.
P22B	I thought that you know the isolated space thing the personal space LM: Yeah) I thought it's useful as an idea, but like I feel for collaboration it can sometimes be a bit, not a, it can be a bit not very useful at the same time because like if I want to see, I could I could vaguely hear what xname [P22A's name] was playing and what he was doing.

		P24B	In other session, we don't have personal space, I think we had discussed music better.
		P24A&B	P24B: It can get a little bit messy like we are just creating personal space in the middle of music kind of thing. P24A: Yeah.
Reporting LeMo system	Advantages of system(31)	P2A	I think it is useful, definitely, cause...
		P2A	I think natural, sorry, the natural gesture and allowing you to be in your own it was good, the way it was just came down yeah.
		P3A	The interaction was good, I think maybe you can add more instruments.
		P5B	Given to us was easy for carrying out the tasks, and we could communicate with each other to make a good project, or good sound.
		P5A	I also like the instructions because they were coherently delivered, we could complete the task without seeing each other completely, so just using vocal communication, so be able to produce a music.
		P9B	It was quite straightforward sometimes you have to maybe.
		P9B	Is normal. They were very intuitive.
		P10B	It is easier to collaborate when, like when we can see each other, like in virtual reality it's bit more real because you don't know what's happening in the real time, but everything else is OK. I, we enjoyed it. Think she.
		P11B	It was just a great experience, I really enjoyed it.
		P11A&B	It was just great watching each other and just working together in virtual reality environment, I mean it is much more interesting than working.. P11A and you are able to perform a task, making a music.
		P11A&B	P11B: It was just great watching each other and just working together in virtual reality environment, I mean it is much more interesting than working. P11A: And you are able to perform a task, making a music.
		P11B	In real time.
		P11A	Working together in virtual reality was a great experience.
		P11A&B	P11A: Yes, other than it, it was good. P11B: Yeah, other than it, it was very good.
		P11A	The sky thing [laugh of P11B] is beautiful.
		P11B	It was really intuitive and like the graphics.
		P12A	But I liked the fact that there were different environments to try out, yeah.
		P14B	It was good, it was really good.
		P14B	Mainly you got straight to the point and you see what you should do and you just got on with it and that was perfectly fine.
		P16A	I really enjoyed it.
		P16B	And also it's just clear enough also like to distinguish what I'm doing and what he's doing, it was clear enough because you see the sequencer you know what you were just doing.
		P20B	Overall, it is a good experience.
		P22A	Than the others because you if you are, we've agreed in the middle you can both work it together (LM: Yes) it's very good.
		P22B	And then like moving, and they're just kind of moving sound into the right direction is incredible, how that's like how it works.
		P22B	(LM: Ah, you mean your real movement is track is mapped with the virtual.) With the sound yeah yeah I think it's so cool.
		P22B	The sound of having four having the sound coming from four different places I think is really cool.
		P22B	To just try and isolate because I think there was something. I think it was in the fourth one that we couldn't understand where it was now, you're putting my head round just to see where it was [P22A: Yeah], there the issue was.
		P23B	Oh I think it's quite cool it's a good idea.
		P24B	Yeah I find it really fun as well and it was quite intuitive after we did like the tutorial session. Which made it quite easy and relaxed.

	P24B	Yeah, that was fun because you know when I do the calibration, I put it on the floor so it knows the position of the floor [P24A yeah] and then when you when you put it on, it know how height you are.
	P24A	I also enjoyed the setting, I liked the background, it really put me in the moon.
Disadvantages of system(59)	P1A	The world was moving sometimes when I would try to increase the volume or the pitch, I would go with my figure but that my finger would go inside the interface too much then I couldn't see how was I.
	P1A	And it wouldn't do anything and that was like quite common but then I figured it that so it's just like. Yeah.
	P1B	No, it was like the first time I saw the hand, it didn't match exactly like it, so our brains like right I'm doing this I'm doing this and then my eyes are like no you are not, oh, what's happening? Right?
	P3B	I think the environment the space should be in more detail, because I am not sure where to go and I am afraid of [getting] hit I guess. That possibly should be more detail, the surroundings.
	P3B	In the last session in my VR headset, the interface crashed many times.
	P4A	...the fact that we have seven minutes only doesn't have like I like the interface it like because I have seven minutes and I know I will have, I will need a lot of time to test my ideas and first explore like all the sounds to create a pattern [P4B: Uh huh] and I was more relying on him like OK what can we do and how do I do with that, like, yeah that's all, really.
	P4A	Need more, more time to explorers to really create something good, that's all.
	P4B	I already put it in my feedback but something like having a master volume control, you have at all times because if I had an idea that I want to communicate to you, it would stress me out that we hear everything so loudly I know that we can move back or create the personal space.
	P4A	I wonder if us being a couple affects the results (LM: Possibly).
	P4B	We initially thought that would be gonna like a actual video footage like.
	P6A	Only issue I had was like some kind of confirm button for erasing the sheets because sometimes I get a grabbed a sheet and accidentally hit erase and like lose that the sheet I've made.
	P7A	It had some difficult to read my gestures.
	P7A	The video went off at times.
	P7A	Sometimes the gesture is like instead of traffic tapping the ball, we automatically do like this [he is showing a gesture], naturally do like this [another gesture] so maybe because of using the smartphone.
	P8B	Because I had an issue with the hand recognition, so I cannot interact with it, once I make it a personal space I cannot interact with the dots.
	P9A	Just in some case, you had to tap it more than once because either you press the wrong button, or the next one, but I think it is normal.
	P10B	It is easier to collaborate when, like when we can see each other, like in virtual reality it's bit more real because you don't know what's happening in the real time, but everything else is OK. I, we enjoyed it.
	P11A&B	P11A: I think like the bubble thing or the moving thing P11B: There are few glitches in it.
	P12B	I found that there were some glitches, difficult when wanting to press something like it would not happen, and that like maybe took away some of our time.
	P13B	The only issue I had is sometimes my hand would um, Like I would be using one finger and sometimes like bum would pop out and then I accident pressed like two at the same time...

P13A	Yeah, precision sometimes, like I'd press one, then I note I pressed two at the same time. But that didn't happen that much (P13B: Yeah).
P13A	Also in the last one like I was trying to get this one point and then the rest of my fingers were like glitching a little bit and I tried to like undo it but it wouldn't, but eventually it worked, but yeah, I wouldn't that.
P14A	Sometimes, like when I was in the last session, I was trying to like get my hand into kind of great it wasn't really like, it wasn't really moving [correctly] I know be like because...
P14B	Thing is usually with the leap motion is that especially when it sees, you know when it sees your hand.
P14B	I don't know which sensor is but whenever either, because whenever I overlap with her on this diagonal, if one of the cameras gets locked or like
P16B	It was not so different...
P16A	With the drums and I wasn't aware of what drums were which,
P16B	It didn't track my hands very well.
P16B	Then we pressed the right ones so there was a bit of a challenge to keep calm. We was OK but still yeah I was come sometimes a bit like slowing down my work, so if that can be improved (P16A: Yes, same) that would be quite a beneficial thing. (LM: OK) and of course the tactile missing thing, but then well that's OK, I mean for.
P17A	Because you have the chance to create more and more interfaces and then...
P17A&B	P17A: Um, I think the bubble is, hum, I don't know, I found it difficult. P17B: I found it difficult to make a new bubble.
P18B	I mean in terms of the world, I think we get, we get too close to each other maybe, but I'm um, way definitely started losing whose hand was whose, and like sometimes one hand to show up as the other hand.
P18B	Kind of weird. But it is not designed for people don't like really going in front of each other's face, like we couldn't poke at each other and stop in front of that face...[unclear].
P18A	I think sometimes when I wanted to click off some of them, when I'd click them rather than having to delete all of it, I wanted to be like oh I don't want that one it wasn't recognising it.
P18A	All that confusion well kind of went wrong and the system kind of went [to wrong] it kind of like messed everything up I think at that point. But only for a few seconds, but it was quite disorientating.
P18B	When I spent a while trying to focus in the back.
P20B	Sometimes, there are, it is not very precise, I wanted to click or remove the click (LM: It does not recognise your finger).
P20B	Sometimes it clicks but it just doesn't turn off so I see it clicked, but it doesn't turn off.
P20A&B	P20A: It happened, like it's pressing but not removing. P20B: Sometimes, my hand turns reverted or other places.
P21A	I think it was hard once we started so because we couldn't hear each other very well, so I...
P21B	I kind of round it up because no information about the place for the other person's body is, I think if I knew where the body was that have more sense.
P21B	I think it was tricky because when we got that close to do that like our hands.
P21B	I think I only realised that in the last session, I could bring the bubbles down here, like I could drag them down.
P21A	I found with the interaction was like, like it was not so easy for me to press exactly the same, the point I wanted to press. I would reach into the grid and.
P21A	And then I might hit three buttons at the same time.
P21A	It doesn't give me a feedback either once I pressed it, it's like

		P21A	That works as well, especially the guitar is, really made a sound like, I don't know, music from a castle in the Middle Ages, so it was quite specific something really (P21B: yeah).
		P21B	Yeah like I think if there were things that kind of like another thing band or notes, something from synthesisers, it would be easier to illustrate the animations.
		P21B	The animations are quite like that was the thing like yeah the animations lasts for a longer amount of time than the [musical interface].
		P22A	Because maybe looking at only a daughter[unclear], a person it's not inspired by anything.
		P25B	It was quite difficult to hear the other one (P25A: Yeah).
		P25A&B	P25B: Is there something, was the tracker really close with her P25A: But I always do have some distance right?
		P25A	But you need that tracker, right? This one the HTC headsets can only look like themselves.
		P26B	In three second or something like that just pudding if you want
		P26B	While I was putting [dragging] two bubbles at the same place [time] it exploded yeah yeah.
		P26A	One more thing I found was configuring the top-most row with my finger yeah I could not reach it.
		P26A	That was somewhat tilted but I was trying to make it straight.
		P26A	Not really but there was one cross button at the top, at the bottom so when I pressed when I click that everything wiped out.
		P26A	I was not actually familiar with the functionality and sometimes I was dragging it and for some bubbles I was configuring them from the reverse side so it was a transition for me to visualise it from the reverse side and during that it went wrong and unknowingly I press that cross but...
(52)	Suggestions	P2A	Maybe more colours though more colours with the video? Yeah I, I tried to like minimise the colour, because it can get complicated.
		P3A&B	P3B: More pitch, right? P3A: Or more pitch, but it was all right.
		P4A	Like if it's for the game like because it couldn't be for, I mean not necessarily to compose, it could also be used in a school to say something, and children could, um, yeah but then you...
		P4A&B	P4A: Like on the drums like maybe having, because there are lots of different instruments [she meant drum beats] in there, right? (LM: Yeah.) P4B: Maybe having like a little picture next to its row, so it is difficult to find out.
		P4A	Maybe having colours would be help too because then you have Ideas about whether to have warm music or something thing more on the dark side, then it is just something with little movement.
		P7A	...And music I think music, more controls would have been better.
		P7A	It doesn't have more options for instruments, pitch, we could have more, like a tempo.
		P9A	(LM: Do think it would be helpful if I give you another slider to adjust the attenuation, the degree of attenuation.) P9A: Yeah, to me, yes, or...
		P10B	I had several problems, which was I didn't feel the need to create the bubbles, you know all of that you could just...
		P12A&B	(LM: So how about if I add a function, allow you to adjust the attenuation, the level of attenuation). P12A&P12B: That would be very useful.
		P12A	(So you want it to be a local temple slider, not a global one?) P12A: Yeah.
		P13B	I think it could be interesting if the tempo is not universal.
		P13A	I feel like think start off with like the same like a default tempo, or I don't know.
		P13A	(LM: So it's like normal speed double speed, four times speed and...) Yeah, that could work.

P13B	I think that would be easier, and also like with different tempos will force you to interact with your partner, have one but a local one like one (LM: Yeah) that would be like you would be matching it with your partner's tempo, (P13A yeah) it will force you to interact more, to discuss.
P14B	Maybe with the headphone as well, normally you would have like air cup headphones.
P14B	Normally how you do with it I remember is they have like they would have like a QR code or something as well, so what they would have is each person to just put a sticker on it...
P15A	Maybe you should do that like you have to work on it together like stand next to each other and like do it like at the same time together.
P15A	I mean like, but like make it like more mandatory like maybe one of the sessions because then it kind of encourages more collaboration and not like because some people don't ideas of what we should do, but I mean the best way to make sure that we both do what we both think is best in like one piece is to kind of kind of make us work on the same thing and...
P15A&B	P15B: Force them to work together in one of the sessions P15A: Yeah, because I mean like some of the sessions encourage collaboration but there's no...
P16B	Having a handle maybe something relative, instead of on-off thing, I would can could imagine that...
P16A	But I think that a slider would be from an emotional point of view...
P16A	We might have the mixer, yeah.
P16A	Yeah I think the type of interface at that point would really matter, I mean the graphic icon that is used or...
P16A	An embodied point of view I mean the idea of having a slider for example would be a bit.
P16A	So so what about like a cross-fade you know when you DJ...
P17A	If the voice is included in the VR world, then it would be helpful to collaborate as well, because otherwise you take your headset out to talk to hear the other person.
P17A	There's another thing maybe I thought when you create the personal space there could be an easier like sort of on/off button which could kind of like temporarily open your door and you can hear someone else's work too, and you can sync it and then press the button again it will close and then you can resume your work, so that will be a good way of collaboration without having to talk to each other too much because someone else is creating something already.
P17A	(LM: The personal space should only work for the owner of it rather than work for both, like, well, whether she has her personal space or not you should always be able to hear [hears]?) Yeah, exactly.
P17A	Sometimes I found it difficult to press things (LM: OK) however, on thing could be useful to actually have notes.
P19A	We can create some effects like if you, he was placing the drums and piano, we can parallel to each other, if you move your heads the different sound effects that were getting created in the head.
P19B	I'd like to have option to put one away and kind of mute it, back and forth...
P19B	I think more indication over what measures, I could see the division in the fourth session but the three recessions I didn't notice it. So it's more easier to see the divisions of four over sixteen steps and also it'd be quite useful if you could see the note or the drum on each.
P19B	I was hoping there was one and I only realised there was one in the fourth session.

	P19A	Yeah, it is too happy this thing, but because there was some like the one with the firing neurons kind of a thing, that that wasn't getting created, all the matrix thing that was in the last day the visual that was there and couldn't create that with (P19B yeah) C major.
	P19B	What could even be nice is haven is an effect bubble. Audio effect
	P21B	I kind of round it up because no information about the place for the other person's body is, I think if I knew where the body was that have more sense.
	P21B	Yeah, somehow yeah I've kind of felt um it's like it feels like you're kind of isolated until you see the other person, maybe if I could see like legs or feet even, you know actually I found a tricky note that because you told us to go to the pointers.
	P21B	And I think yeah get feet or more parts of body somehow would be easier to get sense of collaborator.
	P21B	Would have liked for the for the tempo to be different on each (LM: Yeah, so individual tempos])...
	P21B	More like I suppose space for spontaneity in a way instead of just come back and forth (LM: Yeah yeah yeah) or even something to make a kind of a sound that wasn't knocked into the patch.
	P21B	But for making maybe it's I think it's just better to be able to hear what's happening [all the interfaces] or maybe there's an option to turn it on and off I don't know.
	P21B	More instruments maybe.
	P21B	Yeah like I think if there were things that kind of like another thing band or notes, something from synthesizers, it would be easier to illustrate the animations...
	P22A	I would say maybe for in the displays instead of the images, maybe a sound can be implemented to help people because sometimes maybe without the knowledge of music or in general.
	P22A	Yeah different graphics and maybe adding even a sound that types you.
	P22A	...Displays, instead of only images you can also hear sounds that guide you in creating your sound later, maybe not a music, maybe I don't know, rain for example that helps you to...
	P22A	Or maybe instead of dots or pure image like landscape or other image, that's my only suggestion for the rest is a very good year.
	P26A	You can introduce Mike to your system.
	P26B	...in three second or something like that just pudding if you want
	P26A	Introduce some colour code or you can have some buttons at the top, controllers at the top or.
	P26B	If you could write the instrument name at the top or somewhere like what we playing, we can't escape, we can't remember the colour code for every instrument each time so if you just write noting else...please sign the backside of the consent form your mind two five notes yeah yeah that's it for you no meaning no problem we should best chocolate for you we should pass tough luck for you.
Other(23)	P2A&B	P2A: OK, do you feel the differences of videos helps, like help you to generate some new thoughts or new creativity. P2B: Yeah, it was like fun being able to like try to match music with videos. P2A: I think as long as you're not trying to focus too much in it and just looking at and then going off instead of thinking too much I think it helps you the tempo or like generate idea, that's how it helped.
	P2A&B	Maybe more colours though more colours with the video?
	P12A	I liked the fact that there were different environments and different pieces of max-exclusive images[the animations], I thought the images had an impact on what we could produced together, or work alone, but I liked the fact that there were different environments to try out, yeah.

P16B	How I enjoyed it and how I got the technicalities were more to, so not about is taking technicalities, were more due to circumstances of the animation or.
P3B	(LM: How do you think about the collaboration experience?) I think it was good.
P7A	I don't have patience to wait I just walked into his personal space and listened to it, I came back to modify mine, so it was easy personal space what is personal space.
P7B	We didn't know each other, we never done music before.
P11B	It was the, the first time we used virtual reality headsets.
P11B	The difference was the video was different in each.
P14B	Maybe, I am not perfectly sure (LM: Because you haven't tried that, right?) Yeah, no, we haven't tried that.
P14A&B	P14B: Last session we did personal space first then did collaborative work. P14A: Yes, so the order switched.
P14B	Yes, and when she had to fix her music and when I had fixed my music, I knew exactly what to fix.
P14A&B	P14B: So whenever I used to be like oh let me check up on her so whenever she wants to check out with mine. P14A: We [then] kind of like clashing.
P16B	Yeah also, sometimes we decided who was caring about which instrument (LM: I see), that sort of separation.
P16A	Because you could not only decide to isolate yourself but also if the other person is isolating because he wants to explore something else and you still want to listen to what he is doing you can.
P16A	Think also the fact that we use the space in a nice static way also depends on our background, because we are both from a musical and experimental music background but someone else that doesn't have the type of.
P16B	It's more that we were forced to use the isolation thing, I think a little bit more.
P17B	Um, I would have liked to have collaborated.
P17A	Potentially yes in this scenario, yeah.
P22B	I think where we messed up in session four is that we were too close, and we didn't make use of the, the wider spaces.
P24B	I didn't notice huge amount of difference, apart from the personal space, where we work. a little bit more in the last session.
P25B	You feel like you have personal space do your own stuff.
P26A	I designed the layout in 2D and then makes 3D models so I do have several schemes this is the last one I chose.

Table C.1: Thematic Analysis Results of Study III

Appendix D

Implementation Prototypes

Before starting the implementation of LeMo, several prototypes were made to explore the technical features of the final prototype, solving the key development issues.

D.1 Prototype I - Getting the Network of Unity Run

Prototype I, which is called Shooting, helped me test and understand the underlying logic of Unity HLAPI. It is a very simple game, two capsule-like players could see each other and walk around and shoot, see Figure D.1. Through building this prototype and a further prototype (shown in Figure D.2), the basic concept of the Unity HLAPI, the setup of Network Manager, the synchronisation and sending messages between clients and server were mastered.

D.2 Prototype II - A Multiplayer Pixel Chess Board

Basically, a step sequencer is an integration of a matrix of cubes and a play line. Prototype II was built to test how to store and sync the condition of each cubes, see Figure D.3. The condition of each cube could be switched to bright or to shadow/black by clicking the small button. Note the two bigger rectangular buttons have different functions, one is to activate all the buttons, and one is to deactivate them. Besides, the conditions of all the buttons are kept synchronised among the server and the clients. Through this prototype, the knowledge of synchronisation mass data via using HLAPI was obtained and practised.

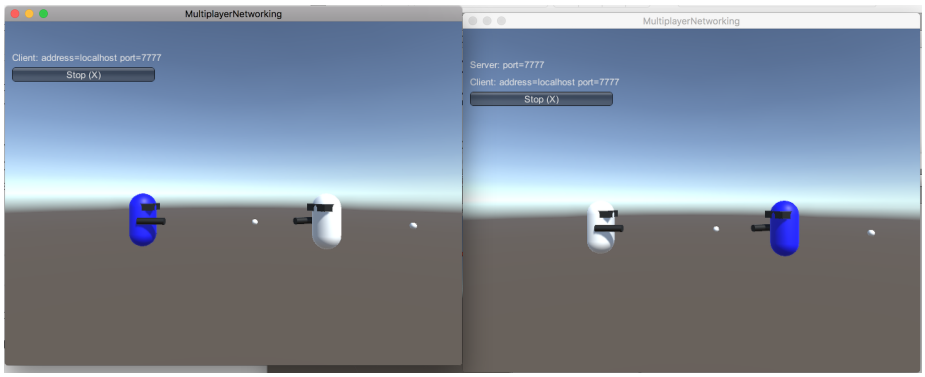


Figure D.1: Screenshots of prototype 1 Shooting (The window on the right side plays the role of server and the left side is the client. The local player is in blue, and the remote player is in white.

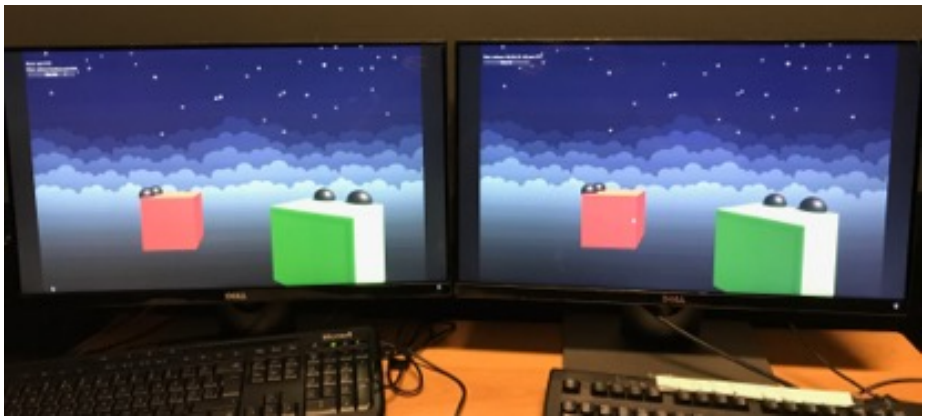


Figure D.2: A prototype built to test communications between two PCs via LAN.

D.3 Prototype III - A Non-Immersive Music Step Sequencer

This prototype mainly aimed at creating and testing the loop mechanism of the step sequencer, in other words, the mechanism of how to check the conditions in each column and play the corresponding notes. See the prototype in Figure D.4, each activated button are shown in black colour, and the column the play line is checking is shown in a white spot light.

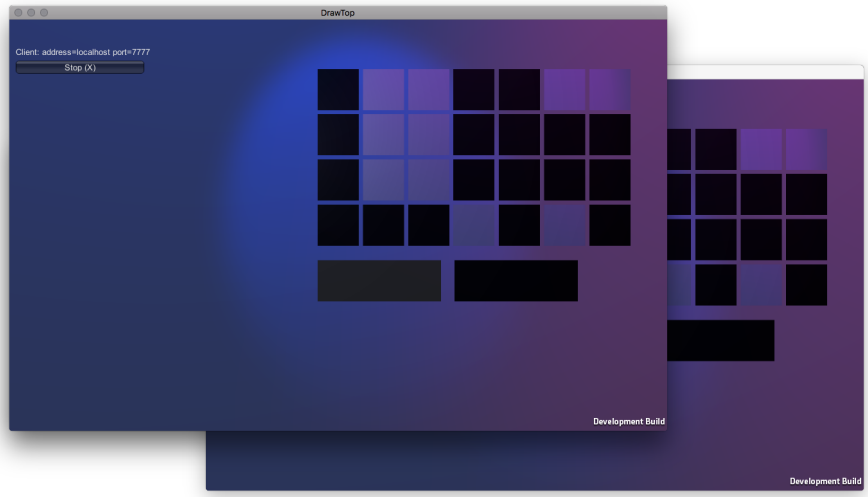


Figure D.3: A 2D multiplayer pixel chess board.

D.4 Prototype IV - Integration with Leap Motion

To the interact with the step sequencer by using bare hands, Leap Motion was integrated. Prototype IV was built to test achieve this, see Figure D.5. In this prototype, the hands' movement was tracked by Leap Motion and the data was then fed to the system, the system checked the collision between the buttons and certain fingers and then switch the button conditions. Gesture detection was also integrated into this prototype, buttons' position could be moved by pinching and dragging.

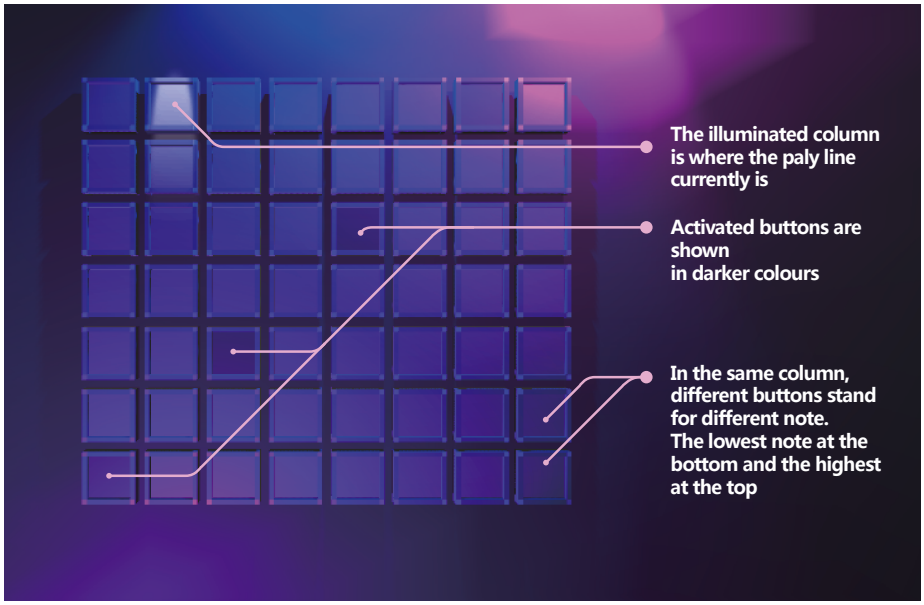


Figure D.4: A mouse-click-based step sequencer

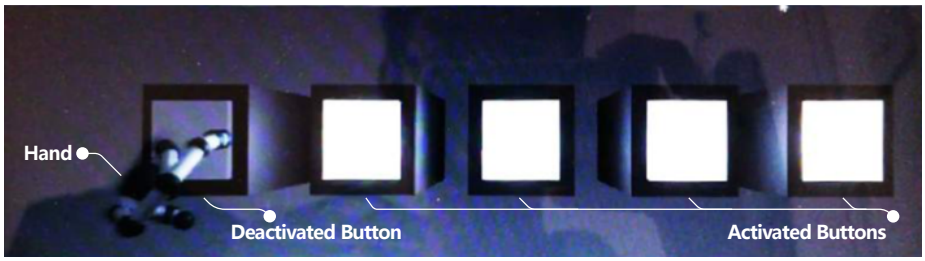


Figure D.5: A Mouse-click-based step sequencer.

Appendix E

Addressing Ethical Concerns in VR Experiments

It is the researchers' responsibility and morality to protect those involved in the experiment, protecting their well-being and avoiding causing harm. Due to the involvement of human participants, the ethical issues in this research do exist and must be well addressed. We believe the ethical issues in these research mainly contain three parts: general issue, data protection and VR-related issues.

E.1 General Issues

Before each of the experiments, forms that explain the experiments and applications have been made, seeking the ethic approval from through Queen Mary Ethics of Research Committee. In each of the experiments, to make sure all the participants take part in the experiment voluntarily, know the experiments procedures (how many steps, where will the experiments happen and what they will do), understand the potential risks (data-related and VR-related ones), the Instructions for Participants Form will be made, explaining the experiment procedures to the participants, and helping the participants fully realise their rights (e.g. withdraw the experiment at any time with no consequences). Only when the participant agrees to continue and then signs a consent form, will the experiment start.

E.2 Data Protection

According to the 1998 U.K. Data Protection Act Principles (Act, 1998), data must be: i) fairly and lawfully processed; ii) processed for limited purposes; iii) Adequate, relevant and not excessive; Accurate; iv) Not kept for longer than is necessary; Processed in line with data subjects' rights; Secure; v) not transferred to countries outside the European Economic Area unless there are adequate safeguards. In this research, data related to individual participants contains background information (age and experience of using VR), video recording of the experiments process, and the audio recording of the interview. For all of these, the participants have the right to know and are free to reject. Thus, before starting the experiments, the participants will be fully told by the experimenter what data will be recorded, how it will be recorded, where it will be stored, how it will be analysed and how long it will be kept. All the data will be processed anonymously to minimise the potential risks. Only when the participants accept the recording, will the experiment continue to be operated. The data recorded will only be used for this study and for research aim.

E.3 VR-Related Concern

Compared with traditional paradigms in experiment psychology, VR poses a novel risk (Madary & Metzinger, 2016). This is mainly due to the fact that the VR could build a higher-level immersion and thus fake the sense of presence. What happened virtually may have the same effect with what happened in reality if people could hardly distinguish between the two. In the obedience experiments carried out by (Milgram, 1974), subjects administered orders that they believed to cause serious injury and pain. Slater et al. (2006) reproduced this famous Milgram obedience experiment in VR, and found the subject reacted similarly even they knew it was not real. "...VR can create a situation in which the user's entire environment is determined by the creators of the virtual world... [This] introduces opportunities for new and especially powerful forms of both mental and behavioural manipulation, especially when commercial, political, religious, or governmental interests are behind the creation and maintenance of the virtual worlds"(Madary & Metzinger, 2016). Despite the direct effect of high level immersion, the illusions of embodiment in VR can have a lasting effect even after users have left the VE, and a handful of recent experiments have shown this. Hershfield et al. (2011) found that subjects tended to allocate more money for retirement after embodying avatars that look like aged themselves . After leaving the VE, subjects who had experienced flying like a superman, acted more helpfully than those who experienced in a helicopter (Rosenberg

et al., 2013). In VR field, four different kinds of risks were concluded and six recommendations were given by Madary & Metzinger (2016). Some of these recommendations are relevant to this study and will be taken into account to minimise the risks. However, it must be noted that obeying recommendations does not mean the risks could be completely eliminated. Thus, it is wise and essential to clearly notify the participants all the potential risks (e.g. the lasting effect) in the experiment instruction form. Also as mentioned above, similar to Magi Barrier Tape (Cirio et al., 2009), the play stage was decorated in different materials, showing the safe roaming areas.

Appendix F

Other Related Materials

F.1 A Reanimation System to Reanimate the System-Logged Data

To ease the difficulty of interpreting the system-logged data, a reanimation (programmed in Processing, graphics designed and made in Adobe Illustrator) was built, see Figure F.1. The system shows the locations, directions of participants and music interfaces. Buttons are available to switch groups and conditions. Time-line at the bottom and play/pause button are available to set time.

F.2 Calculating the Size of Territory

We used a computer vision way to calculate the size of the territory, specifically, we filter each pixel of the image, decide which territory that pixel belongs to (green/red for personal territory and grey for group territory, white for empty), then we count the number of each type of pixels and convert the counts to size measured in square meter, see Figure F.2.

F.3 Related Videos

Links to related videos are listed below.

- An introduction video of the Study I: `https://goo.gl/W6a6jk`.
- A short video clip of the Study II: `https://youtu.be/nk781TFleZI`
- A short video clip of the LeMo II: `https://goo.gl/n9ZhPf`

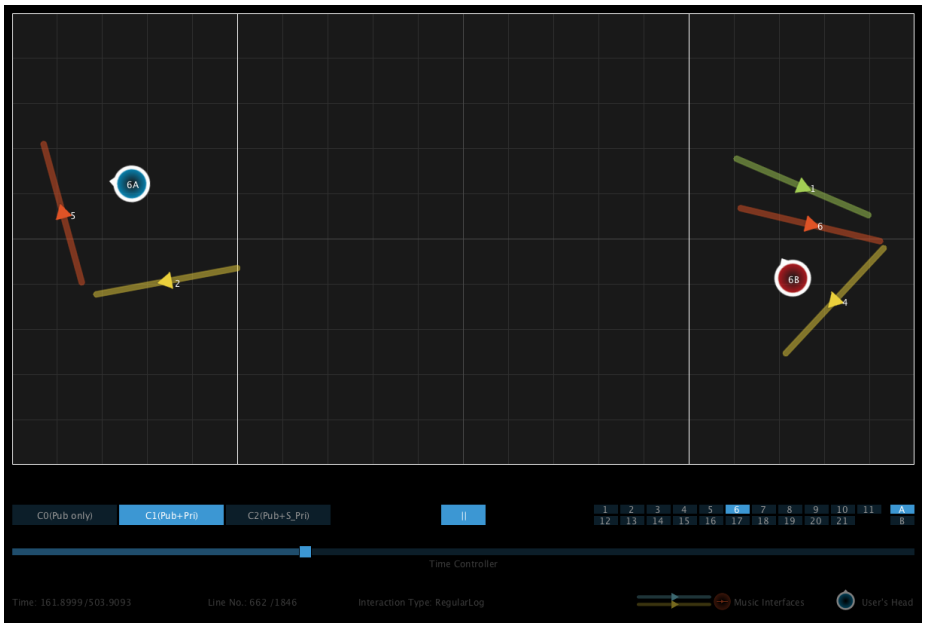


Figure F.1: The reanimation system for system-logged data.

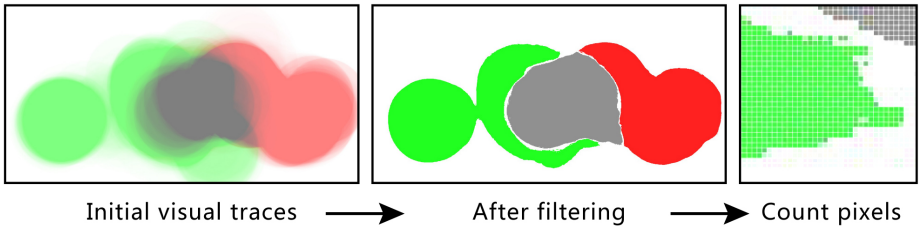


Figure F.2: Calculating the size of territory.

References

Altman, I. (1975). *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding.* ERIC.

Arthur, K. W., & Brooks Jr, F. P. (2000). *Effects of Field of View on Performance with Head-Mounted Displays* (PhD dissertation). University of North Carolina at Chapel Hill.

Bach, C., & Scapin, D. L. (2004). Obstacles and perspectives for evaluating mixed reality systems usability. In *Acte du Workshop MIXER, IUI-CADUI* (Vol. 4).

Baecker, R. M. (1993). *Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration.* Morgan Kaufmann.

Bailenson, J. N., Beall, A. C., & Blascovich, J. (2002). Gaze and task performance in shared virtual environments. *The Journal of Visualization and Computer Animation*, *13*(5), 313–320.

Bailenson, J. N., & Yee, N. (2006). A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *15*(6), 699–716.

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. *Collaborative Learning: Cognitive and Computational Approaches*, *31*, 63.

Ball, R., & North, C. (2005). Effects of tiled high-resolution display on basic visualization and navigation tasks. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems* (pp. 1196–1199).

Barbosa, Á. (2003). Displaced soundscapes: a survey of network systems for music and sonic art creation. *Leonardo Music Journal*, *13*, 53–59.

Barfield, W., Zeltzer, D., Sheridan, T., & Slater, M. (1995). Presence and performance within virtual environments. *Virtual Environments and Advanced Interface Design*, 473–513.

Bavelas, J. B., & Chovil, N. (2006). Nonverbal and verbal communication: Hand gestures and facial displays as part of language use in face-to-face dialogue. In *The SAGE Handbook of Nonverbal Communication* (pp. 97–115). Sage Publications, Inc.

Beebe, S. A., Beebe, S. J., Redmond, M. V., et al. (2000). *Interpersonal Communication*. Scarborough, Ont.: Prentice-Hall Canada.

Begault, D., & Trejo, L. (2000). *3D Sound for Virtual Reality and Multimedia*. Academic Press Professional: Cambridge, MA, USA.

Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for" mini-c" creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *1*(2), 73.

Benford, S., Bowers, J., Fahlén, L. E., Greenhalgh, C., & Snowdon, D. (1995). User embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 242–249).

Benford, S., Bowers, J., Fahlen, L. E., Mariani, J., & Rodden, T. (1994). Supporting cooperative work in virtual environments. *The Computer Journal*, *37*(8), 653–668.

Benford, S., Greenhalgh, C., Rodden, T., & Pycock, J. (2001). Collaborative virtual environments. *Communications of the ACM*, *44*(7), 79–85.

Berthaut, F., Desainte-Catherine, M., & Hachet, M. (2010). Drile: an immersive environment for hierarchical live-looping. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME)* (pp. 192–197).

Billinghurst, M., & Kato, H. (2002). Collaborative augmented reality. *Communications of the ACM*, *45*(7), 64–70.

Billinghurst, M., Poupyrev, I., Kato, H., & May, R. (2000, July). Mixing realities in shared space: an augmented reality interface for collaborative computing. In *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)* (Vol. 3, p. 1641-1644 vol.3). doi: 10.1109/ICME.2000.871085

Bin, S. A., Bryan-Kinns, N., & McPherson, A. (2017). Hands where we can see them! Investigating the impact of gesture size on audience perception. In *International Computer Music Conference.*

Bishop, G., & Fuchs, H. (1992). Research directions in virtual environments: report of an nsf invitational workshop, march 23-24, 1992, university of north carolina at chapel hill. *ACM SIGGRAPH Computer Graphics*, *26*(3), 153–177. Retrieved from `http://doi.acm.org/10.1145/142413.142416` doi: 10.1145/142413.142416

Blaine, T., & Fels, S. (2003). Contexts of collaborative musical experiences. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME), pages=129–134.*

Blandford, A. (2013). Semi-structured qualitative studies. In M. Soegaard & R. Dam (Eds.), *The Encyclopedia of Human-Computer Interaction* (2nd ed.). Interaction Design Foundation. Retrieved from `https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/semi-structured-qualitative-studies`

Boden, M. A. (1998). Creativity and artificial intelligence. *Artificial Intelligence*, *103*(1), 347 - 356. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0004370298000551` (Artificial Intelligence 40 years later) doi: https://doi.org/10.1016/S0004-3702(98)00055-1

Boden, M. A. (2003). *The Creative Mind: Myths and Mechanisms.* Routledge.

Bødker, S., Nielsen, C., & Petersen, M. G. (2000). Creativity, cooperation and interactive design. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (pp. 252–261).

Bouvin, N. O., Zellweger, P. T., Grønbæk, K., & Mackinlay, J. D. (2002). Fluid annotations through open hypermedia: Using and extending emerging web standards. In *Proceedings of the 11th International Conference on World Wide Web* (pp. 160–171).

Bowman, D., Kruijff, E., LaViola, J., Poupyrev, I., & Stuerzlinger, W. (2010). 3D user interfaces: Design, implementation, usability. In *Course Notes of the 28th International Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Brent, E., & Thompson, G. A. (1999). Sociology: Modeling social interaction with autonomous agents. *Social Science Computer Review*, *17*(3), 313–322.

Bressan, F., Vets, T., & Leman, M. (2017). A multimodal interactive installation for collaborative music making: From preservation to enhanced user design. In *Proceedings of the European Society for Cognitive Sciences of Music (ESCOM) Conference, Ghent University* (pp. 23–26).

Brown, A., & Dillon, S. (2007). Networked improvisational musical environments: Learning through on-line collaborative music making. *Music Education with Digital Technology*, 96–106.

Bryan-Kinns, N. (2004). 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* (pp. 135–144). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1013115.1013135` doi: 10.1145/1013115.1013135

Bryan-Kinns, N. (2011). Annotating distributed scores for mutual engagement in Daisyphone and beyond. *Leonardo Music Journal*, *21*, 51–55.

Bryan-Kinns, N. (2013). Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, *71*(1), 76–90.

Bryan-Kinns, N., & Hamilton, F. (2012). Identifying mutual engagement. *Behaviour & Information Technology*, *31*(2), 101–125.

Bryan-Kinns, N., Healey, P. G., & Leach, J. (2007). Exploring mutual engagement in creative collaborations. In *Proceedings of the 6th ACM SIGCHI Conference on Creativity & cognition* (pp. 223–232).

Burkhardt, J.-M., Détienne, F., Hébert, A.-M., Perron, L., & Leclercq, P. (2009). An approach to assess the quality of collaboration in technology-mediated design situations. In *European Conference on Cognitive Ergonomics: Designing beyond the Product—Understanding Activity and User Experience in Ubiquitous Environments* (p. 30).

Butz, A., Beshers, C., & Feiner, S. (1998). Of vampire mirrors and privacy lamps: Privacy management in multi-user augmented environments. *Proceedings ACM User Interface Software and Technology (UIST'98)*, 171172.

Cadoz, C., Luciani, A., & Florens, J. L. (1993). CORDIS-ANIMA: a modeling and simulation system for sound and image synthesis: the general formalism. *Computer Music Journal*, *17*(1), 19–29.

Candy, L., & Hori, K. (2003). The digital muse: HCI in support of creativity: creativity and cognition comes of age: towards a new discipline. *Interactions*, *10*(4), 44–54.

Carlsson, C., & Hagsand, O. (1993, Sep.). DIVE A multi-user virtual reality system. In *Proceedings of IEEE Virtual Reality Annual International Symposium* (p. 394-400). doi: 10.1109/VRAIS.1993.380753

Case, D. A., Ploog, B. O., & Fantino, E. (1990). Observing behavior in a computer game. *Journal of the Experimental Analysis of Behavior*, *54*(3), 185-199. Retrieved from `https://onlinelibrary.wiley.com/doi/abs/10.1901/jeab.1990.54-185` doi: 10.1901/jeab.1990.54-185

Churchill, E. F., & Snowdon, D. (1998). Collaborative virtual environments: An introductory review of issues and systems. *Virtual Reality*, *3*(1), 3–15.

Cirio, G., Marchal, M., Regia-Corte, T., & Lécuyer, A. (2009). The magic barrier tape: a novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology* (pp. 155–162).

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In *Perspectives on Socially Shared Cognition.* (pp. 127–149). Washington, DC, US: American Psychological Association. doi: https://doi.org/10.1037/10096-006

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*(2), 259 - 294. Retrieved from `http://www.sciencedirect.com/science/article/pii/0364021389900086` doi: https://doi.org/10.1016/0364-0213(89)90008-6

Coates, G. (1992). Program from Invisible Site: a virtual show, a multimedia performance work. *Presented by George Coates Performance Works*.

Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education.* Routledge.

Cook, P. R. (2001). Principles for designing computer music controllers. In *Proceedings of the 2001 Conference on New Interfaces for Musical Expression (NIME)* (pp. 1–4).

Cook, P. R. (2002). *Real Sound Synthesis for Interactive Applications.* AK Peters/CRC Press.

Cook, P. R. (2009). Re-designing principles for computer music controllers: a case study of SqueezeVox Maggie. In *Proceedings of the 2009 Conference on New Interfaces for Musical Expression (NIME)* (Vol. 9, pp. 218–221).

Cruz, A., Paredes, H., Fonseca, B., Morgado, L., & Martins, P. (2014). Can presence improve collaboration in 3D virtual worlds? *Procedia Technology*, *13*, 47–55.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 64–73.

Cugini, J., Damianos, L., Hirschman, L., Kozierok, R., Kurtz, J., Laskowski, S., & Scholtz, J. (1997). Methodology for evaluation of collaboration systems. *The Evaluation Working Group of the DARPA Intelligent Collaboration and Visualization Program, Rev*, *3*.

Damianos, L. E., Hirschman, L., Kozierok, R., Kurtz, J. L., Greenberg, A., Walls, K., ... Scholtz, J. (1999). Evaluation for collaborative systems. *ACM Comput. Surv.*, *31*(2es), 15.

Davis, S., Nesbitt, K., & Nalivaiko, E. (2014). A systematic review of cybersickness. In *Proceedings of the 2014 Conference on Interactive Entertainment* (pp. 1–9).

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: a meta-analysis. *Journal of applied Psychology*, *95*(1), 32.

Détienne, F. (2006). Collaborative design: Managing task interdependencies and multiple perspectives. *Interacting with Computers*, *18*(1), 1–20.

De Vignemont, F. (2007). Habeas corpus: the sense of ownership of one's own body. *Mind & Language*, *22*(4), 427–449.

Dinh, H. Q., Walker, N., Hodges, L. F., Chang Song, & Kobayashi, A. (1999, March). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)* (p. 222-228). doi: 10.1109/VR.1999.756955

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work* (pp. 107–114).

Draper, J. V., Kaber, D. B., & Usher, J. M. (1998). Telepresence. *Human Factors*, *40*(3), 354–375.

Dryer, D. C. (1999). Getting personal with computers: How to design personalities for agents. *Applied Artificial Intelligence*, *13*(3), 273–295.

Ellemers, N., & Rink, F. (2005). Identity in work groups: the beneficial and detrimental consequences of multiple identities and group norms for collaboration and group performance. In *Social Identification in Groups* (pp. 1–41). Emerald Group Publishing Limited.

Ellis, S. R. (1991). Nature and origins of virtual environments: a bibliographical essay. *Computing Systems in Engineering*, *2*(4), 321–347.

Ellis, S. R. (1995). Virtual environments and environmental instruments. *Simulated and virtual realities: Elements of Perception*, 85–101.

Fels, S., Gadd, A., & Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, *7*(2), 109–126.

Fencott, R., & Bryan-Kinns, N. (2010). Hey man, you're invading my personal space! Privacy and awareness in collaborative music. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME)* (pp. 198–203).

Fernandes, A. S., & Feiner, S. K. (2016, March). Combating VR sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)* (p. 201-210). doi: 10.1109/3DUI.2016.7460053

Ford, D. Y., & Harris, J. J. (1992). The elusive definition of creativity. *The Journal of Creative Behavior*.

Frécon, E., Greenhalgh, C., & Stenius, M. (1999). The divebone—an application-level network architecture for internet-based cves. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (pp. 58–65). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/323663.323672` doi: 10.1145/323663.323672

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14–21.

Gaver, W. W. (1992). The affordances of media spaces for collaboration. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work (CSCW)* (pp. 17–24).

Gelineck, S., Böttcher, N., Martinussen, L., & Serafin, S. (2005). Virtual reality instruments capable of changing dimensions in real-time. *Proceedings Enactive*.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Exford, England Houghton Mifflin.

Gibson, J. J. (2014). *The Ecological Approach to Visual Perception* (classic ed.). Psychology Press.

Gil-Gómez, J.-A., Manzano-Hernández, P., Albiol-Pérez, S., Aula-Valero, C., Gil-Gómez, H., & Lozano-Quilis, J.-A. (2017). USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. *Sensors*, *17*(7), 1589.

Graham, J. (1995). *Inter-Professional Collaboration in the Special School* (Unpublished doctoral dissertation). Institute of Education, University of London.

Greenbaum, P. (1992). The lawnmower man. *Film and Video*, *9*(3), 58–62.

Greenberg, S., Boyle, M., & LaBerge, J. (1999). PDAs and shared public displays: Making personal information public, and public information personal. *Personal Technologies*, *3*(1-2), 54–64.

Grudin, J. (1994, May). Computer-supported cooperative work: history and focus. *Computer*, *27*(5), 19-26. doi: 10.1109/2.291294

Gutwin, C., & Greenberg, S. (2004). The importance of awareness for team cognition in distributed collaboration. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the Factors that Drive Process and Performance* (p. 177-201). Washington, DC, US: American Psychological Association.

Hall, E. T. (1966). *The Hidden Dimension* (Vol. 609). Garden City, NY: Doubleday.

Hamalainen, R. (2008). Designing and evaluating collaboration in a virtual game environment for vocational learning. *Computers & Education*, *50*(1), 98–109.

Haywood, N., & Cairns, P. (2006). Engagement with an interactive museum exhibit. In *People and Computers XIX - The Bigger Picture* (pp. 113–129). Springer.

Healey, P. G., Leach, J., & Bryan-Kinns, N. (2005). Inter-play: Understanding group music improvisation as a form of everyday interaction. *Proceedings of Less is MoreSimple Computing in an Age of Complexity*.

Heath, C., Svensson, M. S., Hindmarsh, J., Luff, P., & Vom Lehn, D. (2002). Configuring awareness. *Computer Supported Cooperative Work (CSCW)*, *11*(3-4), 317–347.

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, *1*(2), 262–271.

Hendrix, C., & Barfield, W. (1996a). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators & Virtual Environments*, *5*(3), 274–289.

Hendrix, C., & Barfield, W. (1996b). The sense of presence within auditory virtual environments. *Presence: Teleoperators & Virtual Environments*, *5*(3), 290–301.

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research*, *48*(SPL), 23–37.

Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., & Kassell, N. (1997). Usability Analysis of 3D Rotation Techniques. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (pp. 1–10). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/263407.263408` doi: 10.1145/263407.263408

Holmquist, L. E., Falk, J., & Wigström, J. (1999, Mar 01). Supporting group collaboration with interpersonal awareness devices. *Personal Technologies*, *3*(1), 13–21. Retrieved from `https://doi.org/10.1007/BF01305316` doi: 10.1007/BF01305316

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies*, *64*(2), 79–102.

Hoyt, C. L., & Blascovich, J. (2003). Transformational and transactional leadership in virtual and physical environments. *Small Group Research*, *34*(6), 678–715.

Hunt, C. (2018). *Field of view face-off: Rift vs Vive vs Gear VR vs PSVR.* `https://www.vrheads.com/field-view-faceoff-rift-vs-vive-vs-gear-vr-vs-psvr`. (Accessed: 2018-01-28)

IJsselsteijn, W. A., De Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: concept, determinants, and measurement. In *Human Vision and Electronic Imaging V* (Vol. 3959, pp. 520–530).

Imam, B., Miller, W. C., McLaren, L., Chapman, P., & Finlayson, H. (2013). Feasibility of the Nintendo WiiFit for improving walking in individuals with a lower limb amputation. *SAGE Open Medicine*, *1*, 1–11.

Jackson, R. L., & Fagan, E. (2000). Collaboration and learning within immersive virtual reality. In *Proceedings of the Third International Conference on Collaborative Virtual Environments* (pp. 83–92).

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human factors in computing systems* (pp. 201–210).

Johnston, A., Candy, L., & Edmonds, E. (2008). Designing and evaluating virtual musical instruments: Facilitating conversational user interaction. *Design Studies*, *29*(6), 556–571.

Kalawsky, R. S. (1999). VRUSEa computerised diagnostic tool: for usability evaluation of virtual/synthetic environment systems. *Applied Ergonomics*, *30*(1), 11–25.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*(1), 1–12.

Kaufman, J. C., & Sternberg, R. J. (2010). *The Cambridge Handbook of Creativity (Eds.)*. Cambridge University Press. doi: https://doi.org/10.1017/CBO9780511763205

Kizony, R., Katz, N., Rand, D., & Weiss, P. L. T. (2006). Short Feedback Questionnaire (SFQ) to enhance client-centered participation in virtual environments. In *Cyberpsychology & Behavior* (Vol. 9, pp. 687–688).

Kizony, R., Katz, N., & Weiss, P. L. (2003). Adapting an immersive virtual reality system for rehabilitation. *The Journal of Visualization and Computer Animation*, *14*(5), 261–268.

Kizony, R., Raz, L., Katz, N., Weingarden, H., & Weiss, P. (2003). Using a video projected VR system for patients with spinal cord injury. In *Proceedings of the Second International Workshop on Virtual Rehab* (pp. 82–88).

Kohlrausch, A., & van de Par, S. (1999). Auditory-visual interaction: from fundamental research in cognitive psychology to (possible) applications. In *Human Vision and Electronic Imaging IV* (Vol. 3644, pp. 34–44).

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior*, *19*(3), 335–353.

Krueger, M. W. (1991). *Artificial Realities II.* Addison-Wesley Publishing Co., Reading, Massachusetts.

Kruger, R., Carpendale, S., Scott, S. D., & Greenberg, S. (2004). Roles of orientation in tabletop collaboration: Comprehension, coordination and communication. *Computer Supported Cooperative Work (CSCW)*, *13*(5-6), 501–537.

Lanier, J. (1992). Virtual reality: the promise of the future. *Interactive Learning International*, *8*(4), 275–79.

LaViola Jr, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin*, *32*(1), 47–56.

Lea, R., Honda, Y., & Matsuda, K. (1997). Virtual society: Collaboration in 3D spaces on the Internet. *Computer Supported Cooperative Work (CSCW)*, *6*(2), 227–250.

Loomis, J. M. (1992). Distal attribution and presence. *Presence: Teleoperators & Virtual Environments*, *1*(1), 113–119.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 557–564.

Lubart, T. (2005). How can computers be partners in the creative process: classification and commentary on the special issue. *International Journal of Human-Computer Studies*, *63*(4-5), 365–369.

Luciani, A. (2007). Virtual reality and virtual environment. In *Enaction and Enactive interfaces : a Handbook of Terms* (p. 299-300). Enactive Systems Book. Retrieved from `https://hal.archives-ouvertes.fr/hal-00980481`

Madary, M., & Metzinger, T. K. (2016). Real virtuality: a code of ethical conduct. recommendations for good scientific practice and the consumers of vr-technology. *Frontiers in Robotics and AI*, *3*, 3.

Mäki-Patola, T., Laitinen, J., Kanerva, A., & Takala, T. (2005). Experiments with virtual reality instruments. In *Proceedings of the 2005 Conference on New Interfaces for Musical Expression (NIME)* (pp. 11–16).

Men, L., & Bryan-Kinns, N. (2018, March). LeMo: Supporting collaborative music making in virtual reality. In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)* (p. 1-6). doi: 10.1109/SIVE.2018.8577094

Men, L., Bryan-Kinns, N., Hassard, A. S., & Ma, Z. (2017, March). The impact of transitions on user experience in virtual reality. In *2017 IEEE Virtual Reality (VR)* (p. 285-286). doi: 10.1109/VR.2017.7892288

Men, L., & Bryan-Kinns, N. (2019). LeMo: Exploring virtual space for collaborative creativity. In *Proceedings of the 2019 on Creativity and Cognition* (pp. 71–82). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/3325480.3325495` doi: 10.1145/3325480.3325495

Men, L., Bryan-Kinns, N., & Bryce, L. (2019, November). Designing spaces to support collaborative creativity in shared virtual environments. *PeerJ Computer Science*, *5*, e229. Retrieved from `https://doi.org/10.7717/peerj-cs.229` doi: 10.7717/peerj-cs.229

Menon, J. (1997). Collaborative visualization and modeling. In *Proceedings of 1997 International Conference on Shape Modeling and Applications* (pp. 178–187).

Milgram, S. (1974). Compliant subjects.(book reviews: Obedience to authority. an experimental view). *Science*, *184*, 667–669.

Minsky, M. (1980, June). Telepresence. *Omni*, *2*, 45-51.

Morreale, F., & McPherson, A. (2017). Design for longevity: Ongoing use of instruments from NIME 2010-14. In *Proceedings of the 2017 Conference on New Interfaces for Musical Expression (NIME).*

Nabavian, S., & Bryan-Kinns, N. (2006). Analysing group creativity: a distributed cognitive study of joint music composition. *Proc. of Cognitive Science*, 1856–1861.

Naef, M., Staadt, O., & Gross, M. (2002). Spatialized audio rendering for immersive virtual environments. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (pp. 65–72). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/585740.585752` doi: 10.1145/585740.585752

Nassiri, N., Powell, N., & Moore, D. (2010). Human interactions and personal space in collaborative virtual environments. *Virtual reality*, *14*(4), 229–240.

National Advisory Committee on Creative and Cultural Education (NACCCE). (1999). *All our futures: Creativity, culture and education.* London: DfEE/DCMS.

Nedel, L., de Souza, V. C., Menin, A., Sebben, L., Oliveira, J., Faria, F., & Maciel, A. (2016). Using immersive virtual reality to reduce work accidents in developing countries. *IEEE Computer Graphics and Applications*, *36*(2), 36–46.

Nguyen, T. T. H., & Duval, T. (2014, March). A survey of communication and awareness in collaborative virtual environments. In *2014 International Workshop on Collaborative Virtual Environments (3DCVE)* (p. 1-8). doi: 10.1109/3DCVE.2014.7160928

Noguera, M., Hurtado, M. V., Rodríguez, M. L., Chung, L., & Garrido, J. L. (2010). Ontology-driven analysis of UML-based collaborative processes using OWL-DL and CPN. *Science of Computer Programming*, *75*(8), 726–760.

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *12*(5), 481–494.

Pausch, R., Proffitt, D., & Williams, G. (1997). Quantifying immersion in virtual reality. In *Computer Graphics (SIGGRAPH) Annual Conference Series* (pp. 13–18). CUMINCAD.

Phalip, J., Edmonds, E. A., & Jean, D. (2009). Supporting remote creative collaboration in film scoring. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition* (pp. 211–220).

Pinelle, D., Gutwin, C., & Greenberg, S. (2003). Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *10*(4), 281–311.

Plucker, J. A., & Beghetto, R. A. (2003). Why not be creative when we enhance creativity. *Rethinking Gifted Education*, 215–226.

Purcell, A. (1998). Drawings and the design process. *Design Studies*, *19*(4), 389–430.

Rabin, S. (2005). Introduction to game development (game development). *Inc., Rockland, MA*.

Raffestin, C. (2012). Space, territory, and territoriality. *Environment and Planning D: Society and Space*, *30*(1), 121–141.

Ramos, G., & Balakrishnan, R. (2003). Fluid interaction techniques for the control and annotation of digital video. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (pp. 105–114).

Ray, P. (2002). Computer supported cooperative work (CSCW). *Cooperative Management of Enterprise Networks*, 27–46.

Reason, J. T., & Brand, J. J. (1975). *Motion Sickness.* Academic press.

Rice, R. E. (1993). Media appropriateness: Using social presence theory to compare traditional and new organizational media. *Human Communication Research*, *19*(4), 451–484.

Rodden, T. (1991). A survey of CSCW systems. *Interacting with Computers*, *3*(3), 319–353.

Romano, D. M., Brna, P., & Self, J. A. (1998). Collaborative decision-making and presence in shared dynamic virtual environments. In *Presence in Shared Virtual Evironments Workshop.*

Rosenberg, R. S., Baughman, S. L., & Bailenson, J. N. (2013). Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PloS One*, *8*(1), e55003.

Roussos, M., Johnson, A. E., Leigh, J., Vasilakis, C. A., Barnes, C. R., & Moher, T. G. (1997). NICE: combining constructionism, narrative and collaboration in a virtual learning environment. *Computer Graphics-New York-Association for Computing Machinery*, *31*, 62–63.

Rovai, A. P. (2002). Building sense of community at a distance. *The International Review of Research in Open and Distributed Learning*, *3*(1).

Sack, R. D. (1983). Human territoriality: a theory. *Annals of the Association of American Geographers*, *73*(1), 55–74.

Sarmiento, W. J., & Collazos, C. A. (2012, Jan). CSCW systems in virtual environments: a general development framework. In *2012 10th International Conference on Creating, Connecting and Collaborating through Computing* (p. 15-22). doi: 10.1109/C5.2012.17

Sawyer, R. K. (2011). *Explaining Creativity: the Science of Human Innovation.* Oxford university press.

Schroeder, R. (2012). *The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments.* Springer Science & Business Media.

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). Embodied presence in virtual environments. In *Visual Representations and Interpretations* (pp. 269–278). Springer.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, *10*(3), 266–281.

Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. (2001). Research on presence in virtual reality: a survey. *CyberPsychology & Behavior*, *4*(2), 183–201.

Scott, S. D., Carpendale, M. S. T., & Inkpen, K. M. (2004). Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer-Supported Cooperative Work (CSCW)* (pp. 294–303).

Scott, S. D., & Carpendale, S. (2010). Theory of tabletop territoriality. In *Tabletops-Horizontal Interactive Displays* (pp. 357–385). Springer.

Serafin, S., Erkut, C., Kojs, J., Nilsson, N. C., & Nordahl, R. (2016). Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal*, *40*(3), 22–40.

Serim, B., & Jacucci, G. (2019). Explicating "implicit interaction": an examination of the concept and challenges for research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (p. 417).

Shah, J. J., Vargas-Hernandez, N., Summers, J. D., & Kulkarni, S. (2001). Collaborative sketching (C-Sketch)an idea generation technique for engineering design. *The Journal of Creative Behavior*, *35*(3), 168–198.

Shen, C., Everitt, K., & Ryall, K. (2003). UbiTable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *International Conference on Ubiquitous Computing* (pp. 281–288).

Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments*, *1*(1), 120–126.

Sherman, W. R., & Craig, A. B. (2018). *Understanding Virtual reality: Interface, Application, and Design*. Morgan Kaufmann.

Shiu, E. (2014). *Creativity Research: An Inter-disciplinary and Multi-disciplinary Research Handbook*. Routledge.

Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications. *The Social Psychology of Telecommunications*.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557.

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., ... Sanchez-Vives, M. V. (2006). A virtual reprise of the stanley milgram obedience experiments. *PloS One*, *1*(1), e39.

Slater, M., Lotto, B., Arnold, M. M., & Sánchez-Vives, M. V. (2009). How we experience immersive virtual environments: the concept of presence and its measurement. *Anuario de Psicología, 2009, vol. 40, p. 193-210*.

Slater, M., & Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators & Virtual Environments*, *9*(5), 413–434.

Slater, M., Usoh, M., & Chrysanthou, Y. (1995). The influence of dynamic shadows on presence in immersive virtual environments. In *Virtual Environments* (pp. 8–21). Springer. doi: 10.1007/978-3-7091-9433-1_2

Slater, M., Usoh, M., & Steed, A. (1994). Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *3*(2), 130–144.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *6*(6), 603–616.

Smith, J. O. (2006). *Physical Audio Signal Processing: for Virtual Musical Instruments and Digital Audio Effects*. Stanford, CA: W3K Publishing.

Sternberg, R. J. (1988). *The Nature of Creativity: Contemporary Psychological Perspectives*. CUP Archive.

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. *Handbook of Creativity*, *1*, 3–15.

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication*, *42*(4), 73–93.

Sugimoto, M., Hosoi, K., & Hashizume, H. (2004). Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 41–48).

Summers, C., Lympouridis, V., & Erkut, C. (2015). Sonic interaction design for virtual and augmented reality environments. In *2015 IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)* (pp. 1–6).

Tang, J. C. (1991). Findings from observational studies of collaborative work. *International Journal of Man-Machine Studies*, *34*(2), 143–160.

Taylor, C. W. (1988). Various approaches to and definitions of creativity. *The Nature of Creativity*, 99–121.

Taylor, R. B. (1988). *Human Territorial Functioning: an Empirical, Evolutionary Perspective on Individual and Small Sroup Territorial Cognitions, Behaviors, and Consequences* (No. 8). Cambridge University Press.

Thiebaut, J.-B., Healey, P. G., & Bryan-Kinns, N. (2008). Drawing electroacoustic music. In *ICMC*.

Thornhill-Miller, B., & Dupont, J.-M. (2016). Virtual reality and the enhancement of creativity and innovation: Under recognized potential among converging technologies? *Journal of Cognitive Education and Psychology*, *15*(1), 102–121.

Titon, J. T., & Slobin, M. (1996). The music-culture as a world of music. *Worlds of Music: an Introduction to the Music of the World's Peoples. New York: Schirmer Books*.

Towell, J., & Towell, E. (1997). Presence in text-based networked virtual environments or "MUDS". *Presence: Teleoperators & Virtual Environments*, *6*(5), 590–595.

Tromp, J. G., Steed, A., & Wilson, J. R. (2003). Systematic usability evaluation and design issues for collaborative virtual environments. *Presence: Teleoperators & Virtual Environments*, *12*(3), 241–267.

Tsakiris, M., Schütz-Bosbach, S., & Gallagher, S. (2007). On agency and body-ownership: Phenomenological and neurocognitive reflections. *Consciousness and Cognition*, *16*(3), 645–660.

Turchet, L. (2019). Smart musical instruments: Vision, design principles, and future directions. *IEEE Access*, *7*, 8944-8963. doi: 10.1109/ACCESS.2018.2876891

Turchet, L., Fischione, C., Essl, G., Keller, D., & Barthet, M. (2018). Internet of musical things: Vision and challenges. *IEEE Access*, *6*, 61994-62017. doi: 10.1109/ACCESS.2018.2872625

Usoh, M., Alberto, C., & Slater, M. (1996). *Presence: Experiments in the Psychology of Virtual Environments.* Retrieved from `http://www8.cs.umu.se/kurser/TDBD12/VT06/articles/precense-paper-teap_full96.pdf` (Accessed: 2018-1-12)

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence: Teleoperators & Virtual Environments*, *9*(5), 497–503.

Välimäki, V., & Takala, T. (1996). Virtual musical instrumentsnatural sound using physical models. *Organised Sound*, *1*(2), 75–86.

Wallace, P., & Maryott, J. (2009). The impact of avatar self-representation on collaboration in virtual worlds. *Innovate: Journal of Online Education*, *5*(5), 3.

Wang, G. (2009). Designing Smule's Ocarina: the iPhone's magic flute. In *Proceedings of the 2009 Conference on New Interfaces for Musical Expression (NIME)* (pp. 303–307).

Wang, G. (2014). Principles of visual design for computer music. In *ICMC*.

Ware, C., & Jessome, D. R. (1988). Using the bat: a six-dimensional mouse for object placement. *IEEE Computer Graphics and Applications*, *8*(6), 65–70.

Warren, S. D., & Brandeis, L. D. (1890). Right to privacy. *Harv. L. Rev.*, *4*, 193.

Watts, L., Monk, A., & Daly-Jones, O. (1996). Inter-personal awareness and synchronization: assessing the value of communication technologies. *International Journal of Human-Computer Studies*, *44*(6), 849 - 873. Retrieved from `http://www.sciencedirect.com/science/article/pii/S1071581996900361` doi: https://doi.org/10.1006/ijhc.1996.0036

Weakley, A., Deverell, K., & Yuille, J. (2007). WEB 2.0 in support of sketching in architectural practice. In *Proceedings of the First International Conference on Semantic Web and Web 2.0 in Architectural, Product and Engineering Design-Volume 294* (pp. 57–62).

Weisband, S. (2002). Maintaining awareness in distributed team collaboration: Implications for leadership and performance. In P. Hinds & S. Kiesler (Eds.), *Distributed Work.* (pp. 311–333). Cambridge, MA, US: MIT Press.

Wessel, D., & Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Computer Music Journal*, *26*(3), 11–22.

Wexelblat, A. (2014). *Virtual Reality: Applications and Explorations.* Acadeic Press.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence*, *7*(3), 225–240.

Wozniewski, M., Bouillot, N., Settel, Z., & Cooperstock, J. R. (2008). Large-scale mobile audio environments for collaborative musical interaction. In *Proceedings of the 2008 Conference on New Interfaces for Musical Expression (NIME)* (pp. 13–18).

Xambó, A., Hornecker, E., Marshall, P., Jordà, S., Dobbyn, C., & Laney, R. (2013). Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *20*(6), 36.

Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., & Merget, D. (2007). The unbearable likeness of being digital: the persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior*, *10*(1), 115–121.

Yin, R. K. (2017). *Case Study Research and Applications: Design and Methods.* Sage Publications.

Zhang, X., & Furnas, G. W. (2003). The effectiveness of multiscale collaboration in virtual environments. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems* (pp. 790–791).