Designing Privacy for Collaborative Music Making in Virtual Reality

2 authors, including:

Liang Men

Liverpool John Moores University; Queen Mary, University of London

10 PUBLICATIONS   **101** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Collaborative Virtual Environment View project

Virtual Reality View project

Designing Privacy for Collaborative Music Making in Virtual Reality

Liang Men
Liverpool John Moores University
Liverpool, United Kingdom
l.men@ljmu.ac.uk

Danqi Zhao
Independent Researcher
Liverpool, United Kingdom
dq.zhao@hotmail.com

ABSTRACT

Privacy is needed to schedule personal activities and social interaction, and to prepare and create wonderful things before confidently sharing with others. How to build privacy in virtual reality remains an open question, though it is potentially as important as it is in reality. This paper presents two studies to build privacy in VR with a focus on providing sonic privacy to aid collaborative music making (CMM). The first study uses soundproof walls to provide sonic privacy, the second study applies augmented sound attenuation as a way to produce sonic privacy to aid collaborative music making. Results show privacy is needed and both ways can produce sonic privacy. Privacy provided in the first study is binary and discrete, whilst privacy provided in the latter is continuous, step-less and adjustable. By comparing and discussing between these two approaches, we propose a methodological approach to build privacy to aid collaboration in shared virtual reality. In particular, three channels for building privacy have been identified and three implications have been proposed for designing privacy for collaborative tasks in shared virtual environments, collaborative music making in particular.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Virtual reality**; **Collaborative content creation**; • **Applied computing** → **Sound and music computing**.

KEYWORDS

Sonic Privacy, Virtual Reality, Sonic Interaction, Virtual Environment, Audio Attenuation, Virtual Reality Musical Instruments, HCI, Interaction Design, Music Computing

---

1 INTRODUCTION

Everyone needs privacy, it helps us to manage both personal activities and social interactions [24]. Privacy not only means protecting private information, but more importantly, it means secluding ourselves to create and prepare things for sharing with others [10]. Privacy is important to effective group functioning [1, 2, 31, 33]. Previously, privacy has been defined as: 'the selective control of access to the self or to one's group' [1]. Privacy can be gained in multiple ways, usually not by removing oneself from others' presence, but by controlling the amount of contact with others, and as such privacy can be regarded as a boundary control process in which the individual regulates with who, when, what type of and how much the interaction will be [22]. This boundary control involves both restricting and seeking interaction to achieve the desired degree of access to the self (or one's group) by others at a particular time and in a given set of circumstances. Both personal, social, or physical factors can contribute to the boundary control [15]. Physical environmental factors entail features such as barriers, location, layout, and distance [22].

These works present us with an overall understanding of privacy, its important role in collaboration, and potential ways of building privacy. However, none of these works considers privacy in a Virtual Reality setting, where people may have similar behavioural patterns and needs of privacy. Given the intersection between music and Virtual Reality has grown significantly [17–19, 29, 30], we are particularly interested in how to build sonic privacy in VR and if it plays a supporting role. In this paper, we explore how to construct privacy, sonic privacy in particular, discuss the ways to balance privacy and openness, and propose implications to inform privacy design for shared virtual environments (SVEs). In particular, we provide a methodological approach to design privacy feature of shared virtual objects, based on three channels: visual privacy, audio privacy and accessibility.

2 RELATED WORK

2.1 Private Space vs Shared Space

Virtual environments (VEs) constitute illusive but meaningful virtual spaces [27]. These virtual spaces provides richness and affordance to inhabit people and their interactivity. "Space", either in reality or virtual reality, 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, cf. [23]. Territory helps people to mediate their social interaction [1], which is argued to be a key element to collaboration [16]. A "private 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.

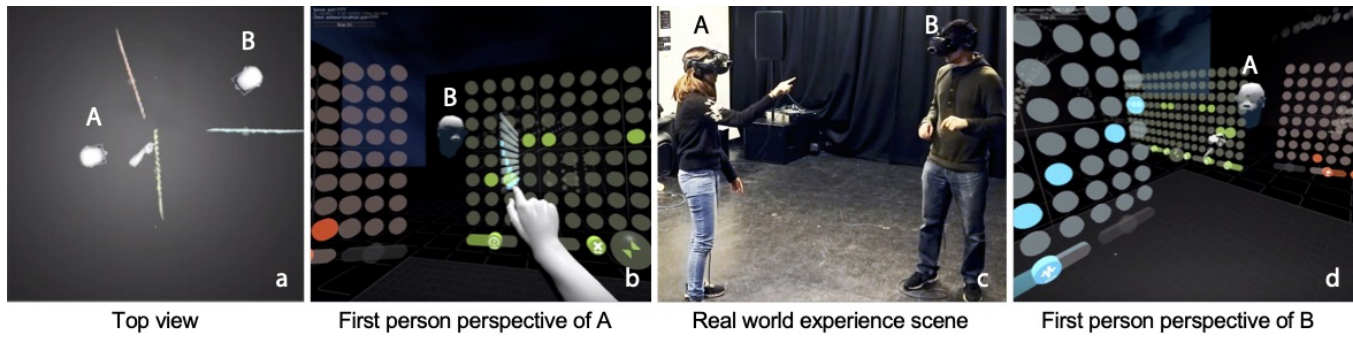| Top view | First person perspective of A | Real world experience scene | First person perspective of B |

Figure 1: Participants creating music together in LeMo.

an experiment). Providing private space is argued to be useful to support collaborative creativity [25]. Is is important to integrate private and group spaces in a way to allow users to work individually in their private spaces at their pace, cooperatively work together in the shared space, and smoothly switch between both of the spaces [11, 28]. As a starting point of this exploration, [11] developed a PDA-based prototype, observed how users shifted between these two spaces and recommended against a rigid notion of "private". Instead, they suggested the boundary between private and public should be provided with gradations in subtle and lightweight ways, supporting a fluid transition between private and public. Following that, [26] 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.

2.2 Privacy and Collaboration

Private workplaces are thought to be an important element of collaboration and have been investigated for a long time, e.g. [4, 7] have investigated providing private space in group text editing. In the field of collaboration, research has explored the provision of personalised audio feedback with multi-touch interfaces [21] and personal information spaces in multi-user tabletop computers [26]. Research identifies that people shift between individual activities and tightly coupled collaboration during collaboration [9, 13], and claims the necessity of remaining informed of collaborators' activities [13]. [10] explores privacy, awareness and roles in collaborative music making, tests varying degrees (3 levels) of privacy, and identifies that participants chose to prepare music contributions before sharing to group. Preservation of different levels of anonymity and privacy in group music making has been explored in the project JamSpace [12]. For example, the highest level of privacy, called private space is provided when a user is discounted to the JamSpace server. In this mode, the user is invisible to other users and can own their own instruments, create, play and record, music. He further defined four types of space depending on the level of privacy, from low to high, they are: private space, personal

space, shared space, and public space, with each showing an increasing level of information and presence to collaborators. Two methods were presented to manipulate the privacy condition of objects in augmented reality [8], if all objects are public by default, a *privacy lamp* or *public lamps* can be positioned over the objects that users want to make it private/public. Another tool is called *vampire mirror*, which 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 may not fit virtual reality, as it breaks the illusion that VR is trying to build – that all the collaborators are present in the same VE and manipulating the same set of virtual objects. i.e. so it is important to build sonic privacy without impacting the visibility of objects. These works are good in informing the design and management of privacy for collaborative music making, however, none of them is exploring the role of privacy in the intersection of collaborative music making and Virtual Reality, which is why we are interested to investigate in this paper.

3 LEMO – A TOOL FOR COLLABORATIVE MUSIC MAKING IN VR

To explore privacy in CMM in SVEs, a musical collaboration system named LeMo[1] was created [20]. LeMo allows two users to make music together in VR. When using LeMo, two users can manipulate virtual music interfaces together to create a 16-beat music loop simultaneously, as shown in Fig.1. LeMo was programmed in Unity using C#, models and textures were made in Cinema 4D and Photoshop respectively. The run-time apparatus includes 2 HTC Vive VR headsets[2], each with a hand tracker (Leap Motion[3]) mounted, see Fig. 1c. The movements of heads, and hand gestures are tracked and then mapped to a virtual embodiment. Two PCs running LeMo are connected and synchronised via a LAN cable.

LeMo is made of three subsystems: (i) Music interface for music creation - LeMo supports two players to generate, remove, position and edit virtual music interfaces. These interfaces have two modes: *sphere* and *matrix* (Figure 2a). Users can generate *spheres* with pinch and stretch gesture. The music interface can be switched

---

[1]LeMo: https://sites.google.com/view/liangmen/projects/LeMo
[2]HTC Vive: https://www.vive.com
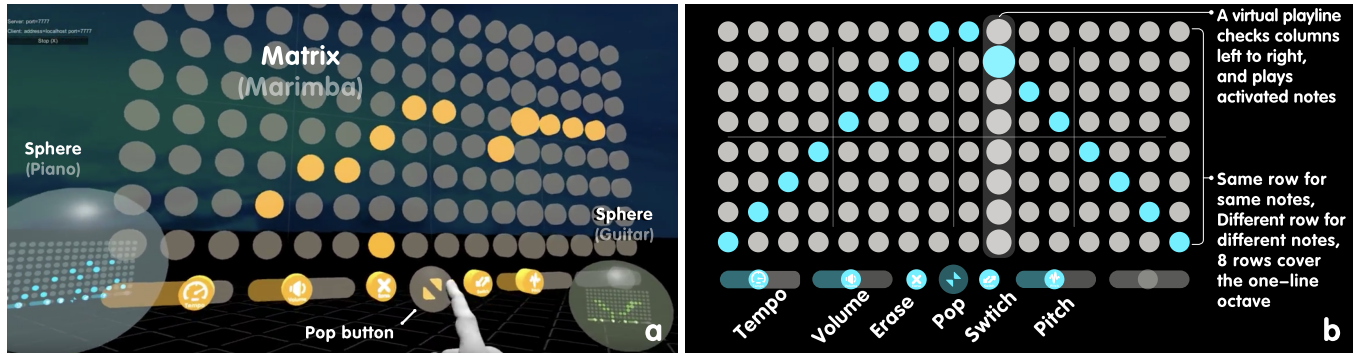[3]Leap Motion: https://www.ultraleap.com

Figure 2: (a) *Matrix* (opened interface) and *sphere* (packed interface), double click the pop button to switch in between. (b) A C major scale, starting from C4 and finishing at C5, and going back to C4.

between the two modes, re-positioned or removed by manipulating the *sphere* or the pop button of the *matrix* with corresponding bare-hand-based gestures. As shown in Fig. 2, the *matrix* interface contains a grid of 16 * 8 cells, each row represents the same pitch, starting from C4 and finishing at C5 from bottom to top, covering a whole octave. Users can edit notes with a simple tap using fingers or palms. A play-line moves from left to right playing corresponding notes recurrently. In this way, each interface produces a 16-note music loop. Three controllers (tempo, volume and pitch) and two functional buttons (erase and switch) are located at the bottom of the *matrix* interface. Bare-hands interaction is developed and integrated to enable users to generate, position and remove music interfaces. The maximum number of music interfaces is set to 8 to ensure a proper frame rate. (ii) Avatars - Each user has an avatar, including a head and hands (Fig. 1). Avatars are synchronised with users' real movements in real time, including position and rotation of heads, as well as gestures. LeMo synchronises the virtual environment, all virtual objects and avatars across the network, providing participants with the sense of being together in the same virtual environment and manipulating the same set of interfaces. (iii) A virtual environment that includes a grey stage with a grid pattern (part of it is shown in Fig. 1d). Highlighted boundary of the stage reminds users of the safe walking area. NB: Besides these 3 elements, LeMo also has a data-log system to log time-stamped users' interactive activities. The data of attention measurements used in Study II of this paper was collected based on this system.

4 STUDY I - PRODUCING PRIVACY USING BARRIERS

Studies have shown that different types of territory exist and serve different functions in table-top collaboration [25, 34]. In this paper, we explore the territory and territorial behaviour that emerged in a VR setting. In creative group-work, enabling people to shift between individual creativity and tightly coupled collaboration is needed [9, 13], so that the collaborators can develop their own ideas without affecting others. Studies have also shown that adding personal workspace is helpful for collaboration, and visibility of co-workers' workspace is preferred but has no notable positive effect on the uses of personal workspace and may even have some negative effects
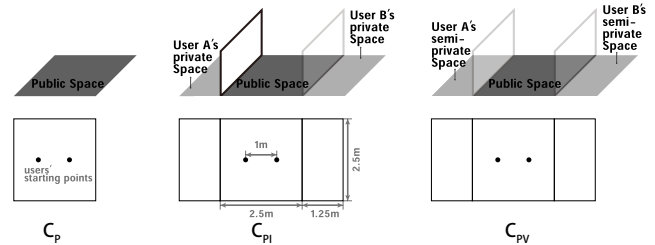

Figure 3: Three settings of spaces in Study I, directional view(upper), top view (bottom).

[10]. Physical environmental factors entail features such as barriers, location, layout, and distance can be used to produce privacy [22]. As such, Study I aims to explore if virtual barriers can produce privacy and see if the privacy is helpful in CMM.

4.1 Experiment Conditions

To do this, three experimental conditions are designed, as shown in Figure 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. For example, when user B is inside their personal space, user A cannot hear nor see user B and what's happening inside B's personal space. Note both of them can always hear and see what's happening in the public space, Figure 3 (middle) shows the setting from user B's perspective.

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.

4.2 Participants and Procedure

Students at the authors' university were recruited through group emails[4]. Each participant was compensated 10 GBP for their time. Twenty-one pairs of participants took part (25 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.

After reading and signing the information form and consent form, each pair of participants first received an explanation of music interface of LeMo (see Figure 2). Then one experimenter demonstrated all the types of interaction gestures supported in LeMo. By linking the demonstration with the first-person view shown on monitors, participants understood and grasped the tricks of how to play LeMo. 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 experimental conditions were explained to the participants before entering VR and trialled by them in the tutorial session. The time length of tutorial session is flexible to ensure participants with diverse musical knowledge could grasp LeMo.All the conditions were explained to the participants in the before entering VR and trialled in the tutorial session. Participants were then asked to have three sessions of composing music together, each lasting 8 minutes. To avoid the impact of adding personal spaces and have a pure observation on how participants form their proximity in the public space, all pairs started with C_P, and then C_{PI} and C_{PV} in a random sequence. Each session was followed by a post-session questionnaire. An interview was carried out at the end of experiment.

4.3 Results

4.3.1 C_P - Public Space only with No Personal Space.
The interview was recorded and transcribed, and a thematic analysis was carried out to analyse the qualitative data. e.g. participant reported having no privacy in C_P result in some issues. Participant 9B (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] (P_{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 private space, P_{5A} reported they could "only passively participate" as their partner P_{5B} played a "dominant" role in the creation. P_{9A} even reported C_P to be "the least collaborative session", which is quite different from other participants

4.3.2 C_{PI} & C_{PV} - Personal Space Available.
Participants held positive attitude on the addition of personal spaces in both conditions (C_{PI}:M=6.19, SD=2.67; C_{PV}: $M = 6.88$, $SD = 2.33$; 1 means not helpful at all, 10 means very helpful), and reported it to be helpful in the interview (P_{1A}[5], P_{2A}, P_{3A}, P_{4A}, P_{4B}, P_{5A}, P_{6A}, P_{9A}, P_{9B}, P_{10A}, P_{10B}, P_{11A}, P_{14B}, P_{15A}), it provides a chance to work independently without interfering each other. With personal space

available, P_{4A} and P_{10A} were able to "work independently" 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. With the addition of personal space, participants reported feeling "more comfortable taking risks" (P_{2A}), by enabling them to "do on their own thing" (P_{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}). They got a chance to complete "some of [their] ideas in [their] space". This became especially true when one participant was more dominant in the collaboration, e.g. while P_{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".

When being asked which is the best setting, C_{PV} was rated as the best setting by a significant number of participants (Binomial Test, 21 out of 42; 0.5 > 0.33, $p = 0.0187$, 1-sided), where the personal space is visible to collaborator. Conversely, when personal space became invisible in C_{PI}, participants had significantly shorter length of drawing attention (C_{PI}: $M = 52.53$, $SD = 62.03$, C_{PV}: $M = 98.25$, $SD = 60.76$; Wilcoxon Rank Sum Test: $W = 310$, $p = 0.00103$) and fewer times of drawing attention toward their collaborator's locations compared with C_{PI} (C_{PI}: $M = 15.82$, $SD = 11.80$, C_{PV}: $M = 30.59$, $SD = 14.98$; Wilcoxon Rank Sum Test: $W = 249.5$, $p = 5.70e-05$), leading a weaker sense of coworker's presence and activities, a significant portion of participants (20 out of 42) reported to have the weakest sense of collaborator's presence in C_{PI} (Binomial Test, 0.48 > 0.33, $p = 0.0384$, 1-sided). Reasons behind can be found from the interview. C_{PI} was reported to lead the "weakest sense of communication with the partner" (P_{18A}), and a more "isolated" feeling (P_{17A}). Being able to see each other all the time in C_{PV} made them feel "more communicated" (P_{5B}) and provides both "privacy and teamwork equally" (P_{5A}) and made it easier to work either "individually or cooperatively" (P_{18A}).

5 STUDY II - PRODUCING PRIVACY BY AUGMENTING SOUND ATTENUATION

Sound attenuates as a result of diminishing intensity when travelling through a medium. Our innate spatial abilities rely on this feature to retrieve and localise information and to aid performance (cf. [6]). Whilst it is hard to adjust the acoustic attenuation in the real world, an augmented spatialised sound can be simulated purposely within VR to enhance its potential, as the audio is artificially simulated. Many packages are available for spatial audio simulation, e.g. Oculus Native Spatializer[6], Google VR Audio Spatializer[7], and Steam VR Audio Spatializer[8]. Research has been done on investigating the impacts of spatialised sounds on user experience in VR [14]. However, little research explores how the spatialisation of sound may affect or aid collaboration (e.g. CMM). Considering that 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.

---

[4]The Queen Mary Research Ethics Committee granted ethical approval to carry out the study within its facilities (Ethical Application Ref: QMREC1872)
[5]P_{1A} refers to Participant 1A.

[6]https://developer.oculus.com/documentation/audiosdk/latest/concepts/book-ospnative-unity/
[7]https://developers.google.com/vr/reference/ios-ndk/group/audio
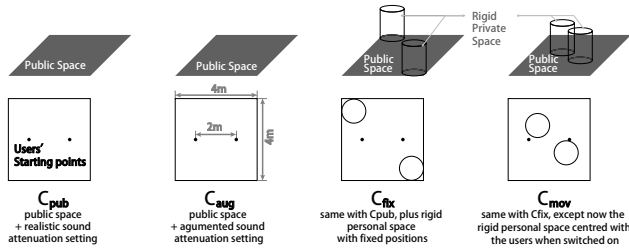[8]https://valvesoftware.github.io/steam-audio/

Figure 4: Four settings of spaces of in Study II, directional view(upper), top view (bottom).

5.1 Experiment Conditions

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. To explore this, 4 conditions are designed:

Condition 1: **Pub**lic space only (referred to as 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 C_{aug}). In addition to C_{pub}, the sound attenuation is augmented. The volume of audio drops much faster, creating sonic privacy, which can be 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 C_{fix}). In addition to C_{pub}, each user is now provided with a cylindrical, translucent, personal space located at the corner of the stage (see Figure 4), which works like an acoustically solid boundary between public space and personal space. In other words, the sonic privacy now shifts rigidly between personal space and public space. 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 C_{mov}). Every feature of this condition is the same as C_{fix}, except now when triggered, the personal space appears centring the user's current head's position.

5.2 Participants and Procedure

Fifty-two (26 pairs) participants were recruited for this study[9] via emails sent to group lists within the authors' school. All participants were aged between 18 and 35, with an average age of 23.00 (SD = 4.37). The procedure is roughly the same with Study I, after briefing and signing consent forms, each pair of participants received a tutorial and took a quick trial of all the 4 conditions. They then had 4 sessions of collaboratively composing music. Each session was followed by a post-session questionnaire. Conditions were

---

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

experienced in a fully randomised sequence to counterbalance the learning effect.

5.3 Results

5.3.1 ***Sonic Privacy Built by Audio Attenuation Augmentation***. According to statistical tests of attention measurements (see Table 1), compared with the other three conditions, in C_{aug} participants paid more attention to their collaborator. Possible reasons for this can be found from the thematic analysis. Compared with realistic acoustic attenuation in C_{pub}, C_{aug}'s augmented acoustic attenuation setting forced or prompted people to work more closely 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 whilst still providing the needed sonic privacy. These indicate that C_{aug} saw a less separated collaboration than C_{fix} and C_{mov}.

5.3.2 ***Preference for C_{aug}***. In the interview, 24 participants reported to be favouring condition C_{aug}, higher than 11 participants for C_{pub}, 8 participants C_{fix}, and 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 C_{aug}, much higher than any other conditions' advantages. C_{aug}'s advantages reported by participants can be grouped into 3 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 something 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}).

Table 1: Study II - Statistics of attention measurements and Wilcoxon Rank Sum Test (two-tailed) of Attention Measurements

Description	Means & SD				Friedman Test	Wilcoxon Rank Sum Test (two-tailed)					
	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 (Chi-square)	p (W)	p (W)	p (W)	p (W)	p (W)	p (W)
Time spent paying attention to collaborator (unit: second)[a]											
	7.19	14.04	5.51	9.43	0.0009759	0.01032	0.364	0.4725	0.001757	0.05722	0.1088
	(7.44)	(15.19)	(9.60)	(13.96)	(16.318)	(957)	(1492)	(1241)	(1833.5)	(1645)	(1105)
Number of times of paying attention to collaborator[a]											
	9.31	14.79	7.31	11.02	0.0009928	0.005451	0.1591	0.446	0.0001145	0.03122	0.02838
	(8.33)	(11.16)	(7.48)	(11.62)	(16.281)	(924.5)	(1568.5)	(1234.5)	(1945)	(1683.5)	(1015)

[a] Paying attention to collaborator means the collaborator is inside the user's central field of view (FOV), size of which roughly covers 27 degrees (horizontally), 28 degrees (vertically) and 1.0 m (depth). Depth 1.0m was chosen based on the distance of personal proximity, horizontal degree and vertical degree are decided based on the capability of HTC Vive headsets, which only provides a clear, comfortable 3D vision roughly within this FOV.

Table 2: Three channels of privacy of virtual artefacts

Channel	Rigid/Discrete way	Gradient/Step-less way
Visibility (Visual privacy)	Using opaque objects to block sights (e.g. walls in Study I)	Lower visibility attenuation space to delivery visual privacy (e.g. using simulated fog to reduce visibility of objects)
Hearability (Sonic privacy)	Using soundproof objects to block sounds (e.g. sound-proof walls in Study I)	Using sound attenuation to block sounds (e.g. C_{aug} in Study II)
Accessibility	Transplanting real accessibility concepts to VEs (e.g. locks and keys)	Control the accessibility with distance, the nearer, the more accessible (e.g. the musical interfaces can be re-positioned in both studies)

6 DISCUSSION

6.1 Managing Privacy and Openness of Shared Objects

Privacy gives people the ability to seclude themselves and schedule their activities independently [10], which is crucial for collaborative tasks. As briefed in Section 2.2 (privacy in collaboration), in the design of any multi-user system, privacy has long been an issue, e.g. how to visually represent users and the privacy states of objects in SVEs [8]. Here in Study I and II, the "privacy" of virtual artefacts (in this case, musical tools in shared virtual environments) include three main aspects:

(i) Visibility (visual privacy)- This concerns whether and how much a virtual artefact is visible to other people in the Virtual Environments, lower visibility of virtual artefacts to others indicates a higher level of visual privacy.

(ii) Hearability (sonic privacy) - This concerns whether and how much a virtual artefact is audibly detectable to other people in the Virtual Environments, lower hearability of virtual artefacts to others indicates a higher sonic privacy. This measurement potentially depends on two features of the objects: (i) Does the artefact produce sound, i.e. does it have sonic emissions in the VE, if so, what is the volume and how far that sound can travel. (ii) Does the artefact has simulated features that change sound travels, i.e. does the item block, reflect or change how sound travels. Both ways can make the object detectable by other people, and in turn, reduce its sonic privacy.

(iii) Accessibility. This concerns whether and how well a virtual artefact is accessible to other people in the virtual environment. In most cases, a closer distance to others means higher accessibility for other people and lower accessibility-privacy, though accessibility to others can be constrained in other ways (e.g. by identity check, only accessible to users who are the owner).

NB: There are also other modalities apart from visibility and hearability, e.g. capability of being smelled, however, LeMo did not involve any technology to support other modalities and currently the majority VEs do not include other sensations either. Therefore, here we only focus on discussing 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", assuming that only explicitly shared things are public, while all other things on a user's computer screen are private by default, simply because most items on the screen inherently have nothing to do with the collaboration [3]. However, what works well in 2D groupware may not work so well when it comes to SVEs, the aim of which is all about building an illusion that all players are present in the same VE, interacting with the same set of virtual objects, i.e. if an object exists in one player's VE, it should also exist in other players' VE. Blocking sight of unshared virtual objects will break this illusion and potentially can harm the sense of co-presence. In a natural sense, everything in the VE should be the same for all the users, including their appearance

and properties, and discoverability, people should hear, see and sense the same set of virtual objects if we want to convince them that they are in the same virtual world. Hence, a way is needed to produce privacy with no/low impact on the discoverability of virtual objects. Study II presents a possible solution.

6.2 Two Ways to Construct Privacy in SVEs

Although Study I approves the soundproof wall can deliver privacy to support creativity, the privacy it provides remains rigid in two ways: (i) the location of the boundary of the privacy is fixed, which restricts where users can use private working space. (ii) the level of privacy is fixed and binary, either a complete privacy or no privacy, i.e. users can either hear or not hear, see or not see, with no ability to adjust in between. Privacy with adjustable levels can facilitate collaborative music making in non-immersive media [10, 12], and we believe such privacy can also help CMM in VR, however, C_{PV} and C_{PI} in Study I does not deliver this. Study II, instead, uses an alternative way to deliver step-less privacy.

To manage the hearability of virtual objects, as concluded in Table 2, two ways have been practised in Study I and Study II: (i) Using rigidly-sound proof objects, like the virtual soundproof wall in Study I. The practice of Study I and II has shown that users are able to understand and get used to the soundproof wall concept. (ii) Using gradient spatial acoustic attenuation, i.e. the acoustic attenuation in C_{aug} in Study II, 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 (NB: this suggestion is very preliminary, as it is not directly drawn from the findings of the two 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 the accessibility of objects, two ways have been practised by the two studies: (i) Completely blocking certain users from a certain area or objects. For example, in Study I, participants were asked not to enter each other's personal space. In this way, the user with 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 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 (NB: this is a preliminary suggestion as it is not directly drawn from the two 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 farther away.

6.3 Balancing Privacy and Openness

Finding a balance between privacy and openness can be important. A total openness was provided in C_P of Study I and C_{pub} of Study

II. In these conditions, users have accessibility to all music interfaces, and can hear and see all the objects. According to the results of Study I, 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, this also echoes part of the findings of [5]. 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 I and II have shown the necessity of having personal space. This echoes the argument that privacy is needed during collaborative work to obtain "the freedom to be left alone" - definition of privacy by [32]. The opposite extreme of privacy is to make everything private, including visibility, hearability and accessibility of objects. C_{PV} of study I can be an example of such setting. The disadvantages are clear, according to findings related to C_{PV}, 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 the two studies, we propose 2 suggestions for the balancing: (i) Construct privacy using the channel for content producing, i.e. audio channel for music production, leaving as many channels as possible public to promote the construction of team-awareness. This is because the two studies have shown the privacy provided via audio channel is sufficient to aid collaboration, further blocking other channels (e.g. visual channels was blocked in C_{PI}) leads to negative impact. (ii) Giving users freedom to adjust the level of privacy. For example, different from other conditions, C_{aug} provides privacy with gradient levels, enabling user to adjust level of privacy simply by adjusting their location to the virtual artefact, and results show that C_{aug} received an overwhelming preference.

7 DESIGN IMPLICATIONS

Following the points made in the discussion, three implications can be proposed to inform the design of privacy in CMM in SVEs.

(i) Choice of modality to deliver privacy - When there are 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 are rigid requirements on audio outputs, using tools based on other modalities is suggested, such as tools based on visual channel.

(ii) Attenuates the modality to deliver privacy - Augment the physics of the virtual space. Augment how the dissemination of modalities (e.g. light/sound/odour) when travelling across the space. For example, in C_{aug} of Study II, the acoustic attenuation was found to be a powerful tool to enhance the music production. And according to the findings of Study II, this light-weight form tool has fewer negative impacts than some other solid form of tools (e.g. the solid soundproof walls in Study I). 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.

(3) Binary vs gradient shift between privacy and openness - Where possible, a smooth shift should be provided to allow users to adjust the level of privacy and stay at the level where they feel comfortable. Study II has shown, for music making, the augmented attenuation can be a very useful tool. This should not be limited to music-related tasks, because visual privacy can potentially be manipulated in a similar way (e.g. adjusting the density of the fog in the VE to deliver visual privacy). Table 2 has listed some example ways.

8 CONCLUSION

Privacy is a powerful tool to aid collaboration. This paper explores how to build privacy for collaboration in music making in Virtual Reality, sonic privacy in particular. Through two studies, two approaches have been explored: a discrete way where there is either privacy or no privacy and a continuous way, where the level of privacy is adjustable. A comparison between these two ways has been made and discussed. Three channels for designing privacy of virtual objects have been identified (visual privacy, auditory privacy and accessibility) and three implications have been proposed for designing privacy for collaborative tasks in shared virtual environments, e.g. CMM.

Future directions would be: exploring privacy design for other modality based tasks, e.g. drawing is based on visual modality, investigating how other modalities can be attenuated to provide privacy, and finding other ways to deliver step-less privacy. Each direction corresponds to and further develops one of the three implications given above. Privacy in this paper is more a spacial, physical, materialised privacy. Although how VR intersects with mental, psychological privacy is also an emerging research area, that is outside the scope of this paper.

REFERENCES

[1] Irwin Altman. 1975. *The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding.* ERIC, Brookes, Monterey, CA, USA.
[2] Irwin Altman and Martin M Chemers. 1984. *Culture and environment.* Cambridge University Press, Cambridge, England, UK.
[3] Ronald M Baecker. 1993. *Readings in Groupware and Computer-Supported Cooperative Work: Assisting Human-Human Collaboration.* Morgan Kaufmann, San Mateo, CA, USA.
[4] Álvaro Barbosa. 2003. Displaced soundscapes: a survey of network systems for music and sonic art creation. *Leonardo Music Journal* 13 (2003), 53–59.
[5] 2020. *Adapting & Openness: Dynamics of Collaboration Interfaces for Heterogeneous Digital Orchestras.* Birmingham, United Kingdom. https://hal.archives-ouvertes.fr/hal-02561064
[6] Mark Billinghurst and Hirokazu Kato. 2002. Collaborative augmented reality. *Commun. ACM* 45, 7 (2002), 64–70.
[7] Tina Blaine and Sidney Fels. 2003. Contexts of Collaborative Musical Experiences. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression* (Montreal, Quebec, Canada) *(NIME '03).* National University of Singapore, SGP, 129–134.
[8] Andreas Butz, Clifford Beshers, and Steven Feiner. 1998. Of Vampire Mirrors and Privacy Lamps: Privacy Management in Multi-User Augmented Environments. In *Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, California, USA) *(UIST '98).* Association for Computing Machinery, New York, NY, USA, 171–172.
[9] Paul Dourish and Victoria Bellotti. 1992. Awareness and Coordination in Shared Workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work* (Toronto, Ontario, Canada) *(CSCW '92).* Association for Computing Machinery, New York, NY, USA, 107–114.
[10] Robin Fencott and Nick Bryan-Kinns. 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).* University of Technology Sydney, Sydney, Australia, 198–203.

[11] Saul Greenberg, Michael Boyle, and Jason LaBerge. 1999. PDAs and shared public displays: Making personal information public, and public information personal. *Personal Technologies* 3, 1-2 (1999), 54–64.
[12] Michael Gurevich. 2006. JamSpace: Designing a Collaborative Networked Music Space for Novices. In *Proceedings of the 2006 Conference on New Interfaces for Musical Expression* (Paris, France) *(NIME '06).* IRCAM — Centre Pompidou, Paris, France, 118–123.
[13] Christian Heath, Marcus Sanchez Svensson, Jon Hindmarsh, Paul Luff, and Dirk Vom Lehn. 2002. Configuring awareness. *Computer Supported Cooperative Work (CSCW)* 11, 3 (2002), 317–347.
[14] Claudia Hendrix and Woodrow Barfield. 1996. The sense of presence within auditory virtual environments. *Presence: Teleoperators & Virtual Environments* 5, 3 (1996), 290–301.
[15] Carl Anderson Johnson. 1974. Privacy as personal control. *Man-environment interactions: evaluations and applications: part* 2 (1974), 83–100.
[16] Karel Kreijns, Paul A Kirschner, and Wim Jochems. 2003. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior* 19, 3 (2003), 335–353.
[17] Liang Men and Bryan-Kinns. 2018. LeMo: Supporting Collaborative Music Making in Virtual Reality. In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE).* IEEE, Reutlingen, Germany, 1–6. https://doi.org/10.1109/SIVE.2018.8577094
[18] Liang Men and Nick Bryan-Kinns. 2019. LeMo: Exploring Virtual Space for Collaborative Creativity. In *Proceedings of the 2019 on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19).* ACM, New York, NY, USA, 71–82. https://doi.org/10.1145/3325480.3325495
[19] Liang Men, Nick Bryan-Kinns, and Louise Bryce. 2019. Designing spaces to support collaborative creativity in shared virtual environments. *PeerJ Computer Science* 5 (Nov. 2019), e229. https://doi.org/10.7717/peerj-cs.229
[20] Liang Men, Nick Bryan-Kinns, Amelia Shivani Hassard, and Zixiang Ma. 2017. The impact of transitions on user experience in virtual reality. In *2017 IEEE Virtual Reality (VR).* IEEE, Los Angeles, CA, USA, 285–286.
[21] Meredith Ringel Morris, Dan Morris, and Terry Winograd. 2004. Individual Audio Channels with Single Display Groupware: Effects on Communication and Task Strategy. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) *(CSCW '04).* Association for Computing Machinery, New York, NY, USA, 242–251.
[22] Darhl M. Pedersen. 1997. Psychological Functions of Privacy. *Journal of Environmental Psychology* 17, 2 (1997), 147–156. https://doi.org/10.1006/jevp.1997.0049
[23] Claude Raffestin. 2012. Space, territory, and territoriality. *Environment and Planning D: Society and Space* 30, 1 (2012), 121–141.
[24] Beate Roessler and Dorota Mokrosinska. 2013. Privacy and social interaction. *Philosophy & Social Criticism* 39, 8 (2013), 771–791. https://doi.org/10.1177/0191453713494968 arXiv:https://doi.org/10.1177/0191453713494968
[25] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori Inpken. 2004. Territoriality in Collaborative Tabletop Workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) *(CSCW '04).* Association for Computing Machinery, New York, NY, USA, 294–303.
[26] Chia Shen, Katherine Everitt, and Kathleen Ryall. 2003. UbiTable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *UbiComp 2003: Ubiquitous Computing. UbiComp 2003. Lecture Notes in Computer Science*, A.K. Dey, A. Schmidt, and J.F. McCarthy (Eds.). Vol. 2864. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-39653-6_22
[27] Jonathan Steuer. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of Communication* 42, 4 (1992), 73–93.
[28] Masanori Sugimoto, Kazuhiro Hosoi, and Hiromichi Hashizume. 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* (Vienna, Austria) *(CHI '04).* Association for Computing Machinery, New York, NY, USA, 41–48.
[29] Luca Turchet, Carlo Fischione, Georg Essl, DamiáN Keller, and Mathieu Barthet. 2018. Internet of Musical Things: Vision and Challenges. *IEEE Access* 6 (2018), 61994–62017. https://doi.org/10.1109/ACCESS.2018.2872625
[30] Luca Turchet, Rob Hamilton, and Anil Çamci. 2021. Music in Extended Realities. *IEEE Access* 9 (2021), 15810–15832.
[31] Anne Vinsel, Barbara B Brown, Irwin Altman, and Carolyn Foss. 1980. Privacy regulation, territorial displays, and effectiveness of individual functioning. *Journal of Personality and Social Psychology* 39, 6 (1980), 1104.
[32] Samuel D Warren and Louis D Brandeis. 1890. Right to privacy. *Harv. L. Rev.* 4 (1890), 193.
[33] Carol M. Werner, Irwin Altman, and Barbara B. Brown. 1992. A Transactional Approach to Interpersonal Relations: Physical Environment, Social Context and Temporal Qualities. *Journal of Social and Personal Relationships* 9, 2 (1992), 297–323. arXiv:https://doi.org/10.1177/0265407592092008
[34] Anna Xambó, Eva Hornecker, Paul Marshall, Sergi Jordà, Chris Dobbyn, and Robin Laney. 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 (2013), 36.