Designing spaces to support collaborative creativity in Shared Virtual Environments

Liang Men [Corresp., 1], **Nick Bryan-Kinns** [1], **Louise Bryce** [1]

[1] School of Electronic Engineering and Computer Science, Queen Mary University of London, London, United Kingdom

Corresponding Author: Liang Men
Email address: l.men@qmul.ac.uk

Shared Virtual Environments (SVEs) have been researched extensively within the fields of education, entertainment, work, and training, yet there has been limited research on the creative and collaborative aspects of interactivity in SVEs. The important role that creativity and collaboration play in human society raises the question of the way that virtual working spaces might be designed to support collaborative creativity in SVEs. In this paper, we outline an SVE named LeMo, which allows two people to collaboratively create a short loop of music together. Then we present a study of LeMo, in which 52 users composed music in pairs using four different virtual working space configurations. 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 and iii) a configuration that provides a movable personal space was preferred to one that provided no mobility. Following these findings, five corresponding design implications for Shared Virtual Environments focusing on supporting collaborative creativity are given and conclusions are made.

Designing Spaces to Support Collaborative Creativity in Shared Virtual Environments

Liang Men[1], Nick Bryan-Kinns[2], and Louise Bryce[3]

[1,2,3]**School of Electronic Engineering and Computer Science, Queen Mary University of London, United Kingdom**

Corresponding author:
Liang Men[1]

Email address: l.men@qmul.ac.uk

ABSTRACT

Shared Virtual Environments (SVEs) have been researched extensively within the fields of education, entertainment, work, and training, yet there has been limited research on the creative and collaborative aspects of interactivity in SVEs. The important role that creativity and collaboration play in human society raises the question of the way that virtual working spaces might be designed to support collaborative creativity in SVEs. In this paper, we outline an SVE named LeMo, which allows two people to collaboratively create a short loop of music together. Then we present a study of LeMo, in which 52 users composed music in pairs using four different virtual working space configurations. 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 and iii) a configuration that provides a movable personal space was preferred to one that provided no mobility. Following these findings, five corresponding design implications for Shared Virtual Environments focusing on supporting collaborative creativity are given and conclusions are made.

INTRODUCTION

The real world envelops us with space that we share with others, and in this surrounding environment we perceive rich sensory information about objects and events happening around us. Using this information, we interact with this outer world around us via inference, manipulation, and exploration. In a similar fashion we interact with other people. 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 proxemics 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 is *Dungeons & Dragons* which was created in the early 1970s, it provides a computational space that players can visit and experience through text description on a computer screen (Mead and Malcomson, 2003). Though these non-immersive media can involve people to a very high level and generate the experience of flow, none of them enable 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 multisensory 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? We believe that having a better understanding of the role of space and territory within creative collaborations would 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 utilization of a working space can possibly be a crucial factor to a successful collaboration in SVEs too. Thus we are keen on designing and testing working space configurations to see if we can provide more fluid support to creative collaboration in SVEs. We begin by reviewing related work in SVEs, space territory, and territoriality. Then the design of our SVE system will be detailed and the study and results will be presented. Finally, the overall study will be discussed and design implications will be given.

RELATED WORK

Shared Virtual Environments

The term Virtual Environment (VE) can be traced back to the early 1990s (Brooks Jr et al., 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 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) (Schroeder, 2012). Besides "SVEs", other terms used include multi-user virtual environments, multi-user virtual reality (Carlsson and Hagsand, 1993), Collaborative Virtual Environments (CVEs) (Zhang and Furnas, 2003) and Social Virtual Reality (SVR). To align with mainstream usage, we will herein use the term SVEs 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. SVEs can be seen as a convergence of research interests in VR and Computing Supported Cooperative Work (CSCW) (Benford et al., 2001).

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). By enabling multiple people to communicate and interact with each other and providing a natural medium for three-dimensional CSCW (Billinghurst et al., 2000), SVEs are considered emerging tools for a variety of purposes, including community activities (Lea et al., 1997), online education (Roussos et al., 1997), distributed work and training (Nedel et al., 2016), gaming and entertainment (Oculus, 2015; Plante, 2016). Despite this, little research exist in the field of supporting collaborative creativity, leaving many open questions: does collaborative creativity in SVEs follow a similar pattern to real-world collaborative creativity? How should the virtual environment be designed to support creative collaboration (cf. Basdogan et al., 2000)?

Space, Territory, and Territoriality

SVEs constitute virtual spaces, although illusive they are meaningful (Steuer, 1992). We believe gaining a better understanding of the virtual space is an effective way to answer the aforementioned questions. "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), helping people 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; Men and Bryan-Kinns, 2019). Thus, potentially, with more knowledge of the virtual space, we can even manipulate the virtual space to influence the collaboration in SVEs. Note in this paper, the term "space" 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 paper.

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). In CSCWs that focus on supporting collaborative creativity, providing personal space is argued to be useful (Fencott and Bryan-Kinns, 2010; Men and Bryan-Kinns, 2019), and integrating personal and group spaces, allowing users 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 is 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. They 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, by providing a flexible gradient of sharing semantics, Shen et al. (2003) addressed this concern in their project UbiTable. In this study, we want to design virtual spatial configurations that provide a more gradual boundary between personal and public spaces and enable a fluid shift between these.

Territory and Territoriality in Collaboration (SVEs and Tabletop)

Because there is limited research on territoriality in VR, and rich research on this in Tabletop-based collaboration, we review territoriality in collaboration not only in SVEs, but also in Tabletop. In a previous study, we found collaborators formed both personal and group territory during collaborative music making in an SVE, and they also had territorial behaviour, e.g. most musical edits were done inside persona territories (Men and Bryan-Kinns, 2019). By manipulating the virtual spatial configurations of an SVE, the formation of territory and territorial behaviour can be influenced and, and as a result, the collaboration can be influenced (Men and Bryan-Kinns, 2019). Similarly, territoriality also plays an important role in the tabletop collaboration. Collaborators 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 note that removing territorial constraints can promote exploratory group activity (Xambó et al., 2013). Two main types of territory have been identified from research on screen and tabletop mediated collaboration:

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

In terms of designing for territoriality, Scott et al. (2004) proposed four guidelines are suggested 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 and Bryan-Kinns, 2010). However, this can result in overloaded cognitive information, which some people found to be difficult to handle (Fencott and Bryan-Kinns, 2010). To date, little research has explored how such features of territoriality might be designed for and used in SVEs.

EXPERIMENT DESIGN

Creativity domain: why collaborative music making

Music making, as a collaborative activity that relies on shared goals, understanding and good interpersonal communication, has long been a key form of collaborative creativity (cf. Bryan-Kinns and Hamilton, 2012; Titon and Slobin, 1996). Its unique features make it an excellent activity through which to study collaborative activity. In 2003, Blaine and Fels explored the design criteria of collaborative music-making (CMM) systems and pointed out key features including the media used, player interaction, the systems' learning curves, physical interfaces and so on. In the same year, with inspiration from Rodden's (1991) Classification Space for collaborative software, otherwise known as groupware, Barbosa (2003) developed the Networked Music Systems Classification Space, which classifies CMM systems (CMMs) in terms of the time dimension (synchronous/asynchronous) and space dimension (remote/co-located). For instance, Daisyphone (Bryan-Kinns, 2004), which provides shared editing of short musical loops falls into the remote synchronous network music systems in this Classification Space. Other examples include reacTable (Xambó et al., 2013) and BilliArT (Bressan et al., 2017), both of which provide co-located music-making experience, and Ocarina (Wang, 2009), which provides a distributed experience. However, we should

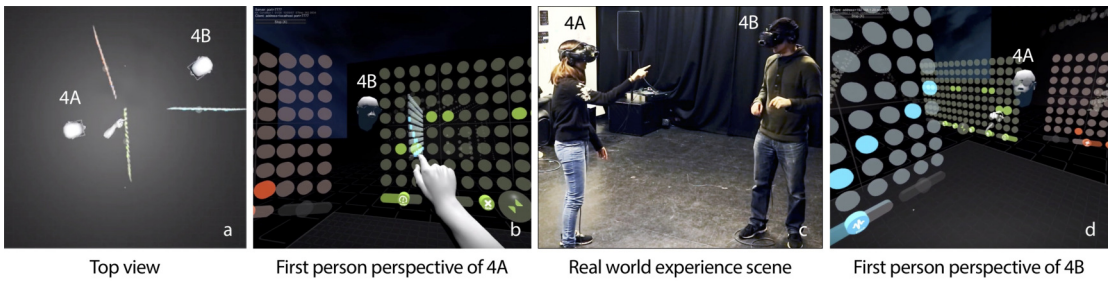

Figure 1. Participant 4A and 4B are creating music together.

note that despite decades of research into CMMs and SVEs, relatively few Shared Virtual Environments (SVEs) that support CMM have been made. As a result many basic but crucial questions in this field are waiting to be answered, e.g. do CMMs in SVEs follow a similar pattern to CMMs in the real world? How can CMMs in SVEs be supported?

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 and 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, we can simulate an augmented spatialised. Research has been done on investigating the impacts of spatialised sounds on user experience in VR, (Hendrix and Barfield, 1996). 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 (Men and Bryan-Kinns, 2019), 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, we can possibly create sonic privacy by augmenting acoustic attenuation, this privacy can then possibly be used as personal space supporting individual creativity in CMM.

LeMo - An SVE for collaborative music making

We created Let's Move (LeMo[1]), which enables two users to manipulate virtual music interfaces together in an SVE to create a 16-beat music loop. Note herein LeMo is an extensively modified version of previous versions (Men and Bryan-Kinns, 2018, 2019). Hereafter, "LeMo" specifically refers to this modified version. LeMo was programmed in Unity, models and textures were made in Cinema 4D and Photoshop respectively. The run-time environment includes two HTC Vive headsets (each with Leap Motion mounted, see Figure 1c) and two PCs connected and synchronised via a LAN cable.

LeMo has three key elements: (1) Music interface - LeMo allows users to generate, remove, position and edit virtual music interfaces, which have two modes: *sphere* and *matrix* (Figure 2b). Users can generate up to 8 *sphere*s with pinch and stretch gesture, see Figure 2a. Both the *sphere* and the *matrix* can be switched between, re-positioned or removed by manipulating the *sphere* of the pop button of the matrix with corresponding gestures. As shown in Figure 3, the *matrix* interface contains a grid of 16 * 8 dots, with controllers at the bottom. Each row represents the same pitch, forming an octave from bottom to top, see Figure 3. 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-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. (2) Avatars - Each user has an avatar, including a head and both hands (Figure 1). Avatars are synchronised with users' real movements in real time, including position and rotation of heads, as well as gestures. LeMo provides visual aids for collaboration by synchronizing 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

[1]Full source available at: https://sites.google.com/view/liangmen/projects/LeMo

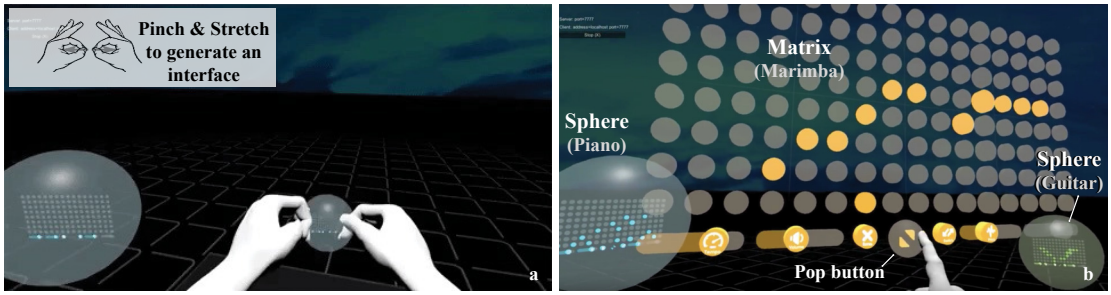

Figure 2. (a) The gesture to generate a new interface; (b) *Matrix* (opened interface) and *sphere* (packed interface), double click the pop button to switch in between (this figure is reproduced from Men and Bryan-Kinns 2019).

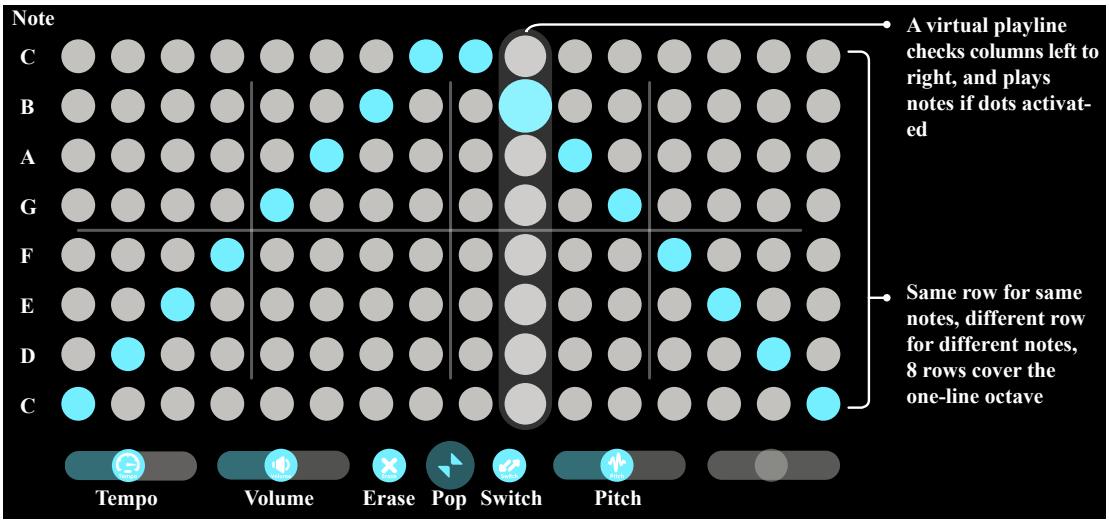

Figure 3. A C major scale, starting from C4 and finishing at C5, and going back to C4.

interfaces. (3) A virtual space that includes a grey stage with a grid pattern (part of it is shown in Figure 1d). Four types of spatial configurations were designed for this study which will be detailed later.

Besides these three fundamental elements, LeMo also has: spatialised audio (volume drops with distance) so that users can hear where the sounds originate; A voice notification system to facilitate the experiments, e.g. in experimental scenario users will hear "1 minute left" and "end of session" notifications; A data-log system to log 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), all information is logged with a time stamp.

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 a previous study (Men and Bryan-Kinns, 2019), following this implication, we built three different spatial configurations (public space only, public space + publicly visible personal space, public space + publicly invisible personal space), and tested 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 (Men and Bryan-Kinns,

2019). We believe 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 we are keen on designing some new types of personal territory to eliminate these disadvantages, and to provide a more flexible, more fluid experience to the collaborators. To increase the flexibility, we want to enable users to use personal space anywhere on the stage in SVE, and see how this flexibility might positively affect the collaboration, thus we developed H1.

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), we thought the attenuation feature could 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 and Bryan-Kinns, 2018), thus by manipulating acoustic attenuation, we can produce sonic privacy. E.g. 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 we developed H3. 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 (cf. Men and Bryan-Kinns 2019) 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.

Independent variable

Spatial configuration is the independent variable in this experiment. To investigate these three hypotheses we designed four space configurations as the independent variable levels, as shown in Figure 4, 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 a sonic privacy, which can be seen as a personal space. As the volume changes gradually with the changes of 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 4), which works like a 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, avoiding making conditions more artificial.

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. The Igroup Presence Questionnaire (IPQ) was used to inform the design

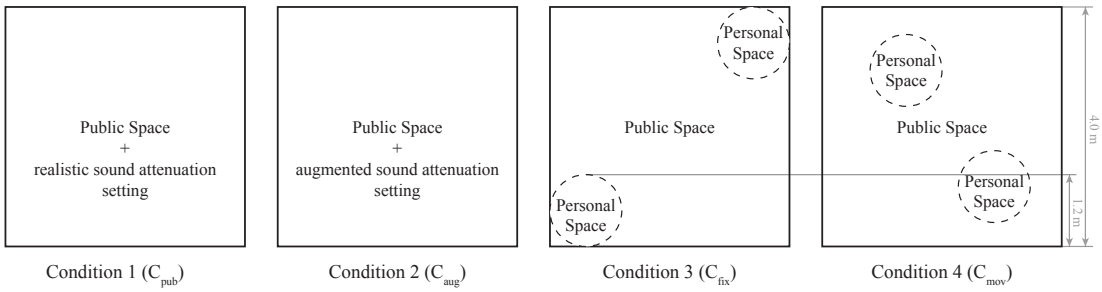

Figure 4. Top view of the four experimental condition settings

of questions about the sense of the collaborator's presence (Schubert et al., 2001). Questions about output quality, communication, and contribution 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) Presence: i) Sense of self-presence, ii) sense of co-worker's presence and iii) sense of collaborator's activities.

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

(3) Content assessment: the satisfaction of the final music created reflects the quality of collaboration, cf. (Bryan-Kinns, 2013; Bryan-Kinns and Hamilton, 2012).

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

(5) Contribution: i) the feeling of self's contribution; ii) the feeling of others' contribution. By setting these measures, we want to see the effects of spatial configurations on the sense of contribution.

These measures will be grouped into a Post-Session Questionnaire (PSQ, see items in Table 1), to be filled after participants experiencing each condition, and a Comparison Questionnaire (CQ, see items Table 2), to be filled at the end of the experiment.

Activity Assessments

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

(1) Contribution: i) number of musical note edits; ii) number of note additions/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).

(2) 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.

(3) Location and territory: i) distribution of participants' locations and interactions; ii) the sizes of personal/group territory if they emerge; iii) note edits fallen in different types of territory; iv) average distance between participants, cf. colocation in (Bryan-Kinns, 2013).

(4) Attention: i) time participants spent paying close/ordinary attention to collaborator; ii) number of times paying close/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.

Participants and Procedure

Fifty-two participants (26 pairs) were recruited for this study[2], 37 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. After reading the information form and signing the consent form, each pair of participants first received an explanation of the music interface of LeMo (see Figure 3). Then one experimenter demonstrated all of the interaction gestures supported

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

Table 1. Results of Post-Session Questionnaire[a] and results of Wilcoxon Rank Sum Test (two-tailed)[b]

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)*
PSQ1 (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)
PSQ2 (support for creativity)- I feel like the spatial configuration in this session was extremely helpful 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)
PSQ3 (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)
PSQ4 (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)
PSQ5 (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)
PSQ6 (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)
PSQ7 (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)
PSQ8 (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)
PSQ9 (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)
PSQ10 (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)
PSQ11 (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)

[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 in this table are calculated based on the data collected from the third and fourth session to counterbalance the learning effect.

in LeMo. By linking the demonstration with the first-person view shown on monitors, participants had a chance to learn how to play LeMo. 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. 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 our pilot study and a previous study (Men and Bryan-Kinns, 2018), we found 7 minutes were sufficient for the task. 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 randomized 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 randomized sequence to counterbalance impacts on the study. Each session ended with a Post-Session Questionnaire (PSQ, see Table 1). After all the four sessions finished, the Comparison Questionnaire (CQ, see Table 2) and a short interview were carried out at the end of the experiment.

RESULTS

Participant Reports

In this section, we report 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 1). Binomial tests were run to see if the number of ratings for each option was significantly different than would be expected by chance, upper-tail, lower-tail or two-tailed tests were used accordingly, see results in Table 2. Next, we will present how we counterbalanced the learning effect on PSQ and then results will be reported following the sub-type of measures.

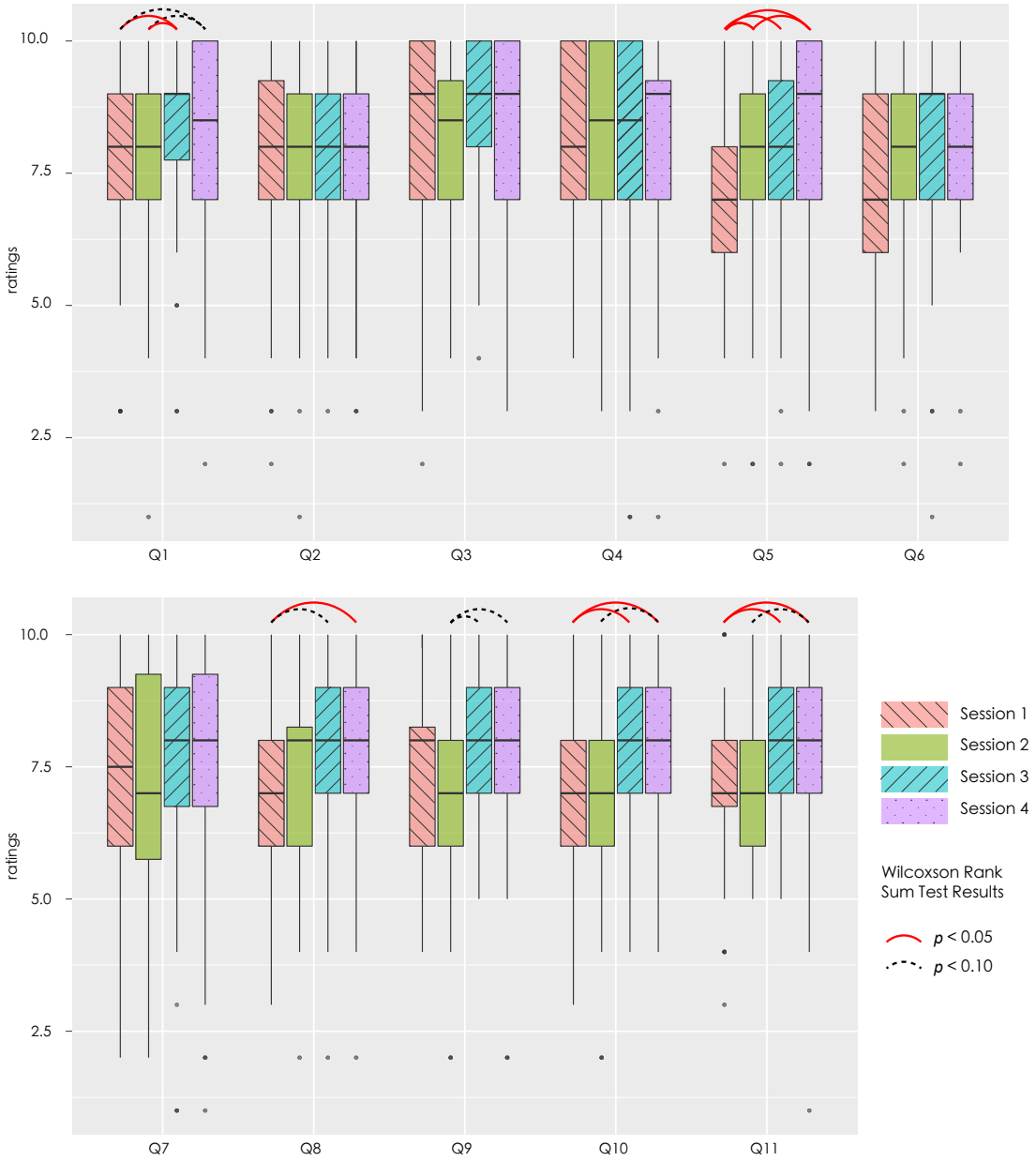

Figure 5. Results of Post-Session Questionnaire, data of all sessions included; arcs showing significant/marginal-significant differences between conditions. These arcs indicate some questions of PSQ were heavily influenced by the sequence of conditions.

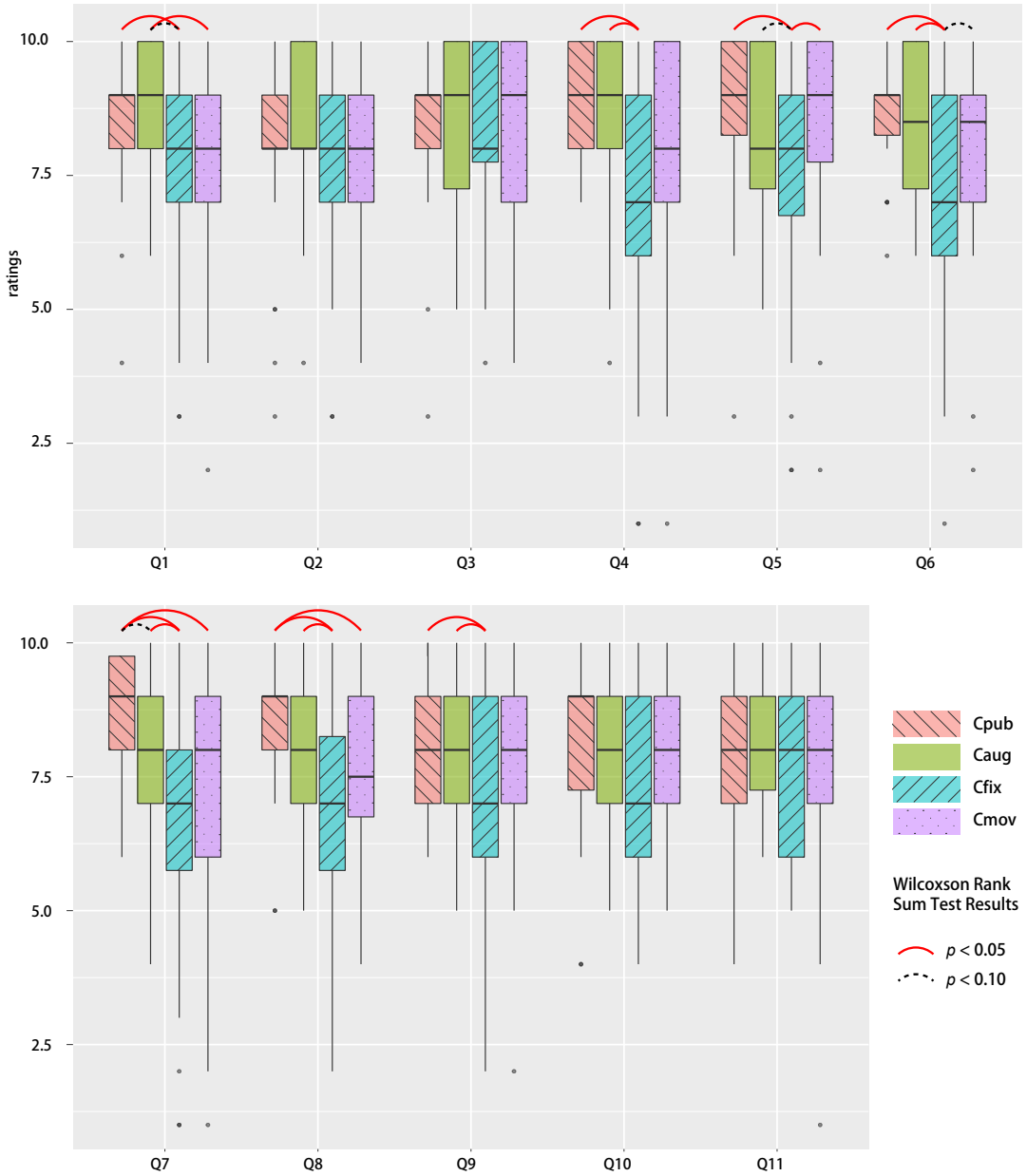

Figure 6. Results of Post-Session Questionnaire, data grouped by experimental conditions (only data collected in the latter two sessions are included; arcs showing significant/marginal-significant differences between conditions).

Table 2. Results of Binomial Test of Comparison Questionnaire (CQ)[a]

Question description	Option	C_{pub}		C_{aug}		C_{fix}		C_{mov}	
		k	p	k	p	k	p	k	p
CQ1 (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
CQ2 (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
CQ3 (coordination) - Which session you found most difficult to track collaborator's activities?									
	most difficult	7		12	0.4469	20	0.02205	13	1.000
	least difficult	22	0.004691	14	0.4262	8	0.06971	8	0.06971
CQ4 (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
CQ5 (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
CQ6 (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
CQ7 (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
CQ8 (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
CQ9 (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

[a] Lower-tailed test when $k < 13$, two-tailed test when $k = 13$, upper-tailed test when $k > 13$.

Counterbalancing the learning effect

As aforementioned, we introduced a fully randomized order of experimental conditions to counterbalance the learning effect. However, it turned out many measurements in the Post-Session Questionnaire were still relatively affected by the sequence, as shown in Figure 5, in which data from all groups were compiled according to how the group was ordered in the session sequence. Wilcoxson Rank Sum tests were run between each two conditions of every question. An orange arc indicates a significant difference between two bars ($p < 0.05$), and a grey arc indicates a trend toward a significant difference ($p < 0.1$). The arcs show that results of some questions (e.g. Q1, Q5, Q10, Q11) 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' (Q1), higher satisfaction with their output (Q5), and both more, and better quality of contributions by themselves and contributors (Q8, Q9, Q10, Q11). This is probably due to the learning effect which has a much stronger effect on these measures compared with the differences between experimental conditions, considering the limited experience participants have in VR and collaborative music making, hence learning effect can strongly promote 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. 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, we chose to only use data collected via PSQ in later two sessions (session 3 and 4) at the expense of the halved sample size. Box-plots were then drawn (Figure 6) and Wilcoxon Rank Sum tests were run (Table1) to compare the conditions against each other.

General feeling (helpfulness of spatial configuration, difficulty of cooperation)

When being asked the helpfulness of spatial configuration (PSQ1), 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 is a trend toward participants rating C_{aug} higher than C_{mov} ($W = 469$; $p = 0.06318$), and rating C_{pub} higher than C_{fix} ($W = 396.5$; $p = 0.07706$). These differences indicating C_{aug} is believed to have relatively higher helpfulness for creativity, and C_{fix} was believed to be less helpful.

When being asked the helpfulness of spatial configuration to support personal idea development (PSQ2), the mean rating of C_{aug} ($M = 8.35$) is higher than the other three conditions (C_{pub}: $M = 7.82$; C_{fix}: $M = 7.71$; C_{mov}: $M = 7.75$), but no significant difference were found. CQ7 of Table 2 shows that C_{pub} was rated by significantly many (21 out of 52; Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided) to be the least difficult to cooperate with their collaborator, significantly few 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 only 7 out of 52 rated it as the least difficult (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided).

Preference

When being asked the enjoyment of the spatial configuration (PSQ3), similar to PSQ2, C_{aug} got a higher rating ($M = 8.65$) however, no significant differences were revealed. When being asked to which session has the most enjoyable spatial configuration, out of 52 participants,16 chose C_{aug}, 16 chose C_{mov}, higher than C_{pub} and C_{fix}, both of which were chosen by 10 participants, though no significant differences were revealed by the Binomial Test (see CQ1 of Table 2). When being 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 CQ6 in Table 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. These indicate the spatial configuration in C_{fix} is more disfavoured and that of C_{mov} is less disfavoured.

Sense of co-presence

Results of PSQ4, PSQ7 (see Table 1) and CQ4 (see Table 2) reveal participants' sense of collaborators' presence and activities. In PSQ4 C_{fix}'s ratings are significantly lower than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, all $p < 0.05$), indicating a lower sense of presence of the collaborator and the collaborator's activities in C_{fix} compared with C_{pub} and C_{aug}. CQ (Table 2)shows a similar result, in CQ4, significantly many participants believed they had strongest sense of collaborators in C_{pub} ($0.52 > 0.25$, $p = 2.807e\text{-}05$, 1-sided) and least strongest in C_{fix} ($0.54 > 0.25$, $p = 8.12e\text{-}06$, 1-sided). Significantly few participants reported they had the least strongest sense in C_{aug} ($0.13 < 0.25$, $p = 0.03317$, 1-sided), and significantly few participants (only 2 out of 52) believed they had the strongest in C_{fix}($0.04 < 0.25$, $p = 5.277e\text{-}05$, 1-sided) or in C_{mov} ($0.12 < 0.25$, $p = 0.1368$, 1-sided), indicating to most participants, C_{pub} is the best to maintain the sense of presence of collaborator, C_{mov} is worse, and C_{fix} is the worst, C_{aug} is worse than C_{pub}, but still much better than C_{mov} and C_{fix}.

Regarding the sense of collaborator's activities (PSQ7), a significantly weaker sense was reported in C_{fix} compared with C_{pub} and C_{aug} (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, CQ3 of the Comparison Questionnaire reveals that significantly many participants (Binomial Test, $0.38 > 0.25$, $p = 0.02205$, 1-sided) felt it to be most difficult tracking collaborators' activities in C_{fix}, and significantly many felt least difficult in C_{pub} ($0.42 > 0.25$, $p = 0.04691$, 1-sided). These indicate that C_{pub} seems to be easier for participants to track collaborators activities, and C_{fix} is more difficult for doing that.

Content assessments

Participants reported a mean rating 7.21 of output quality in C_{fix}, see PSQ5 of Table 1, which is significantly lower than 8.64 in C_{pub} ($W = 448.5$, $p = 0.005155$), and than 8.32 in C_{mov} ($W = 254$, $p = 0.02163$), and quasi-significantly lower than 8.38 in C_{aug} ($W = 473.5$, $p = 0.05549$). Similarly, significantly many participants believed it was in C_{fix} they produced the least satisfying piece of music (Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided), see CQ2 of Table 2. These all indicate that the spatial configuration in C_{fix} led to a music output with lower quality.

Communication assessments

Communication quality was reported significantly lower in C_{fix} ($M = 7.04$) than C_{pub} ($M = 8.68$; $W = 450.5$, $p = 0.004494$) and C_{aug} ($M = 8.50$; $W = 510$, $p = 0.01038$), and near-marginal significantly lower than C_{mov}($M = 8.04$; $W = 274$, 0.05073), see PSQ6 in Table 1 and Q6 in Figure 6. When being asked to compare these sessions, significantly many participants believed they had the best communication quality in C_{pub} and significantly few believed they had the best communication quality in C_{fix}. Conversely,

significantly few believed they had the worst communication quality in C_{pub} and significantly many believed they had worst in C_{fix}.

Contribution

Participants felt they had done a significantly larger amount of contributions in C_{pub} compared with C_{fix} ($W = 439.5$, $p = 0.009236$) or compared with C_{mov} ($W = 412$, $p = 0.03928$), and had done significant more contribution in C_{aug} compared with C_{fix}($W = 479.5$, $p = 0.04281$), see PSQ8. No significant difference was found in CQ8, which is also questioning the feeling of own contribution.

No significant differences were found in the rating of the amount of the collaborators' contribution (PSQ9), only a trend showing that participants reported their collaborator had a lower amount of contribution in C_{fix} than C_{pub} (C_{pub} vs C_{fix}: $W = 394$, $p = 0.08916$) and C_{aug} (C_{aug} vs C_{fix}: $W = 469.5$, $p = 0.06406$). In CQ9, significantly few participants reported they felt their collaborator did the most contribution in C_{mov} (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided). These indicate the addition of personal space in C_{fix} and C_{mov} possibly led to a weaker sense of collaborator's activities.

Activity Assessments

In this section we report on measures focusing on the participants' interactive activities. All measures are listed in Figure 3, Wilcoxon Rank Sum tests were run to compare conditions against each other.

Contribution

(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 note edits in C_{pub}, 80.27 edits in C_{aug}, and 77.69 edits in 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} (both $p < 0.07$, check detailed statistics in AA2, Table 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, 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.

Location and territory

To illustrate how participants used the space, we plotted their locations, directions and musical note edits on a top view of the stage, see Figure 7 as an example. We call these plots visual traces. 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 and Carpendale, 2010). Following this definition, we dye the area within a 0.6-metre radius of the participants' locations (locations here are at 1-second interval for higher accuracy) with different tint colors (red for participant A's personal territory, and blue for B's) to indicate territories. We chose 0.6 metres as it falls into the range of close phase of personal distance, which permits one participant to touch each other or the same music interface (Hall, 1966), most of the musical note edits also fell inside this range.

(1) Distribution of locations and interactions. The redder/bluer the area is, the more presence the corresponding participant had shown in that location. The overlap is coloured grey, indicating appearances

Table 3. Statistics and Wilcoxon Rank Sum Test (two-tailed) of Activity Assessments (AA)

Measure	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)*
AA1 - 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)
AA2 - 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)
AA3 - 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)
AA4 - No. of mutual note modifications[a]										
	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)
AA5 - 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)
AA6 - 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)
AA7 - 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)
AA8 - 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)
AA9 - 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)
AA10 - 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)
AA11 - No. of uses of personal spaces										
	-	-	2.40	2.85	-	-	-	-	-	0.2912
	(-)	(-)	(1.95)	(2.14)	(-)	(-)	(-)	(-)	(-)	(1193)
AA12 - Length of time of using personal spaces (unit: second)										
	-	-	128.60	112.19	-	-	-	-	-	0.4685
	(-)	(-)	(86.95)	(78.67)	(-)	(-)	(-)	(-)	(-)	(1464)
AA13 - Average duration of each entry of personal space (unit: second)[b]										
	-	-	73.07	49.54	-	-	-	-	-	0.008019
	(-)	(-)	(56.55)	(44.83)	(-)	(-)	(-)	(-)	(-)	(1512)
AA14 - No. of note edits in public space										
	-	-	54.48	43.69	-	-	-	-	-	0.5051
	(-)	(-)	(48.69)	(34.96)	(-)	(-)	(-)	(-)	(-)	(1455)
AA15 - No. of note edits in personal space										
	-	-	43.87	34	-	-	-	-	-	0.3869
	(-)	(-)	(40.10)	(25.37)	(-)	(-)	(-)	(-)	(-)	(1485.5)
AA16 - Time spent paying close attention to collaborator (unit: second)[c]										
	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)
AA17 - Times of paying close attention to collaborator[c]										
	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)
AA18 - Time spent paying ordinary attention to collaborator (unit: second)[c]										
	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)
AA19 - Times of paying ordinary attention to collaborator[c]										
	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)

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

[b] 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.

[c] 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 m (depth), whilst FOV of ordinary attention roughly covers 27 degrees (horizontally), 28 degrees (vertically) and 2.7 m (depth).

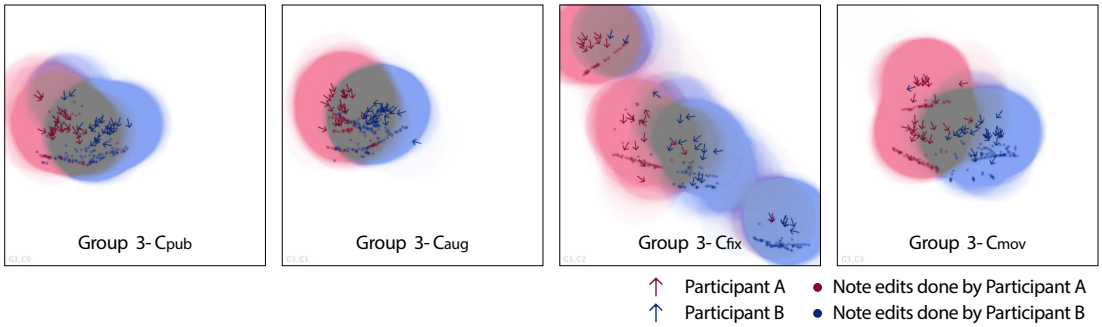

Figure 7. Visual traces of the participants' locations, directions and musical note edits (group 3).

of both participants, this can be seen as group territory, see the traces of Group 3 in Figure 7 as an example, traces of all groups are shown in Figure 8.

(2) 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}. 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}.

Participants had an average of 36.44 group edits in C_{pub}, which is significantly more than that of C_{fix} ($M = 17.50$; $W = 1837$, $p = 0.001448$), and a near-marginal significantly more than that of C_{PI} ($M = 25.23$, $W = 1621.5$, $p = 0.07839$). C_{aug} resulted in a higher average of group edits ($M = 43.04$), though not significantly higher than C_{pub}, it is significantly higher the C_{fix} and C_{mov} (both $p < 0.001$). These indicate the spatial configurations of C_{pub} and C_{aug} are more friendly to group edits.

(3) Sizes of personal territory and personal edits (edits fallen into personal territory). Participant formed a significantly larger personal territory in C_{fix} ($M = 0.7475$ m^2, $SD = 0.1801$) compared with all the other three conditions (Wilcoxson Rank Sum Test, all $p < 0.001$), and had significantly more personal edits in C_{fix} compared with other conditions (all $p < 0.001$). 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 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 independently, whilst C_{pub} and C_{aug}, on the opposite, led to more interactions in the group territory.

(4) Average distance. 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 to each other compared with C_{fix}.

Times and amount of use of personal space

In C_{fix}, participants had an average of 2.40 entries of personal space, each entry on average lasting 73.07 seconds, whilst the average length of time for staying inside personal space is 128.60 seconds. For C_{mov}, the participants did 2.85 entries on average, with a total usage time of 112.19 seconds, No significant difference was found in the number of entries or the usage time. In C_{mov}, the average duration of each entry is 49.54 seconds, which is significantly shorter than that of C_{fix} (Wilcoxon Rank Sum Test, $W = 1512$, $p = 0.008019$), indicating that personal spaces in C_{mov} were used slightly more frequently (mean of C_{mov} is higher) whilst personal spaces in C_{fix} were used for a longer independent creation.

Attention

(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 that of 7.19 seconds in C_{pub}, and 7.31 seconds in C_{fix} (Wilcoxon Rank Sum

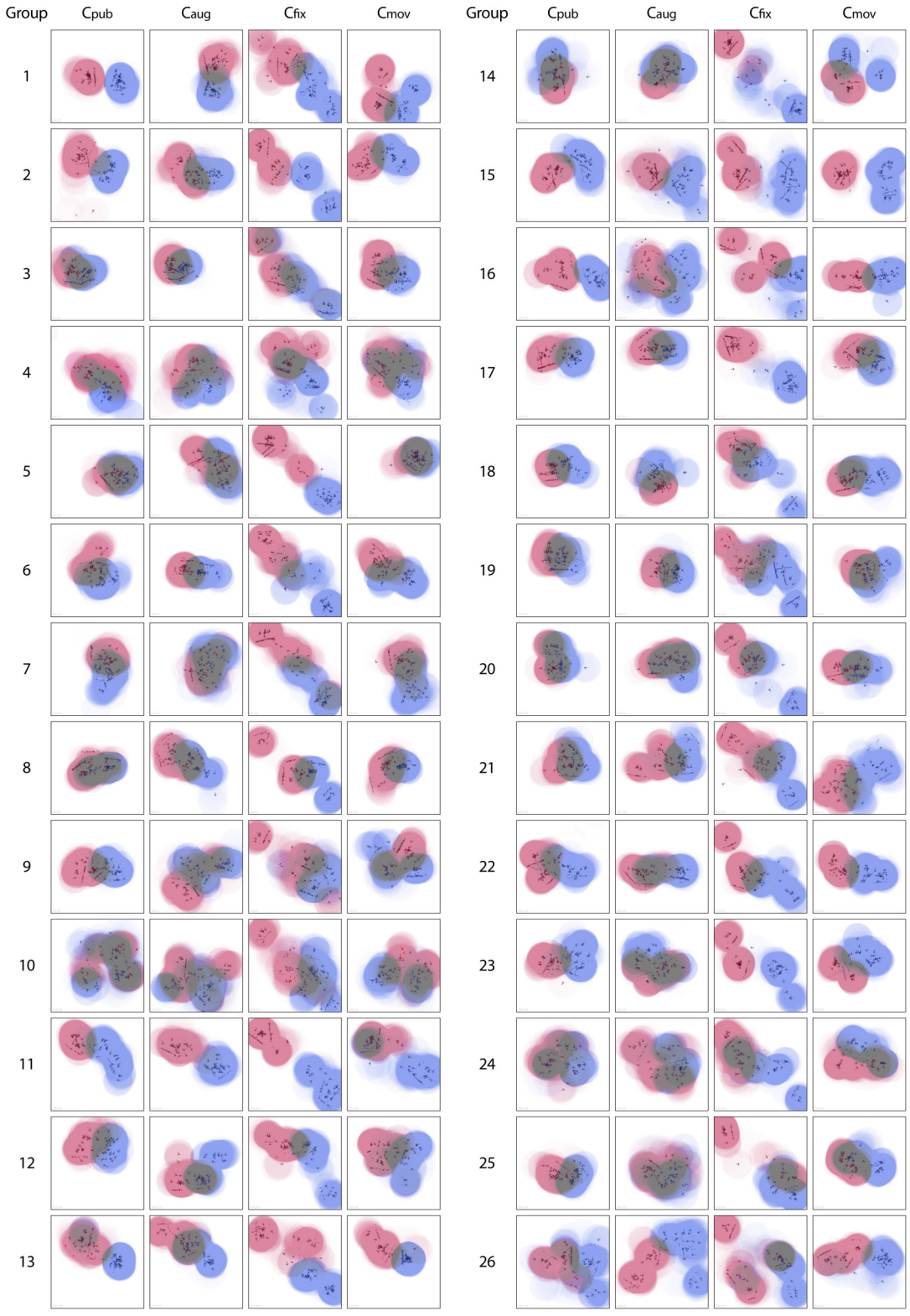

Figure 8. Visual traces of the participants' locations, directions and musical note edits (all groups).

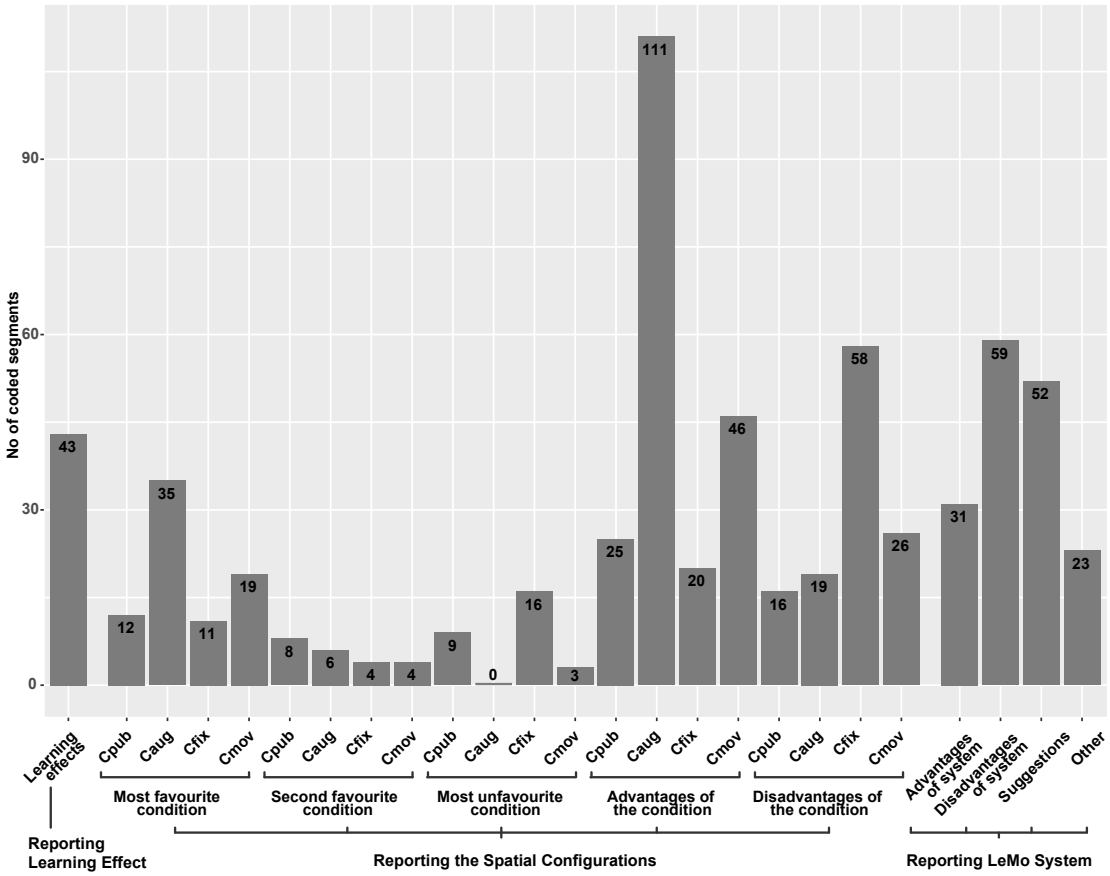

Figure 9. Ingredients of all the coded-segments of the interview, numbers of coded segments are shown along the bars.

Test, both $p < 0.05$), and near-marginal significantly longer than 11.02 seconds in C_{mov} ($W = 1645$, $p = 0.05722$, see AA16 in Table 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} (Wilcoxson 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}: $W = 1683.5$, $p = 0.03122$). Wilcoxson 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 AA17 in Table 3. These results indicate the spatial configuration in C_{aug} significantly promoted participants to pay more close attention to their collaborator, whilst C_{mov} promotes insignificantly and C_{fix} demotes 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} changes the way participants pay their ordinary attention, all at around 70 to 80 seconds out of the 420-second session. C_{fix} greatly reduced participant's ordinary attention paid to each other, they only had an average of 52.89 seconds 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} ($W = 1081$, $p = 0.07865$), more details in AA18 of Table 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 (Wilcoxson Rank Sum Test, all $p < 0.001$), check detailed statistics in AA19 of Table 3), indicating the spatial configuration of C_{fix} greatly reduced participants' paying ordinary attention to each other.

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 and Clarke, 2006; Yin, 2017). The starting point of the thematic analysis was a reading through of the transcript, then we did an inductive analysis of the data, 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 9, in total, 656 coded segments, 24 codes and 3 overarching themes emerged from the thematic analysis:

Learning effects

Members of 18 groups mentioned the effect of the session sequence. Specifically, 43 coded segments contributed by 27 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}), as 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 participants Group 11 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 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 exemplars are: "a bit troubling" (P_{11B}), "music always very loud" (P_{9A}), "it was global music, and there was someone annoying" (P_{2A}), "you are not going to say anything" because possibly make you to be "rude"(P_{2A}). It was easier if there is something helping me "to perceive what I was doing, and not get confused with what [my collaborator] was doing"(P_{15B}), it was too "chaotic" (P_{20A}), "too confusing" (P_{22A}) & P_{22B}), "annoying"(P_{25B}). "cannot concentrate"(P_{25B}) "everything [is] open and quite noisy" (P_{26B}), "don't have the tranquillity to operating your sounds or the everything's come mixed, which is difficult to manage" (P_{22A}).

There were 25 coded segments from 14 participants reporting the positive side of the C_{pub}, some exemplary points 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, as reported by P4A, saying that they needed someone to lead her 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 chose 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 with "personal space" was provided (P_{3B}, P_{18A} & P_{18B}).

(2) **C_{aug} - Overwhelming preference.** There were 35 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 participants C_{fix} and 19 segments contributed by 17 participants for C_{mov} (note sum of participants is who contributed is greater than 52 as a few participant reported more than one favourite conditions during the process of 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} "forces [them] to collaborate more the most because [they] had to stay very close to composing music "(P_{9B}). (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 [the participant] 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" but also being able 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 the [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 sound 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 you want to hear something, you just come closer, like in the real world" (P_{11B} & P_{11B}), "it was good like we were 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 include 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) without 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, common advantages and disadvantages of these two conditions were reported in the interview. Common advantages include: 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 "think about something to add "(P_{9A}), which helped "develop their own ideas" (P_{8A}), and as a result, they "used a lot personal space" (P_{3A} & P_{3B}) 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}), made them "forget" the collaboration/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 each other's music pieces did not fit together when brought up. (P_{4A}) reported they were not familiar with music, and thus they "needed somebody to lead" them, so prefer ed to hear sounds all the time. It was also reported the visual personal space made the stages look "messy" (P_{24B}).

Differences between C_{fix} and C_{mov} - In total, 46 coded segments (from 26 participants) were reporting C_{mov}'s advantages and 27 segments (from 12 participants) reporting its disadvantages, compared with 22 segments (from 14 participants) and 58 segments (from 33 participants) for C_{fix}, indicating in general participants thought C_{mov} was more superseded than C_{fix}. Some exemplary 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 as "the boundary is small" (P_{7B} & P_{7A}). Besides, the relative far distance also helped to "prevent [... them...] from clashing" (P_{7A}).

Reporting LeMo system

Participants reported where they felt right (19 coded segments) and wrong (59 coded segments) with the design and technical part of LeMo, and suggestions were given (51 coded segments). Since these are not directly related to the scope of this paper, these segments will not be detailed here.

DISCUSSION

Based on the results, next we discuss the necessity, and impacts of adding personal space, and specifically compare the differences of adding personal space with/without mobility, and personal space with rigid/fluid boundary.

Necessity of adding personal space

When no personal space was available, C_{pub} was reported to provide the experience of the least difficulty of tracking the collaborator (CQ3 in Table 2), the strongest sense of collaborator (CQ4), best communication quality (CQ5), the least difficulty to cooperate (CQ9). So C_{pub} seems to be the simplest one among these four configurations for participants to learn and get used to. However, the issues of having no personal space are clear. Firstly, especially for the music making task in this study, 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 is crucial. C_{aug}, C_{fix}, C_{mov} solved this problem by providing different solutions. The personal space functioned like a "less stressing" context, within which, they 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", promoting their own creativity, which can then be combined and contributed to the joint piece, which matches the findings in (Men and Bryan-Kinns, 2019), 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 were also reported, e.g. less communication, higher isolation and being messy, most of these limitations are the results of introducing rigid visible personal space, and C_{aug} has addressed these limitations well details will be discussed in later subsections). Next, we discuss the impacts of introducing each of these personal spaces individually.

Impacts of adding personal space

As mentioned above, in the previous study (Men and Bryan-Kinns, 2019), we found 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, fewer times of paying attention to collaborators. We argued these negative impacts are mainly due to that the personal spaces distributed on the opposite side of the group space resulting in a larger distance between participants. So we proposed personal space with different features (e.g. gradual boundary - C_{aug}, mobility - C_{mov}) might reduce these negative effects, below we discuss how these negative effects might be eliminated by these three conditions C_{aug}, C_{fix}, C_{mov} and the impacts of introducing personal spaces in these conditions.

Invisible auditory personal space in C_{aug}

In many ways, C_{aug} is quite similar to C_{pub}, e.g. both do not have a visual boundary for spaces, no triggers to trigger personal space, participants formed a similar pattern in these two conditions (Figure 7). So not surprisingly, no significant differences were found in most of the statistical measures, see Table 1 and 3. The only differences revealed by these Tables are the significant differences found in AA16, AA17 (paying more close attention to collaborator in C_{aug} is even higher than in C_{pub} and a marginal-significant difference in PSQ7 (sense of collaborator's activity is higher in C_{pub} than in C_{aug}.

From another perspective, fewer differences between C_{pub} and C_{aug} indicate the limitations of adding personal space identified in previous work (Men and Bryan-Kinns, 2019) have been successfully minimized. Specifically, the size of group territory and number of group edits maintained similar numbers, in C_{aug} the means are even greater, though not significantly (AA5 and AA7 in Table 3). C_{pub} and C_{aug} saw a similar size of personal territory, personal edits, average distance (respectively, AA6, AA8, AA10 in Table 3), and C_{aug} even saw more close attention paid to each other (AA16 and AA17 of Table 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. Reasons can be that C_{aug} managed to provide a similar interaction experience, in the previous study, personal spaces were located at the opposite side of the public space, participants had to drift apart and walk up and down to use it, this changed the participants' spatial locations. As a result, the group/personal territoriality 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, we managed to provide a maximized-similar user experience with C_{pub}. The second reason is more about the impacts on subjective experience, by making the personal space invisible and gradual, the isolation and difficulty of coordinating that introduced by the additional personal space was minimized. 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.

Movable personal space in C_{mov}

In C_{mov}, participants could trigger and make the personal space appear everywhere in the stage. In this way, a personal space with mobility was provided, and by doing so, some aforementioned negative effects found in (Men and Bryan-Kinns, 2019) were reduced, specifically, these eliminated differences include size of group territory, the average distance have been eliminated, times of paying attention to collaborator (AA5, AA10, AA16, AA17 in Table3). However, some significant differences remained, participants still had fewer mutual note modifications, fewer group edits (though marginal-significantly different) and more personal edits after personal space being introduced in C_{fix} and C_{mov} (see AA4, AA7, and AA8 in Table 3). This is also verified from the result of the thematic analysis of the interview, compared with C_{aug} participants reported a higher sense of isolation in C_{mov} and C_{fix} in which rigid- form personal spaces were provided. In other words, by making the personal space available anywhere in the stage, though we managed to drag participants closer, formed a similar group territory, their behaviour was still affected in many ways, they were still separated to some extent, which can be seen as a disadvantage of adding visible, solid personal space. In other words, C_{aug} minimized the negative impacts of the introduction of personal space better than C_{mov}.

A more rigid personal space in C_{fix}

C_{fix} provided a much more inflexible personal space, it changed participants' behaviour in many ways (see the significant differences between C_{pub} and C_{fix} in Table 1 and Table 3). Not to mention in CQ (Table 2), participants rated these two conditions to be the opposite side. E.g. significantly many participants rated C_{pub} is the session they had the strongest sense of their collaborator's presence. C_{fix} was chosen by significantly many participants to be the one in which they had the least sense of collaborator's presence. Similarly, significantly many participants believed tracking collaborator's activities, best communication quality, least difficult to cooperate with collaborator happened in C_{pub}, whilst C_{fix} was thought conversely by significantly many participants. These can also be verified by the interview, in which only 10 participants reported C_{fix} as their favourite condition and 52 coded segments contributed by 33 participants were reporting the disadvantages of C_{fix}.

Providing personal space with fluid boundary

Measures in Table 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 (AA6) and a bigger group territory (AA5), more mutual modifications (AA4), more group edits (AA7) and fewer personal edits (AA8), a larger distance between collaborators (AA10), more times of paying close attention (AA17) and a longer time of paying close attention (AA16). All these indicate that compared with the rigid personal space in C_{fix} and C_{mov}, the augmented sound attenuation in C_{aug} enabled a closer collaboration, H3 is supported. Its advantages are shown in three ways, next, each will be specified.

Enough support for creativity with minimal impacts

PSQ2 (Table 1) questioned the support each condition gave to individual creativity. Although no significant differences were found, C_{aug} has a higher mean rating. The thematic analysis revealed more insights, C_{aug} provides 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". No major differences were found between C_{pub} and C_{aug}, indicating C_{aug} provides a very mild solution, with limited impacts on people's collaborative behaviour introduced. Whilst still C_{aug} providing sufficient support for individual creativity during collaboration, thus H2 is validated.

Closer collaboration and higher consistency

According to Table 3, in condition C_{aug}, participants paid much more close attention to their collaborator compared with what they did in other conditions. Reasons can be found from the thematic analysis and the measures in Activity Assessments (Table 3). Compared with natural attenuation in C_{pub}, C_{aug}'s augmented

sound 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 in a continuous way" if they want to hear other's work, an invisible gradual boundary in C_{aug} led to less separation, and higher consistency between personal and public space. Compared with rigid personal space 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 (see Table 3), all of these indicate that C_{aug} saw a less separated collaboration than C_{fix} and C_{mov}, H3 is therefore supported. This finding also echos the suggestion that boundary between personal and public space should be provided with gradations in subtle and lightweight ways (Greenberg et al., 1999) to enable a fluid shift.

Popularity

The code "advantage of C_{aug}" have 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. All indicate C_{aug} is the most popular condition. The popularity is also partially verified by that preference measure in PSQ3 (mean of C_{aug} is the highest) and CQ1 (more participants chose C_{aug} as the most joyful setting). We believe the reasons behind this popularity is mainly 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.

Providing personal space without/with mobility

In this subsection we compare C_{fix} with C_{mov}. The clear, sole difference between these two conditions is the mobility of personal space. In C_{fix}, to use personal space at the corners, participants needed to walk to the corner, this might be the reason that C_{mov} saw a closer average distance between collaborators than C_{fix} (AA10, Table 3). A greater distance in C_{fix} resulted in a significantly larger size of personal territories (AA6) and personal edits (AA8). On the contrary, the closer distance in C_{mov} created more chances for paying attention to each other or drawing attention and, as a result, significantly more time was spent paying attention to collaborators (AA17, AA19, 3). With a closer average distance and more attention paid to each other, participants reported they had a better quality of communication in C_{mov} (see the marginal significant difference in PSQ6, Table 1. On the other hand, with participants being far away from each other and less contact, significantly many reported that they had the worst communication quality in C_{fix} (CQ6 Table 2). C_{mov} was also rated to be more enjoyable (CQ1 Table 2, significantly many participants chose C_{fix} as the least enjoyable). C_{fix} also led to a reduced sense of collaborator's contribution (CQ9 in Table 3). As a result of having a better quality of collaboration in C_{mov}, a significantly more satisfying work output was produced (PSQ5, Table 1). Thematic analysis results also echo these findings. There are much more coded segments reporting C_{mov}'s advantages compared with those reporting C_{fix}'s, and there are much more coded segments reporting C_{fix}'s disadvantages than those reporting C_{mov}'s. Also, more participants (19) reported their favourite setting was C_{mov} compared with 11 choosing C_{fix}. Participants reported being able to use personal space anywhere in the stage with the personal space is good as it resulted in a closer distance, which led to more collaboration made it possible to see each other's work, it functioned like a mute button.

To conclude, compared with C_{fix}, C_{mov} resulted in better communication quality, produced better feeling of collaborator's contribution, and was rated more enjoyable, thus saw a closer collaboration produced a more satisfying result, H1 is therefore supported.

Key Findings

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

- Having personal space is suggested as it supports individual creativity, which is an important element of the collaborative creativity.
- C_{aug} minimized the negative impacts introduced by adding personal space (previously identified by Men and Bryan-Kinns, 2019) 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 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 attention was paid between collaborators).

- Compared with stationary personal space, movable personal space 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 it supported collaboration better than the personal space with stationary personal space.

DESIGN IMPLICATIONS

Based on the key points made above, we suggest three design implications for SVEs focusing on supporting collaborative task:

(1) SVEs supporting creative collaborative tasks 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 essential when the output of the task is more disruptive (e.g. audio), co-workers need a space where they can think of and develop own mind and work.

(2) For audio-related task (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 is also light-weight, functions as a personal space well, and can increase close attention paid between participants. We did not find it introduces significantly negative impacts whereas we found rigid personal space did.

(3) Beyond audio-related tasks, when providing personal space in SVEs, lightweight free-form personal space rather than personal space with rigid form should still be firstly considered, as it introduces fewer negative impacts on collaboration and enables a fluid shift, which matches the findings of Greenberg et al. (1999). The light-weight form is not limited to audio, it can be one of other modalities (e.g. visual) or multiple modalities. For example, in this study, augmented attenuation in sound has been verified to provide a useful personal space for CMM in SVEs Similarly, 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 modals, an exemplary task can be making a short animation and creating an accompanying music track for it.

(4) Manipulating the level of augmentation (e.g. the augmented acoustic attenuation in this study) can change the level of feeling personal. In the C_{aug} condition of this study, participant adjusted their distance between themselves and collaborators to obtain a different level of being personal (herein referred as "personalness"), e.g. total isolation can be achieved if both participants are working with a distance greater than 1.2 metres. We believe similarly, when since personal spaces are with gradual 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 if an extremely minimal impact 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/see each other's work), and having no personal space when they have to work together. In this way, users can be 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 is also useful as co-wrokers can stay at any distance whilst still being able to adjust the "personalness" the personal space provides.

(5) When it is hard or impossible to design a gradual, light-weight personal space that applicable to the task due to the type of the task, a rigid-form personal has to be used if providing personal space is wanted. In this situation, it is better to provide rigid personal space with mobility, as it gives users more freedom for accessing the personal spaces, and produces a better user experience with fewer negative impacts

on the collaboration compared with personal space without mobility. This implication also matches the proposal rasied in our previous work (Men and Bryan-Kinns, 2019).

CONCLUSIONS

In this article, we have briefed 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, differences were found and five design implications were given. Specifically, we found 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). We also found that a rigid personal space with mobility serves users' needs better and is preferable over a non-mobile one.

In the future, we are keen to explore how to design and apply personal spaces with fluid boundaries in s wider range of creative scenarios in SVEs, e.g. for collaborative drawing in an SVE, 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 collaborators perceive them. We are also interested in how the boundary might be manipulated and whether the manipulation can result in different impacts on the collaborative behaviour.

ACKNOWLEDGMENTS

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

REFERENCES

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

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

Basdogan, C., Ho, C.-H., Srinivasan, M. A., and Slater, M. (2000). An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(4):443–460.

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

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

Billinghurst, M., Poupyrev, I., Kato, H., and May, R. (2000). Mixing realities in shared space: An augmented reality interface for collaborative computing. In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, volume 3, pages 1641–1644. IEEE.

Blaine, T. and Fels, S. (2003). Contexts of collaborative musical experiences. In *Proceedings of the 2003 conference on New interfaces for musical expression*, pages 129–134. National University of Singapore.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101.

Bressan, F., Vets, T., and 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*, pages 23–26.

Brooks Jr, F., Marcus Brown, U., Burbeck, T. C., Durlach, N., Ellis, M. S., Lackner, J., Robinett, W., Srinivasan, M., Sutherland, M. I., Urban, D., et al. (1992). Research directions in virtual environments. *Computer Graphics*, 26(3):153.

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*, pages 135–144. ACM.

Bryan-Kinns, N. (2013). Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies*, 71(1):76–90.

Bryan-Kinns, N. and Hamilton, F. (2012). Identifying mutual engagement. *Behaviour & Information Technology*, 31(2):101–125.

Bryan-Kinns, N., Healey, P. G., and Leach, J. (2007). Exploring mutual engagement in creative collaborations. In *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, pages 223–232. ACM.

Carlsson, C. and Hagsand, O. (1993). Dive a multi-user virtual reality system. In *Proceedings of IEEE Virtual Reality Annual International Symposium*, pages 394–400. IEEE.

Fencott, R. and Bryan-Kinns, N. (2010). Hey man, you're invading my personal space! privacy and awareness in collaborative music. In *NIME*, pages 198–203. Citeseer.

Gaver, W. W. (1992). The affordances of media spaces for collaboration. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pages 17–24. ACM.

Greenberg, S., Boyle, M., and LaBerge, J. (1999). Pdas and shared public displays: Making personal information public, and public information personal. *Personal Technologies*, 3(1-2):54–64.

Hall, E. T. (1966). *The hidden dimension*, volume 609. Garden City, NY: Doubleday.

Hendrix, C. and Barfield, W. (1996). The sense of presence within auditory virtual environments. *Presence: Teleoperators & Virtual Environments*, 5(3):290–301.

Kreijns, K., Kirschner, P. A., and 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.

Lea, R., Honda, Y., and Matsuda, K. (1997). Virtual society: Collaboration in 3d spaces on the internet. *Computer Supported Cooperative Work (CSCW)*, 6(2):227–250.

Luciani, A. (2007). Virtual reality and virtual environment. In *Enaction and enactive interfaces : a handbook of terms*, pages 299–300. Enactive Systems Book.

Mead, L. and Malcomson, I. (2003). Dungeons & dragons faq. *Wizards of the Coast. Retrieved October*, 3:2008.

Men, L. and Bryan-Kinns, N. (2018). Lemo: Supporting collaborative music making in virtual reality. In *2018 IEEE 4th VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, pages 1–6. IEEE.

Men, L. and Bryan-Kinns, N. (2019). Lemo: Exploring virtual space for collaborative creativity. In *Proceedings of ACM conference on Creativity and cognition 2019*. ACM.

Men, L., Bryan-Kinns, N., Hassard, A. S., and Ma, Z. (2017). The impact of transitions on user experience in virtual reality. In *2017 IEEE Virtual Reality (VR)*, pages 285–286. IEEE.

Nassiri, N., Powell, N., and Moore, D. (2010). Human interactions and personal space in collaborative virtual environments. *Virtual reality*, 14(4):229–240.

Nedel, L., de Souza, V. C., Menin, A., Sebben, L., Oliveira, J., Faria, F., and Maciel, A. (2016). Using immersive virtual reality to reduce work accidents in developing countries. *IEEE computer graphics and applications*, 36(2):36–46.

Oculus (2015). Toybox demo for oculus touch. Accessed: 2019-03-26.

Pinelle, D., Gutwin, C., and 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.

Plante, C. (2016). A playstation VR demo made me cry from laughing. Accessed: 2019-03-26.

Raffestin, C. (2012). Space, territory, and territoriality. *Environment and Planning D: Society and Space*, 30(1):121–141.

Rodden, T. (1991). A survey of CSCW systems. *Interacting with computers*, 3(3):319–353.

Roussos, M., Johnson, A. E., Leigh, J., Vasilakis, C. A., Barnes, C. R., and 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.

Sack, R. D. (1983). Human territoriality: a theory. *Annals of the association of American geographers*, 73(1):55–74.

Schroeder, R. (2012). *The social life of avatars: Presence and interaction in shared virtual environments*. Springer Science & Business Media.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281.

Scott, S. D., Carpendale, M. S. T., and Inkpen, K. M. (2004). Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*,

pages 294–303. ACM.
Scott, S. D. and Carpendale, S. (2010). Theory of tabletop territoriality. In *Tabletops-horizontal interactive displays*, pages 357–385. Springer.
Shen, C., Everitt, K., and Ryall, K. (2003). Ubitable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *International Conference on Ubiquitous Computing*, pages 281–288. Springer.
Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93.
Sugimoto, M., Hosoi, K., and 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*, pages 41–48. ACM.
Tang, J. C. (1991). Findings from observational studies of collaborative work. *International Journal of Man-machine studies*, 34(2):143–160.
Titon, J. T. and 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*.
Turchet, L., Fischione, C., Essl, G., Keller, D., and Barthet, M. (2018). Internet of musical things: Vision and challenges. *IEEE Access*, 6:61994–62017.
Wang, G. (2009). Designing smule's ocarina: The iphone's magic flute. In *NIME*, pages 303–307.
Xambó, A., Hornecker, E., Marshall, P., Jordà, S., Dobbyn, C., and 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.
Yin, R. K. (2017). *Case study research and applications: Design and methods*. Sage publications.
Zhang, X. and Furnas, G. W. (2003). The effectiveness of multiscale collaboration in virtual environments. In *CHI'03 Extended Abstracts on Human Factors in Computing Systems*, pages 790–791. ACM.