

Designing spaces to support collaborative creativity in shared virtual environments

Liang Men, Nick Bryan-Kinns and Louise Bryce

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

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.

Subjects Human-Computer Interaction, Social Computing

Keywords Collaborative creativity, Virtual reality, Shared Virtual Environment, Collaborative music making, Sonic interaction design, Gesture design, HCI

INTRODUCTION

The real world envelops us with space that we share with others; 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 are digital games, e.g., *Star Trek* created in early 1970s provides a computational space that players can visit and experience through text descriptions on a computer screen, see (*Case, Ploog & Fantino, 1990*). Though these non-immersive media can involve people to a very high level and generate the experience of flow, few of them have enabled people to interact in a natural way that is similar to the way that people

Submitted 11 June 2019

Accepted 24 September 2019

Published 4 November 2019

Corresponding author

Liang Men, l.men@qmul.ac.uk

Academic editor

John Carroll

Additional Information and
Declarations can be found on
page 35

DOI 10.7717/peerj-cs.229

© Copyright
2019 Men et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

experience real-world interactions. For instance, the interactions based on keyboards and mouse for input and monitor for output have very different properties to real-world interactions (Gaver, 1992). In contrast, virtual reality (VR) provides a novel space for multisensory experiences (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. 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 utilisation 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 will 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 (Bishop & Fuchs, 1992), it emerged as a competing term to virtual reality; 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, prompting the shared virtual environments (SVEs; Schroeder, 2002). Besides “SVEs”, other terms used include multi-user virtual environments, multi-user virtual reality (Carlsson & Hagsand, 1993), collaborative virtual environments (CVEs; Zhang & Furnas 2003) and social virtual reality (SVR). To align with mainstream usage, 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, 2002). 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, Powell & Moore, 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

¹Toybox Demo for Oculus Touch: <https://www.youtube.com/watch?v=iFEMiyGMA58> (Accessed: 2019-09-03).

²A PlayStation VR demo: <https://www.theverge.com/2016/3/16/11246334/playstation-virtual-reality-social-vr-demo> (Accessed: 2019-09-03).

community activities (Lea, Honda & Matsuda, 1997), online education (Roussos et al., 1997), distributed work and training (Nedel et al., 2016), and gaming and entertainment (see Toybox¹ and a PlayStation VR demo²) Despite this, little research exist in the field of supporting collaborative creativity, leaving an open question: how should shared virtual environments 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 question. “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, Kirschner & Jochems, 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 & 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 physical virtual space rather than the space concept 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 & Bryan-Kinns, 2010; Men & 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 is important (Greenberg, Boyle & LaBerge, 1999; Sugimoto, Hosoi & Hashizume, 2004). As a starting point of this exploration, Greenberg, Boyle & LaBerge (1999) developed a PDA-based prototype named SharedNotes. 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, Shen, Everitt & Ryall (2003) addressed this concern in their project UbiTable by providing a flexible gradient of sharing semantics. Specifically, rather than the binary notion of public and private space, UbiTable provides an additional semi-private space, in which data is visible but not electronically accessible for others. However, both sets of research were carried out based on 2D media (PDA and projector respectively), which made their findings less informative for VR. Moreover, neither of these two explored more gradations between public space and personal space, i.e., the transition between the public space and personal space is step-less and smoother. In this study, we want to design virtual spatial configurations that provide a more gradual boundary between personal and public spaces and see if this can enable a fluid shift and provide better support for collaboration in VR.

Territory and territoriality in collaboration (SVEs and 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 personal territories (Men & Bryan-Kinns, 2019). By manipulating the virtual spatial configurations of an SVE, the formation of territory and territorial behaviour can be influenced and, as a result, the collaboration can be influenced (Men & Bryan-Kinns, 2019). Because there is limited research on territoriality in VR, and rich research on this in tabletop-based collaboration, we review territoriality in tabletop research as a supplement, which might be informative as it is also a computer-mediated collaboration that evolves territory. The term “tabletop” here refers to interactive tabletop *displays*, these usually include high quality projectors, flat panel or plasma displays, and touch-sensitive surfaces (Kruger et al., 2004). These electronic tabletops inherit the collaborative benefits of tables, which greatly compensate computers’ disadvantages in this regard (Scott, Carpendale & Inkpen, 2004). Similarly, territoriality also plays an important role in the tabletop collaboration. Collaborators were found to use different types of territory to serve different needs, including sharing, exchanging or storing working tools and resources (Scott, Carpendale & Inkpen, 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 & 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, Carpendale & Inkpen, 2004).

(2) *Group territory* for performing the main task. In group territory, people create and develop new solutions, transfer resources and provide help (Scott, Carpendale & Inkpen, 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, Carpendale & Inkpen (2004) proposed four guidelines for designing digital tabletop workspaces: (i) visibility of action; (ii) an appropriate size of workspace; (iii) providing functionality in the appropriate locality; (iv) allowing for the grouping of items to facilitate storage. Furthermore, the visibility and transparency of actions have been found to be important in designing group workspaces, as they help collaborators to monitor each others’ actions, maintaining workspace awareness during collaboration (Pinelle, Gutwin & Greenberg, 2003; Fencott & Bryan-Kinns, 2010). However, this can result in overloaded cognitive information, which some people found to be difficult to handle (Fencott & Bryan-Kinns, 2010). To date, little research has explored how such features of workspace might be designed for collaboration 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 & Hamilton, 2012; Titon & Slobin, 1996*). Its unique features make it an excellent activity through which to study collaborative activity. In 2003, *Blaine & 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, Vets & Leman, 2017*), both of which provide co-located music-making experience, and Ocarina (*Wang, 2009*), which provides a distributed experience. However, we should note that, despite decades of research into CMMs and SVEs, relatively few 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., how shared virtual environments should be designed to support creative collaboration, such as CMM.

Acoustic attenuation

Sound attenuates as a result of diminishing intensity when travelling through a medium. This feature of sounds enables humans to use their innate spatial abilities to retrieve and localise information and to aid performance (cf. *Billinghurst & Kato 2002*). Whilst it is hard to adjust the acoustic attenuation of a real medium (e.g., the air) to enhance its potential, within VR, as the audio is simulated, we can simulate an augmented spatialised sound purposely. Research has been done on investigating the impacts of spatialised sounds on user experience in VR, e.g., (*Hendrix & Barfield, 1996*). However, little research explores how the spatialisation of sound may affect or aid collaboration in SVEs (e.g., CMM). Considering sound is both the primary medium and the final output of the creative task CMM (*Men & 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 may then be used as personal space supporting individual creativity in CMM.

LeMo—An SVE for collaborative music making

We created Let's Move (LeMo³), which enables two users to manipulate virtual music interfaces together in an SVE to create a 16-beat music loop, see [Fig. 1](#). LeMo used in this study and *Men & Bryan-Kinns (2019)* is an extensively modified version of the previous version (*Men & Bryan-Kinns, 2018*). Major differences include a new visual

³Full source available at: <https://sites.google.com/view/liangmen/projects/LeMo>

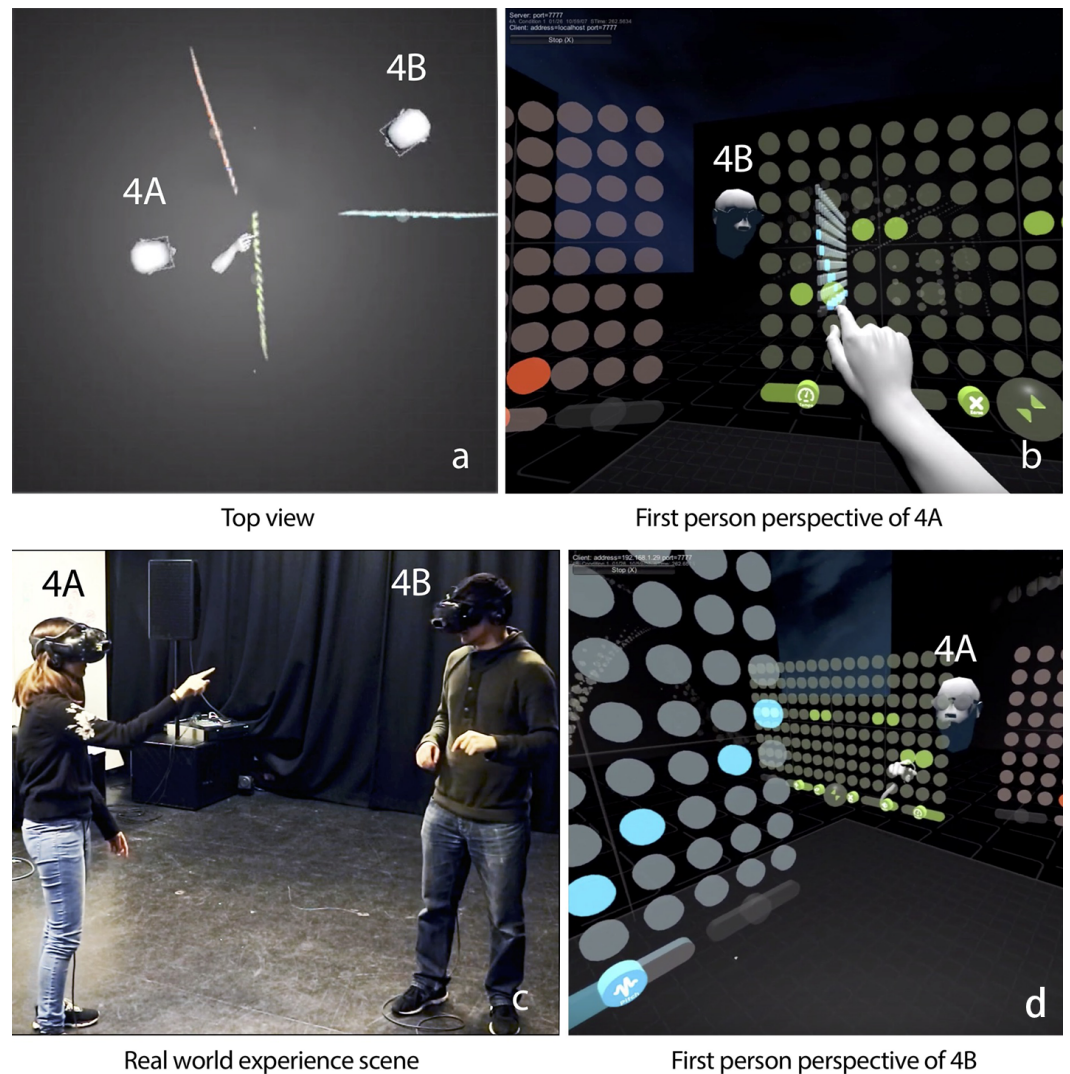


Figure 1 Participant 4A and 4B are creating music together.

[Full-size](#) [DOI: 10.7717/peerjcs.229/fig-1](https://doi.org/10.7717/peerjcs.229/fig-1)

presentation of the interface, the ability to add, delete, and re-position music loops, with more freedom in terms of instruments, tempo, volume and pitch. These changes are made to improve the experience of music making and, more importantly, to enable users to re-position and share music loops, which is essential for this study. 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 a Leap Motion mounted, see Fig. 1C). The position and rotation of headsets are tracked by two tracking cameras set

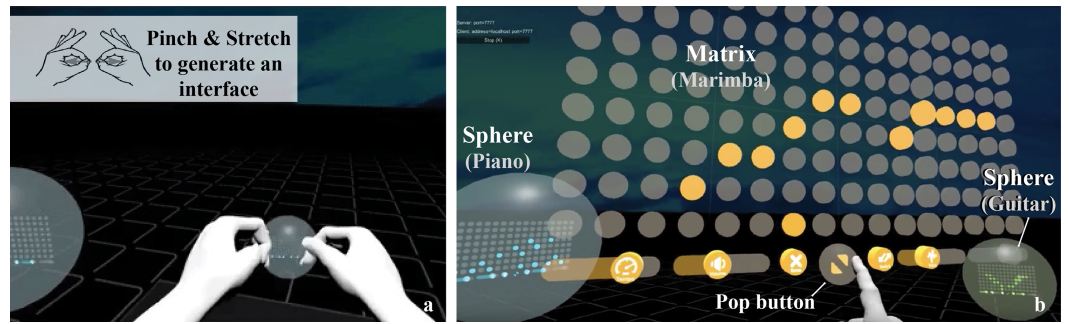


Figure 2 (A) The gesture to generate a new interface; (B) *Matrix* (opened interface) and *sphere* (packed interface), double click the pop button to switch in between (reproduced from Men & Bryan-Kinns 2019).

Full-size DOI: [10.7717/peerjcs.229/fig-2](https://doi.org/10.7717/peerjcs.229/fig-2)

around the scene, and hand gestures are tracked by the mounted Leap Motion. Two PCs are used to run LeMo, these are 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* (Fig. 2B). Users can generate up to eight *spheres* with pinch and stretch gesture, see Fig. 2A. The music interface can be switched between the two modes, re-positioned or removed by manipulating the *sphere* or the pop button of the *matrix* with corresponding gestures. As shown in Fig. 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. Users can edit notes by tapping 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 (Fig. 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 synchronising the virtual environment (virtual space and music interfaces) and avatars across a network, providing participants with the sense of being in the same virtual environment and manipulating the same set of interfaces. (3) A virtual space that includes a grey stage with a grid pattern (part of it is shown in Fig. 1D). Four types of spatial configurations were designed for this study, which will be detailed later.

Beside 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 min left” and “end of session” notifications; A data-log system based on trackers (HTC Vives and Leap Motion) to log time-stamped data from events generated by users' interactions and movements, including positions and rotations of heads, index fingers, thumb fingers, the manipulation with musical interface (addition/deletion/re-positioning

gradual rather than too rigid (*Greenberg, Boyle & LaBerge, 1999*), we think the attenuation feature can be applied to form a gradual personal space and enable a fluid transition between personal space and public space. This is because the sound is both the primary medium of collaborative tasks and the final work of CMM (*Men & Bryan-Kinns, 2018*), by manipulating acoustic attenuation, we can produce sonic privacy. For example, 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, Boyle & LaBerge 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 & 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. As shown in [Fig. 4](#), four space configurations were designed as the independent variable levels to investigate these three hypotheses, including:

Condition 1: Public space only (referred to as C_{pub}): where players can generate, remove or manipulate music interfaces, 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 + Augmented Attenuation Personal Space (referred to as C_{aug}). In addition to C_{pub} , the acoustic 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 in distance, the shift between personal space and public space is gradual.

Condition 3: Public space + Fixed Personal Space (referred to as C_{fix}). In addition to C_{pub} , each user is now provided with a personal space located at the corner of the stage (see [Fig. 4](#)), which works like a acoustically solid (soundproof) boundary between public space and personal space. In other words, the shift between personal space and public space is now rigid. Each user has a handle to activate/deactivate the personal space, the handle appears automatically over their head when they look up.

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

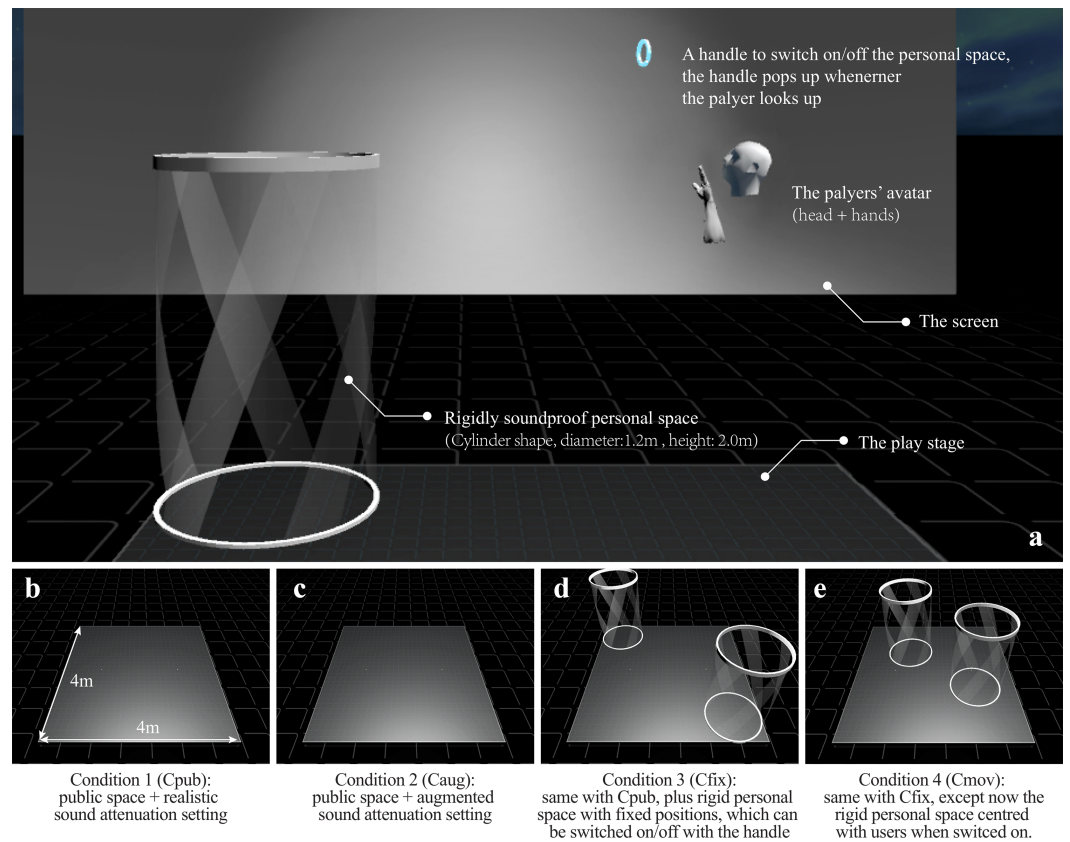


Figure 4 Components of the virtual experiment scene. Note: the rigidly soundproof personal space and the handle switch were only available in C_{fix} & C_{mov} (A); Tilted view of the four experimental condition settings (B, C, D, E).

Full-size DOI: [10.7717/peerjcs.229/fig-4](https://doi.org/10.7717/peerjcs.229/fig-4)

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

Dependent variables

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

Participant 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 of questions about the sense of the collaborator's presence (Schubert, Friedmann & Regenbrecht, 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 co-worker's presence and (ii) sense of collaborator's activities. Note that the measure "sense of self-presence" is not included in this study, because a previous study (*Men & Bryan-Kinns, 2019*) has shown LeMo's capability on producing a high-level of self-presence.

(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 & Hamilton, 2012*).

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

(5) Contribution: (i) the feeling of self's contribution; (ii) the feeling of others' contribution. These measures were set to see the effects of spatial configurations on the sense of contribution.

These measures will be grouped into a Post-Session Questionnaire (PSQ, see its items in [Table 1](#)), to be filled after participants experiencing each condition, and a Comparison Questionnaire (CQ, see its items in [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 a mutual note modification indicate an edit on a certain note, the last update of which was performed by the collaborator, cf. (*Bryan-Kinns, Healey & Leach, 2007*).

(2) Time and amount of use of personal space (only in condition C_{fix} and C_{aug} , as this measure is based on the time-stamped entries of the rigid boundary of personal spaces): (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; (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 study⁴ via emails sent to group lists within the authors' school. All participants were aged between 18 and 35, with an average age of 23.00 ($SD = 4.37$). Participants' mean rating of personal musical theory knowledge is 3.92 ($SD = 2.50$) on a 10-point Likert scale, where higher values indicate increased knowledge. Twenty four (46.15%) played one or more instruments, and the remaining 28 (53.85%) did not. Participants' rating of experience of collaboratively composing music is 2.13 ($SD = 1.56$) on a 10-point Likert scale, where higher values indicate

⁴The 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} M (SD)	C _{aug} M (SD)	C _{fix} M (SD)	C _{mov} M (SD)	C _{pub} vs C _{aug} p (W)	C _{pub} vs C _{fix} p (W)	C _{pub} vs C _{mov} p (W)	C _{aug} vs C _{fix} p (W)	C _{aug} vs C _{mov} p (W)	C _{fix} vs C _{mov} p (W)
PSQ1 (support for creativity)—I think the spatial configuration in this session was extremely helpful for creativity	8.55 (1.44)	8.77 (1.34)	7.61 (2.01)	7.82 (1.94)	0.5695 (259)	0.07706 (396.5)	0.1563 (379)	<u>0.02372</u> (492)	0.06318 (469)	0.695 (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 (1.92)	8.35 (1.50)	7.71 (1.88)	7.75 (1.62)	0.5211 (255.5)	0.6456 (331.5)	0.5029 (342)	0.2172 (434)	0.1452 (446.5)	0.8999 (400)
PSQ3 (preference)—I enjoyed the spatial configuration of this virtual world very much	8.27 (1.61)	8.65 (1.60)	8.18 (1.87)	8.07 (1.86)	0.2622 (233)	0.9358 (303.5)	0.3863 (311.5)	0.3863 (412.5)	0.2165 (433.5)	0.8010 (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 (0.92)	8.54 (1.68)	7.07 (2.52)	7.93 (2.26)	0.7961 (298.5)	<u>0.004813</u> (450)	0.1636 (377.5)	<u>0.01946</u> (497)	0.3229 (420)	0.1368 (302)
PSQ5 (content assessment)—How satisfied are you with the final piece of loop music you two created in this session	8.64 (1.73)	8.38 (1.50)	7.21 (2.22)	8.32 (1.96)	0.4287 (323.5)	<u>0.005155</u> (448.5)	0.5557 (337.5)	0.05449 (473.5)	0.803 (349.5)	<u>0.02163</u> (254)
PSQ6 (communication quality)—How would you rate the quality of communication between you and your collaborator during the session	8.68 (1.09)	8.50 (1.36)	7.04 (2.25)	8.04 (1.97)	0.7644 (300.5)	<u>0.004494</u> (450.5)	0.3038 (359)	<u>0.01038</u> (510)	0.5404 (399)	0.05073 (274)
PSQ7 (sense of collaborator's activity)—I had a clear sense of what my collaborator was doing	8.73 (1.20)	7.96 (1.54)	6.50 (2.52)	7.29 (2.49)	0.08094 (368.5)	<u>0.0003856</u> (487.5)	<u>0.03436</u> (414.5)	<u>0.02786</u> (489.5)	0.5095 (402)	0.176 (310)
PSQ8 (amount of contribution)—The amount of your contribution to the joint piece of music is	8.41 (1.44)	8.15 (1.46)	6.96 (2.15)	7.50 (1.67)	0.4776 (320)	<u>0.009236</u> (439.5)	<u>0.03928</u> (412)	<u>0.04281</u> (479.5)	0.166 (443)	0.4489 (346)
PSQ9 (amount of contribution)—The amount of your collaborator's contribution to the joint piece of music is	8.18 (1.26)	8.23 (1.39)	7.29 (1.96)	7.61 (1.97)	0.8486 (276.5)	0.08916 (394)	0.4025 (350.5)	0.06406 (469.5)	0.3008 (423)	0.4739 (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 (1.70)	7.81 (1.41)	7.36 (1.68)	7.86 (1.53)	0.319 (333.5)	0.1031 (390)	0.4648 (345)	0.3596 (416.5)	0.2829 (327)	0.8599 (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 (1.52)	8.19 (1.20)	7.54 (1.50)	7.75 (2.05)	0.3496 (241.5)	0.5636 (337.5)	0.6459 (284.5)	0.1143 (453.5)	0.6992 (386)	0.3559 (336.5)

Notes.^aWith 10-point-Likert scale, 1 indicates no fulfilment at all with the description of the questionnaire and 10 indicates a full fulfilment.^bNote statistics in this table are calculated based on the data collected from the third and fourth session to counterbalance the learning effect.

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	<u>0.02205</u>	7	<u>0.03317</u>
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	<u>0.01054</u>	9	0.1292
CQ3 (coordination)—Which session you found most difficult to track collaborator's activities?									
	most difficult	7	<u>0.03317</u>	12	0.4469	20	<u>0.02205</u>	13	1.000
	least difficult	22	<u>0.004691</u>	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	<u>2.807e-05</u>	17	0.1322	2	<u>5.277e-05</u>	6	<u>0.01368</u>
	least strongest	4	<u>0.001378</u>	7	<u>0.03317</u>	28	<u>8.12e-06</u>	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	<u>0.02205</u>	17	0.1322	4	<u>0.001378</u>	11	0.3232
	worst quality	6	<u>0.01368</u>	13	1.000	25	<u>0.0002698</u>	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	<u>0.04298</u>	10	0.2146
CQ7 (coordination)—Which session did you find most difficult to cooperate with collaborator									
	most difficult	7	<u>0.03317</u>	12	0.4469	22	<u>0.004691</u>	11	0.3232
	least difficult	21	<u>0.01054</u>	14	0.4262	7	<u>0.03317</u>	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	<u>0.03317</u>

Notes.

^aLower-tailed test when $k < 13$, two-tailed test when $k = 13$, upper-tailed test when $k > 13$.

increased experience. Regarding familiarity with computers, 31 (59.62%) participants rated themselves as computer “experts”, 20 (38.46%) chose “intermediate”, and only 1 participant (0.02%) chose “beginner”. Twenty participants (38.46%) had tried VR 2–5 times before, 20 (38.46%) had only tried once, the remaining 12 (23.08%) had no VR experience before. Thirty-seven participants knew their collaborators very well prior to the experiment, 3 met their collaborators several times, but did not know well, the remaining 12 did not know their collaborators at all prior to the experiment.

The experimental hardware setup was exactly the same with LeMo standard setup, see more details in the previous section “LeMo—An SVE for collaborative music making”. After reading information forms and signing consent forms, each pair of participants first received an explanation of the music interface of LeMo (see Fig. 3). Then one

experimenter demonstrated all of the interaction gestures supported 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 had a trial (5–15 min) 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 complimented an animation loop, each session lasts 7 min. Based on our pilot study and a previous study (Men & Bryan-Kinns, 2018), we found 7 min was sufficient for the task. To counter-balance the learning effect, all four conditions were experienced in a fully randomised sequence to counterbalance the learning effect. In total four animation loops were introduced to trigger participants' creativity, each to be played in one experimental session on four virtual screens surrounding the virtual stage. These clips were played in an independently randomised sequence to counterbalance impacts on the study. Each session ended with a Post-Session Questionnaire (PSQ, see Table 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

This section reports on the results of the questionnaires. Ratings of Post-Session Questionnaire 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, results are listed 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.

Counterbalancing the learning effect

As aforementioned, we introduced a fully randomised order of experimental conditions to counterbalance the learning effect. However, it turned out many measurements in the Post-Session Questionnaire were still affected by the sequence to a certain extent, as shown in Fig. 5, in which data from all groups were compiled according to how the group was ordered in the session sequence. Wilcoxon Rank Sum tests were run between each two conditions for every question. A red solid arc indicates a significant difference between two bars ($p < 0.05$), and a grey dotted arc indicates a trend toward a significant difference ($p < 0.1$). The arcs show that results of four questions (especially PSQ1, PSQ5, PSQ8, PSQ9, PSQ10, PSQ11) 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 (PSQ1), higher satisfaction with their output (PSQ5), and both more, and better quality of contributions by themselves and contributors (PSQ8, PSQ9, PSQ10, PSQ11). This is probably due to the learning effect which has a much stronger effect on these measures compared with the differences between experimental conditions.

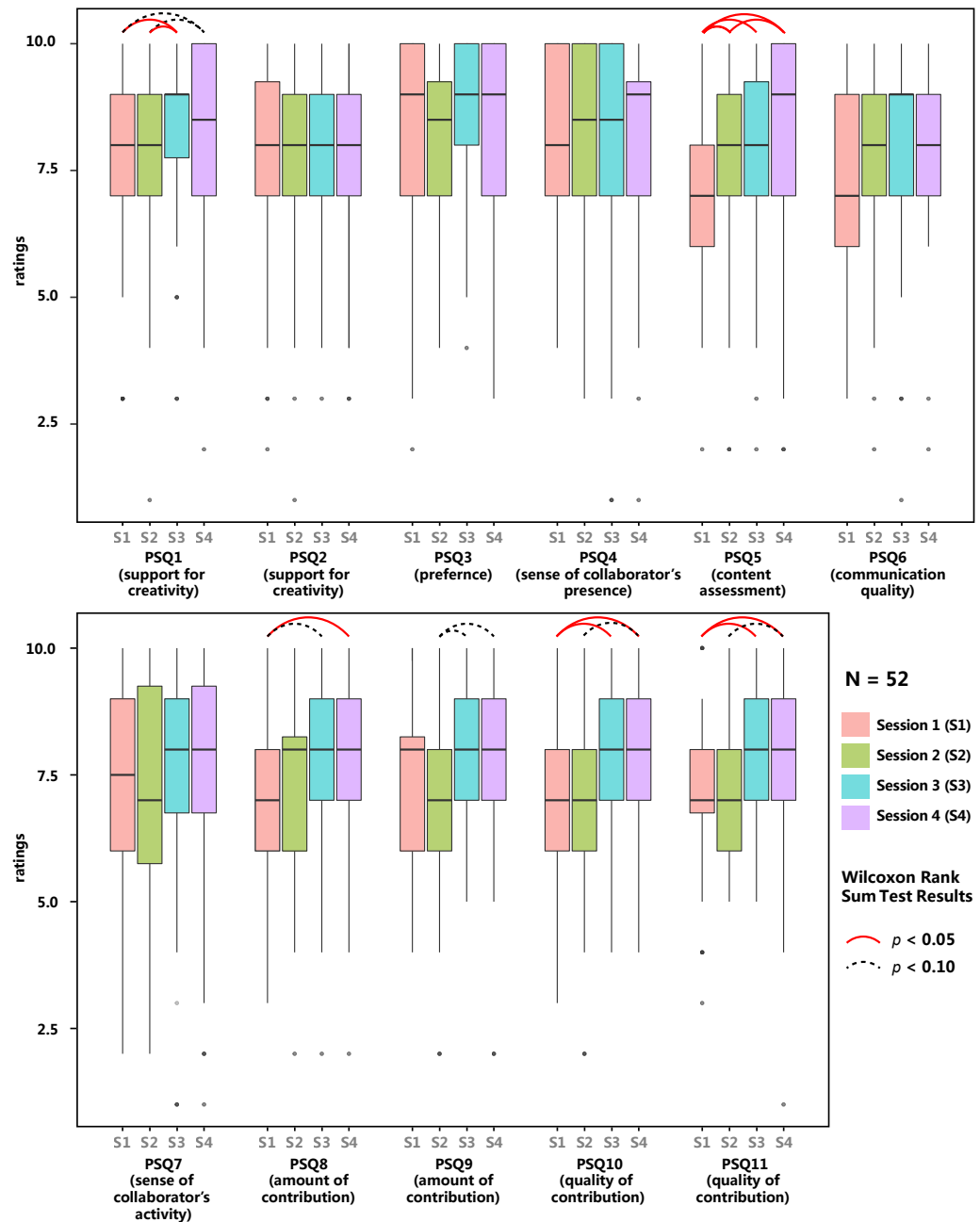


Figure 5 Results of Post-Session Questionnaire ($N = 52$) for all sessions. Arcs show significant (solid line) and marginal-significant (dotted line) differences between conditions, indicating possible ordering effects.

Full-size DOI: 10.7717/peerjcs.229/fig-5

Given the limited experience some participants had in VR and collaborative music making, learning effect could have possibly and greatly promoted participants' skills and knowledge in performing the task, resulting in a better feeling of the spatial configuration of the session, higher quality of output, more contribution with better quality in later sessions.

This learning effect has been also mentioned by some participants in the interview. More details will be discussed in the later subsection “Interviews”.

To better counterbalance the learning effect and habituation on PSQ, only data collected via PSQ in later two sessions (session 3 and 4) will be retained at the expense of the halved sample size. Box-plots were then drawn (Fig. 6) and Wilcoxon Rank Sum tests were run (Table 1) to compare the conditions against each other.

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

When asked the condition’s helpfulness for creativity (PSQ1 in Table 1), on a 10-point Likert Scale, participants gave an average rating of 8.77 in C_{aug} , which is significantly higher than 7.61 given in C_{fix} (Wilcoxon Rank Sum Test, $W = 492$, $p = 0.02372$). There are trends towards significances between participants’ rating of C_{aug} and C_{mov} ($W = 469$, $p = 0.06318$), and between C_{pub} and C_{fix} ($W = 396.5$, $p = 0.07706$). These differences indicate that C_{aug} is better than C_{fix} , and possibly also better than C_{mov} in terms of supporting participants’ creativity.

When asked to rate the helpfulness of spatial configurations to support personal idea development (PSQ2), the mean rating of C_{aug} ($M = 8.35$) is higher than that of the other three conditions (C_{pub} : $M = 7.82$; C_{fix} : $M = 7.71$; C_{mov} : $M = 7.75$), though no significant differences were found. CQ7 of Table 2 shows that C_{pub} was rated by significantly many participants (21 out of 52) to be the least difficult to cooperate with their collaborator (Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided), whilst significantly few participants rated C_{pub} as the most difficult one to do so (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided). On the opposite, C_{fix} was rated by significantly many participants as the most difficult (Binomial Test, $0.42 > 0.25$, $p = 0.004691$, 1-sided), and significantly few participants (7 out of 52) rated it as the least difficult (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided).

Preference

When asked the level of enjoyment of the spatial configuration (PSQ3), similar to PSQ2, C_{aug} got a higher rating ($M = 8.65$). However, no significant difference was revealed. In CQ1 of Table 2, when asked which session has the most enjoyable spatial configuration, out of 52 participants, both C_{aug} and C_{mov} were chosen by 16 participants, more than those choosing C_{pub} and C_{fix} (10 participants each), though no significant difference was found. When asked which session had the least enjoyable spatial configuration, a significant number of participants (20 out of 52) opted C_{fix} ($0.38 > 0.25$, $p = 0.02205$, 1-sided), and significantly few (7 out of 52) opted C_{mov} ($0.13 < 0.25$, $p = 0.03317$, 1-sided). Result of CQ6 in Table 2 indicates that significantly many participants (19 out of 52) believed C_{fix} was the worst setting for creating a good piece of music collaboratively. To summarise, the spatial configuration of C_{fix} is more disfavoured and that of C_{mov} is less disfavoured.

Sense of co-presence

Results of PSQ4, PSQ7 and CQ4 reveal participants’ sense of collaborators’ presence and activities. C_{fix} ’s ratings in PSQ4 and PSQ7 are significantly lower than C_{pub} and

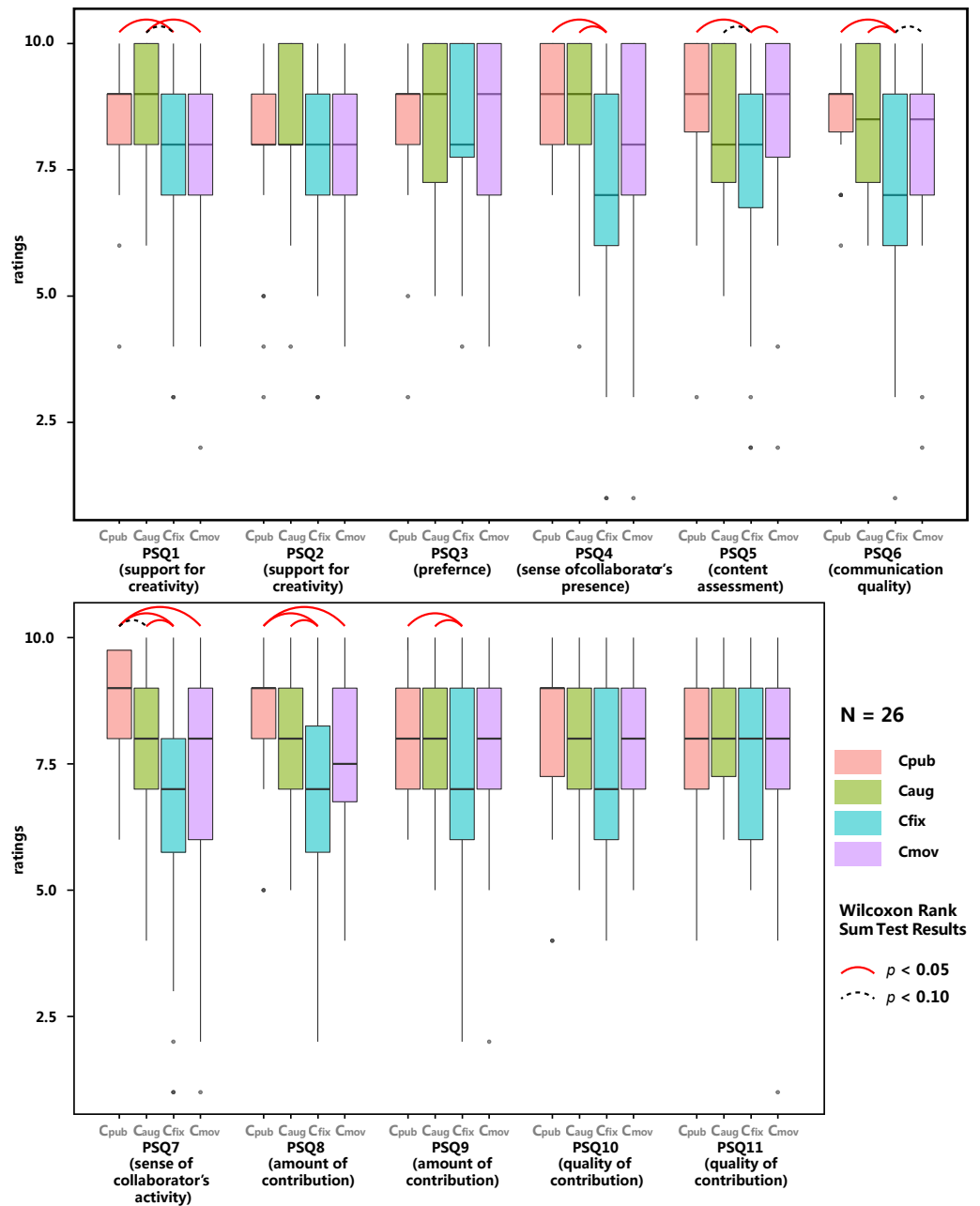


Figure 6 Results of Post-Session Questionnaire ($N = 26$), data grouped by experimental conditions. Only data collected in the latter two sessions are included. Solid arcs show significant differences and dotted arcs show marginal-significant differences between conditions.

Full-size [DOI: 10.7717/peerjcs.229/fig-6](https://doi.org/10.7717/peerjcs.229/fig-6)

C_{aug} (Wilcoxon Rank Sum Test, all $p < 0.05$), indicating C_{fix} saw a lower sense of the collaborator's presence and activities. Similarly, when being asked in which session they had the strongest sense of collaborators (CQ4), significantly many participants (27) chose C_{pub} ($0.52 > 0.25$, $p = 2.807e-05$, 1-sided), 17 chose C_{aug}, significantly few chose C_{mov}

(chosen by 6) and C_{fix} (chosen by 2; Binomial Test, both $p < 0.05$; more details in CQ4 of Table 2). When questions changed to “least sense of collaborator’s presense”, ratings reversed, significantly many chose C_{fix} while significantly few chose C_{pub} and C_{aug} (Binomial Test, all $p < 0.05$; more details in CQ4 of Table 2). These results indicate that in terms of maintaining the sense of collaborator’s presence, $C_{\text{pub}} > C_{\text{aug}} > C_{\text{mov}} > C_{\text{fix}}$.

Regarding the sense of collaborator’s activities (PSQ7), a significantly weaker sense was reported in C_{fix} compared with C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$). C_{pub} also saw a stronger sense compared with C_{fix} (Wilcoxon Rank Sum Test, $W = 414.5$, $p < 0.0346$). No significant difference was found between C_{pub} and C_{aug} nor between C_{fix} and C_{mov} . Similarly, CQ3 of the Comparison Questionnaire reveals that significantly many participants reported tracking collaborators’ activities in C_{fix} was the most difficult, and significantly many felt least difficult to do so in C_{pub} (Binomial Test, both $p < 0.05$). These indicate that C_{pub} seems to be easier for participants to track collaborators’ activities, and C_{fix} is more difficult for them to do so.

Content assessments

Participants reported a mean rating 7.21 of output quality in C_{fix} (PSQ5 of Table 1), which is significantly lower than 8.64 in C_{pub} , and 8.32 in C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.05$), and quasi-significantly lower than 8.38 in C_{aug} ($W = 473.5$, $p = 0.05549$). Similarly, significantly many participants reported that they produced the least satisfying piece of music in C_{fix} (Binomial Test, $0.40 > 0.25$, $p = 0.01054$, 1-sided), see CQ2 of Table 2. In other words, C_{fix} tended to led to a music output with lower quality compared with other sessions.

Communication assessments

As listed in PSQ6, communication quality of C_{fix} ($M = 7.04$) is significantly lower than 8.68 of C_{pub} and 8.50 of C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$), and near-marginal significantly lower than 8.04 of C_{mov} (Wilcoxon Rank Sum Test: $W = 274$, 0.05073), see PSQ6 in Table 1 and PSQ6 in Fig. 6. When asked to compare these sessions, significantly many participants reported the best communication quality was in C_{pub} and significantly few believed they had the best communication quality in C_{fix} (Binomial Test, both $p < 0.05$, see CQ5 of Table 2). Conversely, significantly few had the worst communication quality in C_{pub} and significantly many had worst in C_{fix} (Binomial Test, both $p < 0.05$, see CQ5 of Table 2). In summary, C_{fix} saw a relatively lower communication quality.

Contribution

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

No significant differences were found in the ratings of the amount of the collaborators’ contribution (PSQ9), except a trend reporting their collaborator had a lower amount of contribution in C_{fix} than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, $W = 469.5$, both $p <$

0.1). However, CQ9 reveals significantly few participants felt that their collaborator did the most contribution in C_{mov} (Binomial Test, $0.13 < 0.25$, $p = 0.03317$, 1-sided). These results indicate that the addition of personal space in C_{fix} and C_{mov} possibly led to a weaker sense of collaborator's activities.

Activity assessments

This section reports on measures focusing on the participants' interactive activities. All measures are listed in Table 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 of C_{pub} , 80.27 of C_{aug} , and 77.69 of C_{mov} (Wilcoxon Rank Sum Test, all $p < 0.05$). Note additions, as the main part of note edits, follow a similar pattern. The number of note additions in C_{fix} is significantly greater than that of C_{pub} ($W = 1026$, $p = 0.03429$), and near-marginal significantly greater than that of C_{aug} and C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.07$, check detailed statistics in 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, based on the system-logged data, we plotted the positions and directions of participants' heads and musical note edits on a top view of the stage, see Fig. 7 as an illustrative example of visual traces from arbitrarily selected group 3. Visual traces of all groups are shown in Fig. 8. These figures were made based on system logged data, specifically, the arrows were participants' locations at 20-second intervals for ease of reading the diagram, and dots are the locations of participants' hands

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}	p (W)
	M (SD)	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 (36.59)	80.27 (36.92)	98.35 (48.67)	77.69 (34.61)	0.6988 (1292)	<u>0.02386</u> (1004)	0.7599 (1304.5)	<u>0.03375</u> (1025)	0.8965 (1372.5)	<u>0.02228</u> (1704)	
AA2—No. of note additions	50.23 (27.12)	58.96 (30.03)	72.88 (40.93)	55.98 (25.31)	0.8301 (1318.5)	<u>0.03429</u> (1026)	0.8149 (1315.5)	0.06572 (1068.5)	0.876 (1376.5)	0.05591 (1646.5)	
AA3—No. of note deletions	20.90 (14.46)	21.31 (12.94)	25.46 (18.39)	21.71 (15.15)	0.7108 (1294.5)	0.243 (1172)	0.8376 (1320)	0.3308 (1202)	0.9689 (1358.5)	0.3323 (1501.5)	
AA4—No. of mutual note modifications ^a	4.37 (4.42)	4.23 (5.57)	3.71 (7.69)	2.44 (3.92)	0.6452 (1422.5)	<u>0.01929</u> (1703.5)	<u>0.007331</u> (1754.5)	0.06614 (1627.5)	<u>0.02514</u> (1687.5)	0.7732 (1394.5)	
AA5—Size of group territory (unit: m ²)	0.3465 (0.2443)	0.4331 (0.2446)	0.2339 (0.1878)	0.3103 (0.1942)	0.152 (259)	0.1013 (428)	0.7099 (359)	<u>0.005236</u> (491)	0.09421 (430)	0.2639 (276.5)	
AA6—Size of personal territory (unit: m ²)	0.4282 (0.1690)	0.4547 (0.2193)	0.7475 (0.1801)	0.5067 (0.1894)	0.9559 (1343)	<u>2.25e-12</u> (272)	<u>0.02347</u> (1003)	<u>2.085e-10</u> (374)	<u>0.04421</u> (1042)	<u>2.3e-08</u> (2212)	
AA7—No. of group edits (note edits done in group territory)	36.44 (35.24)	43.04 (34.79)	17.50 (23.79)	25.23 (29.00)	0.2913 (1189.5)	<u>0.001448</u> (1837)	0.07839 (1621.5)	<u>4.043e-05</u> (1977)	<u>0.009044</u> (1751.5)	0.1822 (1151)	
AA8—No. of personal edits (note edits done in own personal territory)	40.50 (44.81)	37.10 (38.42)	80.62 (51.89)	52.42 (38.81)	0.9610 (1360)	<u>1.179e-05</u> (678)	<u>0.0294</u> (1017)	<u>2.157e-06</u> (623)	<u>0.02016</u> (994.5)	<u>0.003695</u> (1799)	
AA9—No. of note edits done in other's personal territory	0.058 (0.42)	0 (0)	0.19 (0.89)	0.019 (0.14)	0.3267 (1378)	0.1797 (1275)	1.000 (1352.5)	<u>0.04343</u> (1248)	0.3267 (1326)	0.1686 (1431)	
AA10—Average distance between collaborators (unit: metre)	1.11 (0.38)	1.12 (0.38)	2.19 (0.58)	1.28 (0.41)	0.8632 (348)	<u>4.731e-11</u> (26)	0.2119 (269)	<u>2.663e-10</u> (34)	0.08045 (242)	<u>2.459e-08</u> (616)	
AA11—No. of uses of personal spaces	- (-)	- (-)	2.40 (1.95)	2.85 (2.14)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	0.2912 (1193)

(continued on next page)

Table 3 (continued)

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}	p (W)
		M (SD)	M (SD)	M (SD)	M (SD)	p (W)	p (W)	p (W)	p (W)	p (W)	p (W)
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	-	-	-	-	-	<u>0.008019</u>
	(-)	(-)	(-)	(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	<u>0.01032</u>	0.364	0.4725	<u>0.001757</u>	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	<u>0.005451</u>	0.1591	0.446	<u>0.0001145</u>	<u>0.03122</u>	<u>0.02838</u>	
	(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	<u>0.03719</u>	0.6655	<u>0.02047</u>	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	<u>0.0006613</u>	0.8888	<u>0.0001704</u>	0.6773	<u>0.002744</u>	
	(19.23)	(19.42)	(13.55)	(20.67)	(1312)	(1876)	(1374)	(1930.5)	(1416.5)	(891)	
AA20—No. of music interface additions	x	x	x	x	x	x	x	x	x	x	
	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	
AA21—No. of music interface additions in public space	-	-	x	x	-	-	-	-	-	x	
	(-)	(-)	(x)	(x)	(-)	(-)	(-)	(-)	(-)	(x)	
AA22—No. of music interface additions in personal space	-	-	x	x	-	-	-	-	-	x	
	(-)	(-)	(x)	(x)	(-)	(-)	(-)	(-)	(-)	(x)	

Notes.

^aMutual note modifications include activation/deactivation, the last update of which was performed by the collaborator.

^bData 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.

^cThe 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).

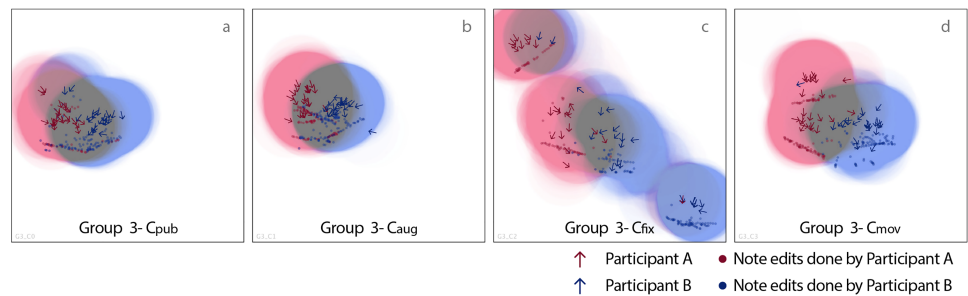


Figure 7 Illustrative example of visual traces of the participants' locations, directions and musical note edits (group 3). Arrows show participant's position and direction at 20 s intervals, dots show participant's hand's position while performing note edits.

Full-size DOI: 10.7717/peerjcs.229/fig-7

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, Carpendale & Inkpen, 2004; Scott & 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 colours (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, and permits participants to touch each other or the same music interface (Hall, 1966), most of the musical note edits also fell inside this range.

(1) Distribution of locations and interactions. The more intensely blue or red the area is, the more presence the corresponding participant had shown in that location. The overlap is coloured grey, indicating the appearance of both participants, which can be seen as group territory.

(2) Sizes of group territory and group edits (edits undertaken in 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 m² of group territory in C_{pub}, 0.4331 m² in C_{aug}, 0.2339 m² in C_{fix} and 0.3103 m² in C_{mov} (AA5 of Table 3). Results of the Wilcoxon Rank Sum Tests show that the size of group territory of C_{aug} is significantly larger than that of C_{fix} ($W = 491$, $p = 0.005236$), and near-marginal significantly larger than C_{mov} ($W = 430$, $p = 0.09421$). No significant difference was found between C_{pub} and C_{aug}.

AA7 of Table 3 shows that participants had an average of 36.44 group edits in C_{pub}, which is significantly more than 17.50 of C_{fix} (Wilcoxon Rank Sum Test, $W = 1837$, $p = 0.001448$), and a near-marginal significantly more than 25.23 of C_{mov} ($W = 1621.5$, $p = 0.07839$). C_{aug} resulted in a higher average of group edits ($M = 43.04$), though not significantly higher than C_{pub}, it is significantly higher than numbers of C_{fix} and C_{mov} (Wilcoxon Rank Sum Test, both $p < 0.01$). These indicate the spatial configurations of C_{pub} and C_{aug} are more friendly to group edits.

(3) Sizes of personal territory (AA6 of Table 3) and personal edits (edits fallen into personal territory; AA8). Participants formed a significantly larger personal territory in C_{fix}

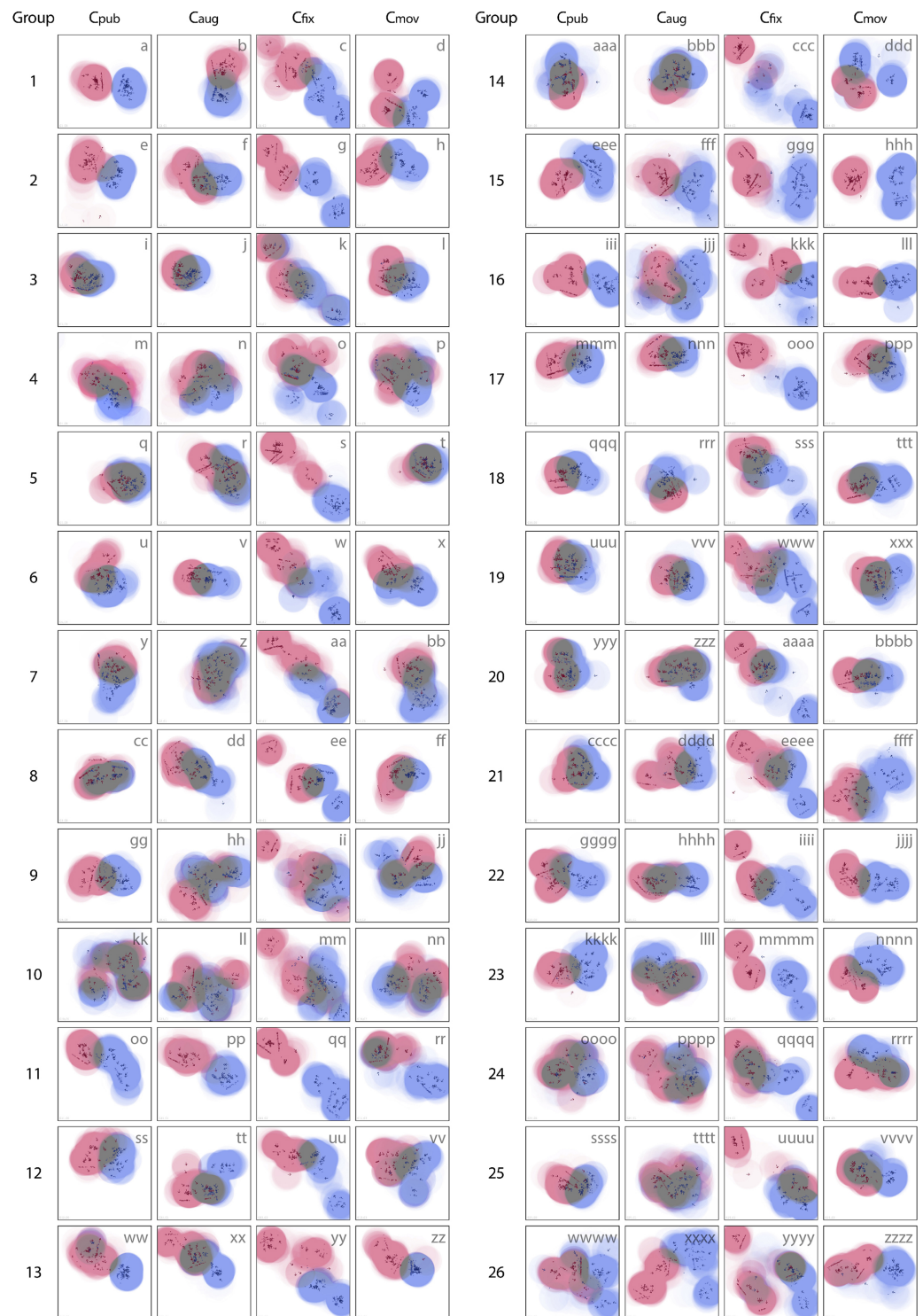


Figure 8 Visual traces—the participants' locations, directions and musical note edits shown on a top view of the stage (based on system-logged data of all groups).

Full-size  DOI: [10.7717/peerjcs.229/fig-8](https://doi.org/10.7717/peerjcs.229/fig-8)

($M = 0.7475 \text{ m}^2$, $SD = 0.1801$) compared with all the other three conditions (Wilcoxon Rank Sum Test, all $p < 0.001$), and had significantly more personal edits in C_{fix} compared with other conditions (Wilcoxon Rank Sum Test, all $p < 0.01$; AA8). Similarly, a larger size of personal territory was formed in C_{mov} and more personal edits were done in C_{mov} compared with C_{pub} and C_{aug} (all $p < 0.05$). No significant differences were found between C_{pub} and C_{aug} , neither in the size of personal territory nor in the number personal edits. To summarise, C_{fix} results in the largest size of personal territory and the largest number of personal edits, the metrics of C_{mov} follows, and C_{pub} and C_{aug} have the least, indicating that C_{fix} led to a much looser collaboration, in which participants worked more independently, whilst C_{pub} and C_{aug} , on the opposite, led to more interactivities in the group territory.

(4) Average distance (AA10 of Table 3). Participants had an average distance of 2.19 metres between themselves and their collaborators in C_{fix} , this is significantly bigger than other three conditions (Wilcoxon Rank Sum Test, all $p < 0.001$). Namely, in the other three sessions participants worked more closely.

Times and amount of use of personal space

As shown in AA11, AA12 and AA13 of Table 3, in C_{fix} , participants had an average of 2.40 entries of personal space on average, with each entry lasting 73.07 s with total duration 128.60 s. For C_{mov} , the participants did 2.85 entries, each lasting 49.54 s on average, with a total usage time of 112.19 s, no significant difference was found in the number of entries or in the sum usage time. However, the average duration of each entry of C_{mov} is significantly shorter than that of C_{fix} (Wilcoxon Rank Sum Test, $W = 1512$, $p = 0.008019$), indicating that the personal spaces of C_{fix} were possibly more used for 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 s in C_{aug} , which is significantly longer than 7.19 s in C_{pub} , and 7.31 s in C_{fix} (Wilcoxon Rank Sum Test, both $p < 0.05$), and near-marginal significantly longer than 11.02 s 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} (Wilcoxon Rank Sum Test, C_{pub} vs C_{aug} : $W = 924.5$, $p = 0.005451$; C_{aug} vs C_{fix} : $W = 1945$, $p = 0.0001145$; C_{aug} vs C_{mov} : $W = 1683.5$, $p = 0.03122$). Wilcoxon Rank Sum Test result also shows participants paid their attention to their partner significantly fewer times in C_{fix} than they did in C_{mov} ($W = 1015$, $p = 0.02838$), see 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} possibly promoted insignificantly and C_{fix} demoted insignificantly compared with C_{pub} .

(3) The time participants spent paying ordinary attention to each other. Different from the impact of spatial configurations on close attention, neither C_{aug} nor C_{mov} significantly changed the way participants paying their ordinary attention, all fell inside the range from 70 to 80 s out of the 420-second session. C_{fix} greatly reduced participant's ordinary

attention paid to each other, on average participants only spent 52.89 s on doing this, which is significantly shorter than C_{pub} and C_{aug} (Wilcoxon Rank Sum Test, both $p < 0.05$) and near-marginal significantly shorter than C_{mov} (Wilcoxon Rank Sum Test, $W = 1081$, $p = 0.07865$), more details are shown in 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 (Wilcoxon 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 & Clarke, 2006; Yin, 2017). The starting point of the thematic analysis was a reading through of the transcripts, 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 Fig. 9, in total, 650 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, 40 coded segments contributed by 24 participants were related to learning effects. Participants reported the sequence is an “important factor” (Participant 15A, hereafter abbreviated to P_{15A}). The first session was felt to be hard as they were “just being introduced to [the system and they were] still adjusting” to it (P_{5A}), trying to “[figure] out how the system was working” (P_{16A}). When they “were progressing into latter sessions, [they] felt easier to communicate and use gestures to manipulate the sound, being able to collaborate more, more used to the system” (P_{5B}), these changes led to a higher level of satisfaction and more enjoyment in later conditions. It should also be noted that interestingly P_{11A} and P_{11B} reported the sequence effect adversely, they enjoyed the first session more because “the first one was an element of surprise, a total surprise” as that was “the first time they were using the system”. That feeling of freshness made that session more exploratory and more joyful to them. These learning effects might possibly affect the results of Post-Session Questionnaire and Comparison Questionnaire and thus should be well counter-balanced.

Reporting the spatial configurations

(1) C_{pub} - **Simple but can be chaotic**. Since there is no personal space, participants could, and had to hear all the interfaces all the time. In total, 16 coded-segments are about the disadvantage of this setting, some examples are: “a bit troubling” (P_{11B}), “music [was] always very loud” (P_{9A}), “it was global music, and there was someone annoying” (P_{2A}), “[they were] not going to say anything” about disagreement because that could possibly make them appear to be “rude” (P_{2A}). It was easier if there is something helpful “to perceive

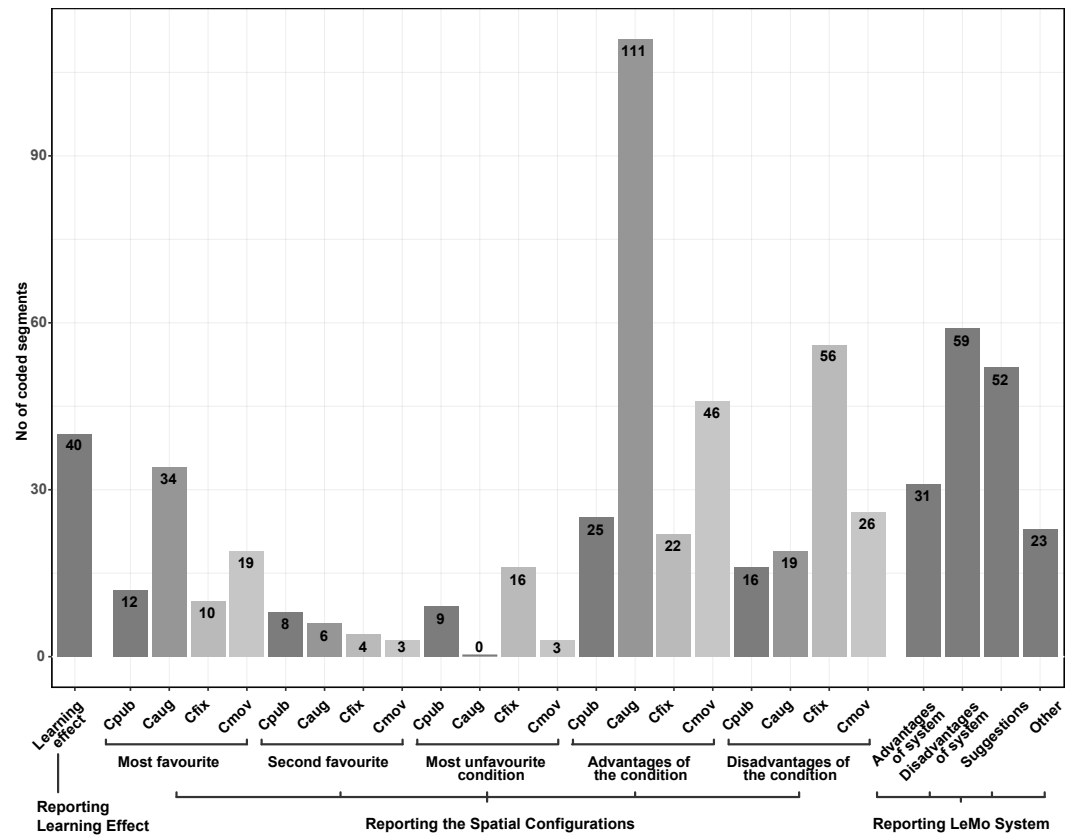


Figure 9 Ingredients of all the coded-segments of the interview, numbers of coded segments are shown along the bars. .

Full-size DOI: 10.7717/peerjcs.229/fig-9

what [they were] doing, and not get confused with what [their collaborator] was doing” (P_{15B}), it was too “chaotic” (P_{20A}), “too confusing” (P_{22A} & P_{22B}), “annoying” (P_{25B}). They “cannot concentrate” (P_{25B}), “everything [was] open and quite noisy” (P_{26B}), they didn’t “have the tranquillity to operating [the] sounds...everything [came] mixed, which [was] difficult to manage” (P_{22A}).

There were 25 coded segments from 14 participants reporting the positive side of the C_{pub}, some examples are: (i) pieces created in “personal space” might clash in a music way (P_{1A}), “better to work when knowing how it sounds all together” (P_{17B}), music pieces might match better; (ii) better for providing help to the other, according to P_{4A}, they needed someone to lead them and thus the ability to hear all the work all the time was helpful; (iii) “space wise”, namely, no space limitation, compared with having to work closer to “hear the sound well” (P_{12A}), C_{pub} does not have this constraint, they could choose to work “anywhere” (P_{24A}); (iv) “easier” to understand the condition (P_{6B}), fewer confusions when simply being able to hear all the things all the time (P_{13A}); (v) “collaborative wise” (P_{13A}), less separation, better collaboration compared conditions where “personal space” was provided (P_{3B}, P_{18A} and P_{18B}).

(2) **C_{pub} - Simple but can be chaotic.** There were 34 coded segments contributed by 24 participants favouring condition C_{aug}, higher than 12 segments contributed by 11 participants for C_{pub}, 10 segments contributed by eight participants for C_{fix}, and 19 segments contributed by 17 participants for C_{mov} (the sum of number of contributors here is greater than 52 as a few participants reported more than one favourite condition in the interview). The reason for the popularity can be concluded from the overwhelming 111 coded segments from 33 participants from 25 groups reporting the advantages of this condition, much higher than the number of segments reporting other conditions' advantages. C_{aug}'s advantages reported by participants can be grouped into four groups: (i) **Higher team cohesion and less sense of separation.** Participants reported that without the rigid personal space, they had to “work with the other person” (P_{6A}). With no rigid personal space, C_{aug} “forced [them] to collaborate more...because [they] had to stay very close” to compose music (P_{9B}). (ii) **An appropriate environment** for creativity, more consistency and convenience. As described by participants, it was “a middle point between personal space and no personal space” (P_{6A}), without even triggering something, “[they] could decide in a continuous way”, “whether [they] were able to listen to the other sound sources or not, [and] to what extent [they] wanted to isolate [themselves]” (P_{16A}). Compared with having to hear all sounds in C_{pub}, this provided them with a “less stressing” (P_{4A}) context, and they could selectively move away to avoid “getting interrupted with the other” (P_{5B}) and overlapping music. Compared with C_{fix} and C_{mov}, being able to still “hear a bit of it in the background but not completely” (P_{20A}) was reported good as this kept them “up to date” (P_{9A}) and helped them to “tailor what [he/she] was making” (P_{22B}) to match the co-created music and to make something new and see if it “fit with” (P_{20A}) the old. C_{aug} provided them with “a little bit of personal space” although not a quite “defined thing” (P_{6A}), which provided the possibility “to work on something individually” and to “share work quite easily” (P_{20A}). (iii) **Easier to identify sounds.** Participants reported it was easier to “locate the source of the sound” (P_{16A}) and “perceive what (they were) doing” (P_{15B}), these factors then helped them “understand instruments better” (P_{7B}) and “not get confused” (P_{15B}); (iv) **More real.** Quite interestingly, instead of C_{pub}, which simulates the acoustic attenuation in the real world, C_{aug} was reported to be similar to the experience in the real world. Participants reported in C_{aug} “if [they] want to hear something, [they] just come closer, like in the real world” (P_{11B} & P_{11B}), “feeling like the real-time experience (P_{26B})”.

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

(3) **C_{fix} and C_{mov} - resemblance and differences.** Regardless of the mobility, the personal space provided by C_{fix} and C_{mov} share the same characteristics. Not surprisingly, the participants reported many common advantages and disadvantages shared by both conditions, including: The addition of rigid personal space was described as an “added advantage” (P_{7A}), it made it “easier to perceive what [they were] doing and not get confused

with what [their collaborator] was doing” (P_{15B}), provided them with a chance to “isolate themselves to create their piece” (P_{22A}), and to “think about something to add” (P_{9A}), helped them to “develop their own ideas” (P_{8A}). As a result, they used personal space “a lot [and] used [their] own creativity much more comparing with [other two sessions]” (P_{3A} & P_{3B}).

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

Differences between C_{fix} and C_{mov}—In total, 46 coded segments (from 26 participants) were reporting C_{mov}’s advantages and 26 segments (from 12 participants) reporting its disadvantages, compared with 22 segments (from 14 participants) and 56 segments (from 34 participants) for C_{fix}, indicating in general participants thought C_{mov} better than C_{fix}. Some example insights behind the preference are: C_{mov} functioned like a “mute button” (P_{4B}), which could be used anywhere (P_{7B}), enabling them to “move around”, work “closer...and see each other’s things” and thus led to “more collaboration” between them (P_{1B}). Though C_{fix} had no advantages on these aspects, the location at the opposite corners provided a more “personal feeling” and a higher sense of belonging (P_{22A} & P_{22B}). Walking to the corner to access personal space was not a big issue for P_{7B} & P_{7A} as “the boundary is small”. Besides, the relatively far distance also helped to “prevent [them] from clashing” (P_{7A}).

DISCUSSION

The key question that this paper tries to address is how shared virtual environments should be designed to support creative collaboration. To help answer this question, a better understanding of the role of personal and public space is needed given that previous research has highlighted the necessity of providing personal space with fluid transition to public space during collaboration (*Scott, Carpendale & Inkpen, 2004; Shen, Everitt & Ryall, 2003; Sugimoto, Hosoi & Hashizume, 2004*) and also that people do form public and personal space in collaboration in VEs (*Men & Bryan-Kinns, 2019*). Next, based on the results, we will discuss (i) the necessity and (ii) the impacts of adding each type of personal space, then make comparisons between (iii) conditions providing personal space with and without mobility, and iv) conditions providing personal space with rigid and 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 for 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, according to the interview results, participants reported that the background could 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. The personal space provided in C_{aug} , C_{fix} , and C_{mov} functioned like a “less stressing” context, within which, participants could better “understand instruments” and not “get confused”. Secondly, participants need an opportunity to develop their own ideas. From the interview results, having personal space was reported to be “an added advantage”, it helped to promote their own creativity, which was then combined and contributed to the joint piece. This echoes the findings in *Men & 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 were possibly the result of introducing rigid visible personal space, and C_{aug} was founded to have addressed these limitations well (details will be discussed in later subsections). Next, we discuss the impacts of introducing each of these personal spaces individually by comparing their condition with C_{pub} .

Impacts of adding personal space

As mentioned above, the previous study (*Men & Bryan-Kinns, 2019*) found that the addition of personal space located at the opposite side of the public space led to a shrunken size of group territory, fewer group note edits, a larger size of personal territory, more personal note edits, a larger average distance between collaborators, and fewer times of paying attention to collaborators. We argued these negative impacts are mainly due to, in that study, the personal spaces being distributed on the opposite side of the group space resulted in a larger distance between participants. So we proposed personal space with different features (e.g., gradual boundary in C_{aug} , mobility in C_{mov}) might reduce, or even minimise, these negative effects. Next, the impacts of introducing newly-featured personal spaces in C_{aug} , C_{fix} , C_{mov} , and how these negative effects might be eliminated or minimised will be discussed.

Invisible auditory personal space in C_{aug}

C_{aug} is quite similar to C_{pub} in many ways, e.g., both do not have a visual boundary for personal spaces, no visual triggers for personal space, similar territorial patterns were formed in these two conditions (Fig. 8). Not surprisingly, no significant differences were found in most of the statistical measures listed in Tables 1 and 3. The only differences revealed in these two Tables are the significant differences found in AA16, AA17—significantly more occurrences and longer duration of close attention were paid to collaborator in C_{aug} , 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 that the limitations of adding personal space identified in previous work (*Men & Bryan-Kinns, 2019*) have been successfully minimised. Specifically, the size of group territory and the number of group edits maintained similar numbers, the means of C_{aug} are even greater, though not significantly (AA5 and AA7 in [Table 3](#)). C_{pub} and C_{aug} saw a similar size of personal territory and personal edits, and similar average distance (respectively, AA6, AA8, AA10 in [Table 3](#)), and C_{aug} even saw more close-attention paid to each other (AA16 and AA17 in [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 to C_{pub} . In the previous study (*Men & Bryan-Kinns, 2019*), to access the personal spaces located at the opposite side of the public space, participants had to drift apart, which might have influenced their spatial locations, changed the formation of group/personal territory they formed and the average distance between collaborators changed, and territoriality-based interaction (group/personal edits) changed. Here in C_{aug} , by enabling participants to use personal space anywhere inside the stage with no specific triggers needed, C_{aug} managed to provide a user experience as similar as possible to C_{pub} . The second reason is more related to the impacts on subjective experience, by making the personal space invisible and gradual, the isolation and difficulty of coordinating that introduced by the additional rigid personal space was minimised. For instance, 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 pop up the personal space anywhere in the stage. In this way, personal spaces was provided with mobility. By doing so, several aforementioned negative effects found in *Men & Bryan-Kinns (2019)* were reduced. Specifically, these include the size of group territory, the average distance, times of paying attention to collaborator (respectively, AA5, AA10, AA16, and AA17 in [Table 3](#)). However, some significant differences remained, participants still had significantly fewer mutual note modifications, marginal-significantly fewer group edits and significantly more personal edits after personal space being introduced in C_{fix} and C_{mov} (respectively, AA4, AA7, and AA8 in [Table 3](#)). This can also be verified by the interview results. Compared with C_{aug} , participants reported a higher sense of isolation in C_{mov} and C_{fix} , both providing rigid-form personal spaces. Namely, C_{mov} , by making the personal space available anywhere in the stage, managed to drag participants closer, saw a similar group territory, however, participants' behaviour was still affected in many ways. Participants were still separated to some extent, which can be seen as a disadvantage of adding visible, solid personal space. In other words, C_{mov} did better than C_{fix} in minimising the negative impacts of adding personal space, but not as good as C_{aug} .

A more rigid personal space in C_{fix}

C_{fix} provided a much more inflexible personal space, which influenced participants' behaviour in many ways (see the significant differences between C_{pub} and C_{fix} in [Tables 1 and 3](#)). Not to mention participants' polarised ratings on C_{pub} and C_{fix} in CQ3, CQ4,

CQ5, CQ7 of [Table 2](#): Significantly many participants reported the least difficulty of tracking collaborator's activities (CQ3), the strongest sense of their collaborator's presence (CQ4), the best communication quality (CQ5), the least difficulty of cooperating with collaborator (CQ7) happened in C_{pub} , whilst C_{fix} was thought conversely by significantly many participants. Their dislike of C_{fix} can also be seen in the interviews, in which the number of coded segments favouring C_{fix} and the number of segments reporting its advantages are the lowest, whilst the number of segments disfavouring it and the number of segments reporting its disadvantages are the highest among the four conditions.

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 provided in C_{fix} and C_{mov} , the augmented acoustic attenuation in C_{aug} enabled a closer collaboration, H3 is therefore supported. C_{aug} 's 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, possibly indicating a higher level of support. It should be noted that all the questions in PSQ were phrased either positively (PSQ1, PSQ2, PSQ3, PSQ4, PSQ7) or neutrally (the rest), with no negative statements, which might have affected participants' ratings positively. However, this imperfection has a limited influence on this study, because PSQ results are mostly used for comparisons between conditions, which are affected equally due to all of the conditions using the same phrasing. More insights regarding C_{aug} 's helpfulness are revealed by the thematic analysis, according to which, C_{aug} provided both "an appropriate background" with which participants felt "less stressed" and were able to "tailor" the individual composing to match the co-work, and a space personal enough to "work on something individually". 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 being introduced, whilst still providing sufficient support for individual creativity during collaboration. Thus H2 is validated.

Closer collaboration and higher consistency

According to measures of attention (AA16, AA17, AA18, AA19 in [Table 3](#)), compared with other conditions, in C_{aug} participants paid more close attention to their collaborator. Possible reasons for this can be found from the thematic analysis and other measures of Activity Assessments ([Table 3](#)). Compared with realistic acoustic attenuation in C_{pub} , C_{aug} 's augmented acoustic attenuation setting forced or prompted people to work closer in order to hear each other's work, as reported by some participants. Compared with

adding personal space with visible rigid boundary, by enabling participants to “decide” whether to hear other’s work or not “in a continuous way”, an invisible gradual boundary in C_{aug} led to less separation, and higher consistency between personal and public space, which matches the finding that people would like to be able to smoothly shift their artifacts from personal to public with intermediate shades in-between (*Greenberg, Boyle & LaBerge, 1999*). Compared with rigid personal spaces in C_{fix} and C_{mov} , C_{aug} saw more mutual note modifications, more group note edits, and larger group territory, a closer average distance between collaborators (respectively, AA4, AA7, AA5, and AA10 in [Table 3](#)), all of these indicate that C_{aug} saw a less separated collaboration than C_{fix} and C_{mov} . Compared with the three levels of privacy provided in UbiTable (*Shen, Everitt & Ryall, 2003*) and the binary levels of privacy provided in SharedNote (*Greenberg, Boyle & LaBerge, 1999*), the step-less sonic privacy provided by C_{aug} in this study possibly managed to better echo the suggestion that a boundary between personal and public space should be provided with gradations in subtle and lightweight ways to enable a fluid shift (*Greenberg, Boyle & LaBerge, 1999*), H3 is therefore supported.

Popularity

The code “advantage of C_{aug} ” has 111 coded segments, which is far more than the segments other codes have. Thirty-five coded segments are “most favourite— C_{aug} ”, higher than all other three conditions. These indicate C_{aug} is the most popular condition. This can also be partially verified by the preference measure. Specifically, C_{aug} has the highest preference rating in PSQ3 ([Table 1](#)), and more participants chose C_{aug} as the setting they most enjoyed in CQ1 (in [Table 2](#)). We believe the reasons behind this popularity are 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

This subsection compares C_{fix} with C_{mov} . The clear, sole difference between these two conditions is the mobility of personal space. In C_{fix} , to access personal spaces at the corners, participants needed to physically walk away from the centre and head to the corner, which might be the reason that C_{mov} saw a closer average distance between collaborators than C_{fix} (AA10, [Table 3](#)). This greater distance in C_{fix} possibly resulted the significantly larger size of personal territories (AA6) and more personal edits (AA8) in C_{fix} . On the contrary, the closer distance in C_{mov} created more chances for participants to pay or draw attention between each other, as a result, significantly longer time was spent paying attention to collaborators (AA17, AA19 in [Table 3](#)). With a closer average distance and more attention paid to each other, participants reported they had a marginal-significantly better quality of communication in C_{mov} (PSQ6 of [Table 1](#)). On the other hand, with participants being far away from each other and less chances for contact in C_{fix} , significantly many participants reported that they had the worst communication quality in C_{fix} (CQ6, [Table](#)

2). C_{mov} was also rated by much fewer participants to be the least enjoyable condition than compared with C_{fix} (CQ1, Table 2). Besides, C_{fix} also led to a reduced sense of collaborator's contribution (CQ9 in Table 3). As a possible result, C_{mov} saw a significantly more satisfying work output (PSQ5, Table 1).

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

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} minimised the negative impacts introduced by adding personal space (previously identified by *Men & 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 significantly less ordinary attention paid between collaborators.
- Providing personal space with a fluid boundary is preferable, it provides enough support for individual creativity with the minimal cost, and can even lead to a closer collaboration (specifically, greater attention was paid between collaborators).
- Compared with stationary personal space, space with mobility led to better communication, produced a better feeling of collaborator's contribution, had a higher rating in enjoyment, and produced a more satisfying output, and thus it supported collaboration better than stationary personal space.

DESIGN IMPLICATIONS

Based on the key points made above, we suggest five design implications for SVEs focusing on supporting creative collaboration (point 2 is for audio-related task only):

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

(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, Boyle & LaBerge (1999)*. The basis of providing light-weight privacy is not limited to audio, it can be provided by other modalitie(s) as well (e.g., visual). 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 example 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, participants 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 personal spaces are provided with gradual and adjustable boundary, manipulating the parameter of the boundary (e.g., the degree of augmented attenuation) can impact the level of “personalness” and therefore adjust the impact of introducing personal space. For example, the augmented attenuation can be set to a very low level if an extremely minimal impact is being pursued. So adding a method enabling users to adjust the level can allow users to shift between having a “very personal” space with total isolation where they could not hear nor see each other’s work, and having no personal space when they have to work together. 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 might also be useful as co-workers can then stay anywhere whilst still being able to adjust the “personalness” that 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, and a rigid-form personal has to be considered, it is better to provide rigid personal space with mobility, as the mobility feature gives users more freedom for accessing the personal spaces, and produces a better user

experience with fewer negative impacts on the collaboration compared with rigid personal space without mobility. This implication also echoes the proposal raised in our previous work (*Men & 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, augmented attenuation can support the individual creativity and the fluid shift between group activity and individual activity well during collaboration in SVE, with minimal negative impacts on collaboration introduced. 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 a 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.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology (EP/L01632X/1) and China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology:
EP/L01632X/1.

China Scholarship Council.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Liang Men conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, performed the computation work, authored or reviewed drafts of the paper, approved the final draft of the manuscript submitted for review and publication.

- Nick Bryan-Kinns conceived and designed the experiments, authored or reviewed drafts of the paper, approved the final draft of the manuscript submitted for review and publication.
- Louise Bryce authored or reviewed drafts of the paper, analyzed the data, approved the final draft of the manuscript submitted for review and publication.

Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

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

Data Availability

The following information was supplied regarding data availability:

The VR system of LeMo is available at: men, liang (2019): Unity Package for LeMo—A Multiplayer Music Making Prototype. figshare. Online resource. <https://doi.org/10.6084/m9.figshare.9816395.v1>.

Data is available at: Men, Liang (2019): Raw Data of Paper “Designing Virtual Spaces to Support Collaborative Creativity”. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.9820262.v1>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.229#supplemental-information>.

REFERENCES

- Altman I. 1975.** *The environment and social behavior: privacy, personal space, territory, and crowding*. Monterey: Brooks/Cole Pub. Co., Inc.
- Barbosa Á. 2003.** Displaced soundscapes: a survey of network systems for music and sonic art creation. *Leonardo Music Journal* 13:53–59 DOI 10.1162/096112104322750791.
- Basdogan C, Ho C-H, Srinivasan MA, Slater M. 2000.** An experimental study on the role of touch in shared virtual environments. *ACM Transactions on Computer-Human Interaction* 7(4):443–460 DOI 10.1145/365058.365082.
- Benford S, Greenhalgh C, Rodden T., Pycock J. 2001.** Collaborative virtual environments. *Communications of the ACM* 44(7):79–85.
- Billinghurst M, Kato H. 2002.** Collaborative augmented reality. *Communications of the ACM* 45(7):64–70.
- Billinghurst M, Poupyrev I, Kato H, May R. 2000.** Mixing realities in shared space: An augmented reality interface for collaborative computing. In: *2000 IEEE international conference on multimedia and expo. ICME2000. Proceedings. Latest advances in the fast changing world of multimedia (Cat. No. 00TH8532)*. Piscataway: IEEE, 1641–1644 DOI 10.1109/ICME.2000.871085.

- Bishop G, Fuchs H. 1992.** Research directions in virtual environments: report of an NSF Invitational Workshop, March 23–24, 1992, University of North Carolina at Chapel Hill. *ACM SIGGRAPH Computer Graphics* **26(3)**:153–177 DOI [10.1145/142413.142416](https://doi.org/10.1145/142413.142416).
- Blaine T, Fels S. 2003.** Contexts of collaborative musical experiences. In: *Proceedings of the 2003 conference on new interfaces for musical expression (NIME)*. National University of Singapore, 129–134.
- Braun V, Clarke V. 2006.** Using thematic analysis in psychology. *Qualitative Research in Psychology* **3(2)**:77–101 DOI [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa).
- Bressan F, Vets T, Leman M. 2017.** A multimodal interactive installation for collaborative music making: from preservation to enhanced user design. In: *Proceedings of the European society for cognitive sciences of music (ESCOM) conference, Ghent University*. 23–26.
- 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*. ACM, 135–144.
- Bryan-Kinns N. 2013.** Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies* **71(1)**:76–90 DOI [10.1016/j.ijhcs.2012.02.004](https://doi.org/10.1016/j.ijhcs.2012.02.004).
- Bryan-Kinns N, Hamilton F. 2012.** Identifying mutual engagement. *Behaviour & Information Technology* **31(2)**:101–125 DOI [10.1080/01449290903377103](https://doi.org/10.1080/01449290903377103).
- Bryan-Kinns N, Healey PG, Leach J. 2007.** Exploring mutual engagement in creative collaborations. In: *Proceedings of the 6th ACM SIGCHI conference on creativity & cognition*. New York: ACM, 223–232.
- Carlsson C, Hagsand O. 1993.** DIVE A multi-user virtual reality system. In: *Proceedings of IEEE virtual reality annual international symposium*. Piscataway: IEEE, 394–400.
- Case DA, Ploog BO, Fantino E. 1990.** Observing behavior in a computer game. *Journal of the Experimental Analysis of Behavior* **54(3)**:185–199 DOI [10.1901/jeab.1990.54-185](https://doi.org/10.1901/jeab.1990.54-185).
- Fencott R, Bryan-Kinns N. 2010.** Hey man, you're invading my personal space! privacy and awareness in collaborative music. In: *Proceedings of the 2010 conference on new interfaces for musical expression (NIME)*. Available at http://www.eecs.qmul.ac.uk/~nickbk/papers/P198_Fencott.pdf.
- Gaver WW. 1992.** The affordances of media spaces for collaboration. In: *Proceedings of the 1992 ACM conference on computer-supported cooperative work*. ACM, 17–24.
- Greenberg S, Boyle M, LaBerge J. 1999.** PDAs and shared public displays: Making personal information public, and public information personal. *Personal Technologies* **3(1–2)**:54–64 DOI [10.1007/BF01305320](https://doi.org/10.1007/BF01305320).
- Hall ET. 1966.** *The hidden dimension*. Vol. 609. Garden City: Doubleday.
- Hendrix C, Barfield W. 1996.** The sense of presence within auditory virtual environments. *Presence: Teleoperators & Virtual Environments* **5(3)**:290–301 DOI [10.1162/pres.1996.5.3.290](https://doi.org/10.1162/pres.1996.5.3.290).
- Kreijns K, Kirschner PA, 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
[DOI 10.1016/S0747-5632\(02\)00057-2](https://doi.org/10.1016/S0747-5632(02)00057-2).
- Kruger R, Carpendale S, Scott SD, Greenberg S. 2004.** Roles of orientation in tabletop collaboration: comprehension, coordination and communication. *Computer Supported Cooperative Work* **13**(5–6):501–537 [DOI 10.1007/s10606-004-5062-8](https://doi.org/10.1007/s10606-004-5062-8).
- Lea R, Honda Y, Matsuda K. 1997.** Virtual society: collaboration in 3D spaces on the internet. *Computer Supported Cooperative Work* **6**(2):227–250
[DOI 10.1023/A:1008672915465](https://doi.org/10.1023/A:1008672915465).
- Luciani A. 2007.** Virtual reality and virtual environment. In: *Enaction and enactive interfaces: a handbook of terms*. Grenoble: Enactive Systems Book ACROE, 299–300.
- Men L, 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)*. Piscataway: IEEE, 1–6 [DOI 10.1109/SIVE.2018.8577094](https://doi.org/10.1109/SIVE.2018.8577094).
- Men L, Bryan-Kinns N. 2019.** LeMo: exploring virtual space for collaborative creativity. In: *Proceedings of the 2019 on creativity and cognition, C&C '19*. New York: ACM, 71–82 [DOI 10.1145/3325480.3325495](https://doi.org/10.1145/3325480.3325495).
- Men L, Bryan-Kinns N, Hassard AS, Ma Z. 2017.** The impact of transitions on user experience in virtual reality. In: *2017 IEEE virtual reality (VR)*. 285–286
[DOI 10.1109/VR.2017.7892288](https://doi.org/10.1109/VR.2017.7892288).
- Nassiri N, Powell N, Moore D. 2010.** Human interactions and personal space in collaborative virtual environments. *Virtual Reality* **14**(4):229–240
[DOI 10.1007/s10055-010-0169-3](https://doi.org/10.1007/s10055-010-0169-3).
- Nedel L, De Souza VC, Menin A, Sebben L, Oliveira J, Faria F, Maciel A. 2016.** Using immersive virtual reality to reduce work accidents in developing countries. *IEEE Computer Graphics and Applications* **36**(2):36–46.
- Pinelle D, Gutwin C, Greenberg S. 2003.** Task analysis for groupware usability evaluation: modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction* **10**(4):281–311
[DOI 10.1145/966930.966932](https://doi.org/10.1145/966930.966932).
- Raffestin C. 2012.** Space, territory, and territoriality. *Environment and Planning D: Society and Space* **30**(1):121–141 [DOI 10.1068/d21311](https://doi.org/10.1068/d21311).
- Rodden T. 1991.** A survey of CSCW systems. *Interacting with Computers* **3**(3):319–353
[DOI 10.1016/0953-5438\(91\)90020-3](https://doi.org/10.1016/0953-5438(91)90020-3).
- Roussos M, Johnson AE, Leigh J, Vasilakis CA, Barnes CR, Moher TG. 1997.** NICE: combining constructionism, narrative and collaboration in a virtual learning environment. *ACM SIGGRAPH Computer Graphics* **31**:62–63
[DOI 10.1145/262171.262264](https://doi.org/10.1145/262171.262264).
- Sack RD. 1983.** Human territoriality: a theory. *Annals of the Association of American Geographers* **73**(1):55–74 [DOI 10.1111/j.1467-8306.1983.tb01396.x](https://doi.org/10.1111/j.1467-8306.1983.tb01396.x).
- Schroeder R. 2002.** Social interaction in virtual environments: key issues, common themes, and a framework for research. In: *The social life of Avatars: presence and interaction in shared virtual environments*. Berlin, Heidelberg: Springer-Verlag, 1–18.

- Schubert T, Friedmann F, Regenbrecht H. 2001.** The experience of presence: factor analytic insights. *Presence: Teleoperators & Virtual Environments* **10(3)**:266–281 DOI [10.1162/105474601300343603](https://doi.org/10.1162/105474601300343603).
- Scott SD, Carpendale S. 2010.** Theory of tabletop territoriality. In: *Tabletops-horizontal interactive displays*. London: Springer, 357–385.
- Scott SD, Carpendale MST, Inkpen KM. 2004.** Territoriality in collaborative tabletop workspaces. In: *Proceedings of the 2004 ACM conference on computer supported cooperative work*. New York: ACM, 294–303.
- Shen C, Everitt K, Ryall K. 2003.** UbiTable: impromptu face-to-face collaboration on horizontal interactive surfaces. In: Dey AK, Schmidt A, McCarthy JF, eds. *UbiComp 2003: ubiquitous computing. UbiComp 2003. Lecture notes in computer science*, vol. 2864. Berlin, Heidelberg: Springer, 281–288.
- Steuer J. 1992.** Defining virtual reality: dimensions determining telepresence. *Journal of Communication* **42(4)**:73–93.
- Sugimoto M, Hosoi K, Hashizume H. 2004.** Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York: ACM, 41–48.
- Tang JC. 1991.** Findings from observational studies of collaborative work. *International Journal of Man-machine studies* **34(2)**:143–160 DOI [10.1016/0020-7373\(91\)90039-A](https://doi.org/10.1016/0020-7373(91)90039-A).
- Titon JT, Slobin M. 1996.** The music-culture as a world of music. In: *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, Barthet M. 2018.** Internet of musical things: Vision and challenges. *IEEE Access* **6**:61994–62017 DOI [10.1109/ACCESS.2018.2872625](https://doi.org/10.1109/ACCESS.2018.2872625).
- Wang G. 2009.** Designing Smule's Ocarina: the iPhone's magic flute. In: *Proceedings of the 2009 conference on new interfaces for musical expression (NIME)*. 303–307.
- Xambó A, Hornecker E, Marshall P, Jordà S, Dobbyn C, Laney R. 2013.** Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction* **20(6)**:Article 36.
- Yin RK. 2017.** *Case study research and applications: design and methods*. Thousand Oaks: Sage Publications.
- Zhang X, Furnas GW. 2003.** The effectiveness of multiscale collaboration in virtual environments. In: *CHI'03 extended abstracts on human factors in computing systems*. New York: ACM, 790–791.