



Social Virtual Reality Platform Comparison and Evaluation Using a Guided Group Walkthrough Method

Qiaoxi Liu¹ and Anthony Steed^{2*}

¹UCL Interaction Centre, University College London, London, United Kingdom, ²Department of Computer Science, University College London, London, United Kingdom

As virtual reality (VR) headsets become more commercially accessible, a range of social platforms have been developed that exploit the immersive nature of these systems. There is a growing interest in using these platforms in social and work contexts, but relatively little work into examining the usability choices that have been made. We developed a usability inspection method based on cognitive walkthrough that we call *guided group walkthrough*. Guided group walkthrough is applied to existing social VR platforms by having a guide walk the participants through a series of abstract social tasks that are common across the platforms. Using this method we compared six social VR platforms for the Oculus Quest. After constructing an appropriate task hierarchy and walkthrough question structure for social VR, we ran several groups of participants through the walkthrough process. We undercover usability challenges that are common across the platforms, identify specific design considerations and comment on the utility of the walkthrough method in this situation.

OPEN ACCESS

Edited by:

Nadia Magnenat Thalmann,
Université De Genève, Switzerland

Reviewed by:

Aryabrata Basu,
Emory University, United States
Mark Billingham,
University of South Australia, Australia

*Correspondence:

Anthony Steed
A.Steed@ucl.ac.uk

Specialty section:

This article was submitted to
Virtual Reality and Human Behavior,
a section of the journal
Frontiers in Virtual Reality

Received: 15 February 2021

Accepted: 22 April 2021

Published: 24 May 2021

Citation:

Liu Q and Steed A (2021) Social Virtual Reality Platform Comparison and Evaluation Using a Guided Group Walkthrough Method.
Front. Virtual Real. 2:668181.
doi: 10.3389/frvir.2021.668181

Keywords: collaborative virtual environment, cognitive walkthrough, usability inspection methods, consumer virtual reality, social virtual reality

1 INTRODUCTION

Because of their body tracking and thus their ability to support fluid non-verbal communication alongside verbal communication, immersive virtual reality (VR) systems have attracted a lot of attention for their potential to support remote collaboration. Some of the earliest commercial VR systems from the late 1980s and early 1990s supported collaboration (Churchill and Snowdon, 1998). For example, the Reality Built for Two system from VPL Research (Blanchard et al., 1990). The recent availability of affordable and effective consumer VR systems based on head-mounted displays (HMDs) has unleashed a wave of creative development that includes many and varied social experiences. With travel and socialization currently being curtailed, these systems have great potential for enriching our personal and professional lives.

To meet this potential, as human-computer interfaces, the social VR systems should be usable. Because the design space of immersive content is very large, the current social VR systems are diverse both in the way the tasks that they support and in their overall visual and interaction style. There is a very significant body of work around specific aspects of the design of social VR systems, such as the role of avatar representation (Biocca et al., 2003; Schroeder, 2010; Latoschik et al., 2017; Kolesnichenko et al., 2019), but such work tends to be done in relatively controlled social situations. Surveys of technical functionality and user affordances go some way to helping

identify usability issues or lack of functionality in specific systems by contrasting the design choices (e.g., (Tanenbaum et al., 2020)).

In this paper we propose to adapt a usability walkthrough method to provide a different way of revealing usability issues in these platforms. We identify that there are common patterns of behavior that social VR systems must support in order to enable group formation, group maintenance, communication and task monitoring. Thus there are common tasks that all the platforms need to support (e.g., rendezvous in a user-chosen location). Usability walkthrough methods such as cognitive walkthrough (Lewis et al., 1990), pluralistic walkthrough (Bias, 1994) and group walkthrough (Pinelle and Gutwin, 2002) (though see **Section 2.3** for other methods) are not typically used as comparative methods, but because social VR platforms need to support these common tasks there is an interesting opportunity to take the formative feedback that would be generated by applying a walkthrough method to one platform, see how it applies to all platforms, and compare and contrast the results. We call our method *guided group walkthrough* as an expert walks through the tasks with other users acting as the group members.

Thus main research question is thus whether there are common usability problems across a set of current social VR platforms. We focus exclusively on immersed users and chose to focus on social VR systems for Oculus Quest, see **Section 3.2** for discussion of the scope. We chose six to compare: AltspaceVR, RecRoom, VRChat, Bigscreen, Spatial and Mozilla Hubs (see **Table 1**). We find that there are common problems across the six platforms that provide challenges to the platform providers and researchers in the field. We also find specific problems on one or a small number of platforms that are design choices that might be avoided in future.

A secondary contribution is the guided group walkthrough method itself. The cognitive walkthrough method was originally designed for an expert to walkthrough through a user interface solo (Lewis et al., 1990). The method has been extended and adapted in many ways previously and has been applied to desktop collaborative virtual environments and groupware systems, see **Section 2.3**. We note that these approaches deal in slightly different ways with the relative freedom that a user has in undertaking a task, and that immersive interfaces exacerbate some of these concerns. We thus re-analyse the task structure that underlies the walkthrough and propose modified sets of questions to ask users at each step of the walkthrough. The full sets of tasks and questions are provided in **Supplementary Material**.

Finally we take the findings and tie them to recent threads of research about adoption of these platforms, platforms comparisons, work on embodiment, etc. and make some recommendations for implementation of these platforms going forward.

2 BACKGROUND

2.1 Virtual Reality and Social Virtual Reality

Immersive VR (henceforth just “VR” unless this needs clarification) systems have long been of interest as interfaces

because the person is mostly encompassed by the displays with those displays presenting from a first-person point of view. Users of VR systems can control the display by moving their bodies. When presented with virtual environments mimicking plausible real situations, users can show behaviors similar to those they might exhibit in a similar real situation, a phenomenon sometimes referred to as *presence* (Sheridan, 1992; Slater, 2009). Thus a lot of evaluations of VR have focused on determinants of presence or how to maximize presence response as it is hypothesized that this is the key determinant of success of VR. Threads of research have looked at either observing user responses such as reacting as if something was real (e.g., was shocking such as a virtual drop (Usoh et al., 1999)), or asking them about their ability to interact with the world as treating this as the indicator of presence (Witmer and Singer, 1998).

Recent work has tied presence in VR to the notion of embodiment as investigated within neuroscience (Slater et al., 2008; Gonzalez-Franco and Peck, 2018). A range of fascinating studies have shown how VR users can believe that the body that they see inside the VR scene in the location of their real body (e.g., when they look down) is treated their own body, and that this can change their perception of themselves (e.g., (Kilteni et al., 2012; Maister et al., 2015)). This then leads to questions of how the user feels agency over their actions, and how representation of their body supports their interactions with the virtual environment (Argelaguet et al., 2016).

Collaborative virtual environments (CVE) is a term that covers a broad class of desktop and immersive VR systems that support collaboration in a common virtual environment where each participant is represented by an avatar (Blanchard et al., 1990; Stone, 1993; Mantovani, 1995; Damer et al., 1997). The avatars allow people to see each other and to see how each other are positioned relative to other objects (Hindmarsh et al., 2000). As soon as the participants embody avatar representations, they can exhibit non-verbal behaviors either implicitly or explicitly (Fabri et al., 1999). This ability in immersive VR to communicate directly through gesture, or in desktop VR to communicate indirectly through activating animations, has generated a very significant body of work on the role of avatars and the adoption of roles for communication and gesturing (e.g., (Yee et al., 2007; Pan and Steed, 2017; Moustafa and Steed, 2018)). We will use the term social VR to refer specifically to immersive systems that prioritize and focus on the in-environment communication, rather than desktop systems or systems that use a CVE alongside other tools.

Turning back to immersive VR, copresence, or social presence, is the feeling of being with other people (Biocca et al., 2003). It has a natural correspondence with presence: if the user feels that they are present in the virtual space, then representations of humans or objects that act like humans, should be treated as if they were human. It has been found that observed social rules from real environment transfer to the VR, and indeed interacting with others can enhance presence (Hai et al., 2018). The interaction between embodiment, social action and its impact on users is now a very active area of research.

2.2 Usability of Virtual Reality

Usability for VR has a close relationship with the issues of presence and co-presence: the goal might be to experience co-presence with users so one requirement of the system is to facilitate that. In some situations this can be solely as a matter of facilitating communication through verbal and non-verbal communication (e.g. (Tanenbaum et al., 2020)). However VR systems do have to include other user interface components to facilitate tasks. These include manipulation, locomotion and system control (e.g., menus) (LaViola et al., 2017). These compensate for the fact that although VR can be said to have a metaphor of non-mediation or direct manipulation, in that the participant uses their own body to interact, there are limitations: the person can't walk long distances; manipulation of objects is typically done using gestures or button clicks rather than full force feedback on the hand; and system controls are necessary to control environment or behaviors of the world, such as loading new worlds, changing avatar clothes, etc. There are multiple ways of implementing these interface components and we should consider the interplay of usability of user-interface elements within the VR as well as the way in which components represented to oneself and others. We will find in the task breakdown (see **Section 3.4**) that even in immersive applications that support primarily 3D interactions and there are many tasks that are effectively done in 2D because they utilize a traditional-looking 2D menu system.

Thus VR systems are within scope of usability practice including aspects of learnability, effectiveness and satisfaction (Gabbard, 1997; Kalawsky, 1999; Marsh, 1999; Sutcliffe and Kaur, 2000). Stanney et al. (2003) noted that traditional usability techniques tend to overlook issues specific to VR, including the 3D nature of the atomic tasks of interaction, immersive and multimodal nature of displays, the impact of presence and the multi-user nature of many VR systems. Geszten et al. have more recently classified the usability factors into three classes: concerning the virtual environment, concerning the device interaction and issues specific to the task within VR (Geszten et al., 2018). Specifically for device interaction a whole field of study on 3D user interfaces has concerned itself with efficient ways to effect locomotion, manipulation and systems control within 3D spaces (Bowman et al., 2002; LaViola et al., 2017).

As an example of a specific usability method adapted to the desktop CVE situation, Sutcliffe and Kaur (Sutcliffe and Kaur, 2000) extended cognitive walkthrough (Lewis et al., 1990; Nielsen, 1994; Hollingsed and Novick, 2007) to assess a desktop VR business park application. A walkthrough method is a structured method of having a user interact with the system asking specific questions at each stage of a task. It is one of a class of expert or guided review methods that are typically used to provide feedback before testing at scale with users (Hartson et al., 2001; Mahatody et al., 2010). In developing their method Sutcliffe and Kaur noted the freedom that users had to locomote about the environment. Thus engagement with any particular interface component could be suspended by moving away, and users needed to explore and approach different components. They thus broke down the interaction into three cycles: task action

cycle, navigation cycle and system initiative cycle; each cycle contains a slightly different set of walkthrough questions. We build upon work described below that extended this to the CVE situation.

2.3 Usability of Social Virtual Reality

There are many social VR systems out there. If we include both desktop and immersive VR systems, over 150 publicly available systems are cataloged in Schulz's blog about social VR (Schulz, 2020) and many more are constructed for experimentation in research labs. The usability of a social VR has many angles. Recent surveys have focused on cataloging the scope of the interface (Jonas et al., 2019), the types of avatar supported (Kolesnichenko et al., 2019) or specific functionality such as facial expressions (Tanenbaum et al., 2020). Other recent work has focused on longer term relationships (Moustafa and Steed, 2018) in social VR, development of trust in social with others (Pan and Steed, 2017) or how users perceive their avatars (Freeman et al., 2020).

Our goal is to take a broader view of usability and look at the issues found as a small group of users exercise the functionality of the interface. We extend the approach of Tromp et al. (2003) who based their work on that of Sutcliffe and Kaur (2000). In their cognitive walkthrough, Tromp et al. added a collaborative cycle to represent those tasks where users interacted. This is a similar tactic taken by other researchers studying the broad category of groupware systems such as the work of Pinelle and Gutwin (2002). They developed a groupware walkthrough by analyzing teamwork and using a hierarchical task model to break tasks down. However, their questions regarding the tasks are focused mainly on effectiveness and satisfaction. Our own work is inspired by the Tromp et al. work as it systematizes the exploration of features of the social VR. Their work dealt with desktop VR systems, so our explorations will be sensitive to the new issues raised by the social VR on avatars and embodiment. As our walkthrough technique involves representative users, it has similarities with group walkthrough and pluralistic walkthrough methods which have been developed and used in a broad range of domains (e.g., see (Bias, 1994; Hollingsed and Novick, 2007; Mahatody et al., 2010; Jadhav et al., 2013)).

3 CONSTRUCTING THE WALKTHROUGH

3.1 Overview

We want to uncover the broad issues of usability encountered when interacting with current social VR platforms. We have chosen six, described in **Section 3.2**. On each platform, we want to investigate the scenario of one person meeting with one or two other people. Thus the main task is simply to meet, then communicate with each other, interact with each other, interact with the scene and then move to another scene. In order to understand and compare the design choices of different social VR platforms, we choose to use a walkthrough method. We try to keep the tasks as natural as possible. Even though the main task is quite straightforward, as we will see in the task analysis (see **Section 3.3**) it already presents a significant set of complexities once expanded to individual task actions.

TABLE 1 | The six social VR platforms chosen for the study.

Platform	Key features
VRChat (VRChat Inc, 2021)	Cross-platform (PC-VR, Quest and desktop). Avatars selected from a very diverse set. More of a focus on user-created content than other platforms
RecRoom (Rec. Room, 2021)	Cross-platform (PC-VR, Quest and desktop). Avatars selected from a relatively constrained style of cartoon-like styles and options. A variety of social spaces with different social activities including games
AltSpaceVR (Microsoft, 2020)	Cross-platform (PC-VR, Quest and desktop). Avatars selected from a relatively constrained style of cartoon-like styles and options. More of a focus on events than some other platforms
BigScreen (Bigscreen, 2021)	Cross-platform (PC-VR, Quest and desktop). Avatars selected from a relatively constrained style of cartoon-like styles and options. Specifically focused on screen sharing (applications or movies)
Spatial (Spatial Systems, 2021)	Cross-platform (PC-VR, Quest and desktop) and also on some AR systems. Avatars are built from a photo of the user's face. Focused around meetings to share media
Mozilla Hubs (Mozilla Corporation, 2021)	A web-based platform that can be explored in some WebXR-enabled web browsers. Supported on Quest. Quite open, but main demos include a small set of avatars and simple worlds that can be shared

We develop a walkthrough method designed to reflect the specific issues of social VR. As noted this extends the work of Tromp et al. (2003) by updating the task analysis to support immersive systems and focusing on groups of user representatives rather than experts performing the walkthrough, as is common in pluralistic walkthroughs (Bias, 1994). Further, our use of the method focuses on common social tasks rather than application specific tasks. Finally, we use the method to explore similarities and differences between platforms that implement these common social tasks.

3.2 The Social Virtual Reality Platforms

We chose to focus on social VR platforms for the Oculus Quest available in the United Kingdom in June/July/August 2020. We excluded multi-player games with a single game focus because we wanted to focus on general social tasks that would not be possible within the constraints of most single games (e.g., group navigation around large environments). We also excluded systems without verbal communication (e.g., Half + Half and The Under Presents) because we considered talking to be a primary task. We also excluded asymmetric systems designed to be played by single immersed players but with others on non-immersive interfaces or external (e.g., Puppet Theater and Keep Talking and No-One Explodes). This left five main published software platforms available at the time of writing as listed in **Table 1**. We added a sixth platform, Mozilla Hubs, which is based on WebXR technologies and thus runs within the web browser on the Oculus Quest. Other web-based platforms would be possible, but none was as mature as Mozilla Hubs. The Oculus Quest was chosen as the hardware platform because the study needed to be run off the university campus as the buildings with appropriate laboratories were closed due to Covid-19 in Q2/Q3 2020. The choice of Oculus Quest was also motivated by its low cost and its being relatively easy to provision and maintain the three devices needed for the study because they are self-contained wireless systems. All of the platforms also support PC-based VR, desktop-style interfaces or augmented reality, and there are many more social VR systems on the PC, but the issues we uncover would apply both to the corresponding PC-based VR interface, because they are substantially the same, or are issues common to the form of the shared environment, such as having avatars acting at a distance (see **Section 5.1.5**).

3.3 Task Analysis

A task analysis was done to form the basis of walkthrough. User task analysis can provide representative user scenarios by defining and ordering user task flows (Bowman et al., 2002). Because we wanted to cover the generic tasks in social VR, the following high-level tasks are covered:

1. Find your offline or online friend in the platform and meet in a private/public room, make friends and interact with each other.
2. Talk about your experience and share related sources, using the objects in the room or in the system to express your emotions.
3. Move to another room together.

For desktop CVEs Tromp et al. (2003) used a hierarchical task analysis to classify the tasks into four groups: navigate, find other users, find interactive objects and collaborate. However, we extended this to cover the broader range of functionality that is typically available in modern systems and the nature of immersive systems. We thus worked with five task groups:

1. *Identification* which is all the actions to identify a target, which could be a person or an object.
2. *Communication* which includes verbal and non-verbal communication, and in this case, interpersonal interaction is covered.
3. *Navigation* means how people plan a route and then how they locomote in space or between the rooms, which includes movement of both individuals and groups. Group moving means moving as a group to the same place. Room transport refers to how people teleport to a new room.
4. *Manipulation* mainly relates to the interactive objects, such as creating, moving, passing and joint manipulating.
5. *Coordination* is the action to gather people, handle conflict, and plan actions.

In order to make the platforms comparable, the tasks chosen had to be supported by all of them. Though they have different ways of functioning, they do all target the general goal of social activities. Based on the previous work and the self-walkthrough, the tasks are listed in **Table 2**. If a task is not supported on a

TABLE 2 | The full set of tasks within each of five main task groups (left column), and which interaction cycle they use for each platform.

	VRChat	RecRoom	AltSpaceVR	BigScreen	Spatial	Mozilla hubs
Identification						
Identify others	Explore	Explore	Goal +3D	Explore	Explore	Explore
Identify speaker	Explore+2D	Explore	Explore+2D	Explore+2D	Explore+2D	Explore
Identify interactor	Explore	3D	Explore	Explore	Explore	Explore
Communication						
Emotion express	2D	2D	2D	–	–	2D
Gesture communicate	Collaboration/2D	Collaboration	Collaboration	Collaboration	Collaboration	Collaboration
Mark friends	2D + Switch	2D + Switch	2D + switch	–	–	–
Text communicate	System+2D	System+2D	2D	–	–	System
Navigation						
Group gathering	Goal	Goal	Goal	Goal	Goal	Goal
Group moving	Goal/2D	2D	Goal/2D	2D	2D	-
Room transport	2D	2D/Switch	2D	2D/Switch	2D/Switch	Switch
Manipulation						
Creating objects	–	2D+3D	–	2D	2D	2D
Moving objects	3D	3D	3D	3D	3D	3D
Passing objects	Collaboration	Collaboration	Collaboration	Collaboration	Collaboration	Collaboration
Coordination						
Room creating	2D	2D	2D	2D	2D	-
Invite others	2D + System	2D/Switch	2D + System	Switch	2D + Switch	Switch
Public room meeting	2D + Goal	2D + Goal	2D + Goal	2D + Goal	–	-
External source sharing	-	-	-	Switch	Switch	-
System notification	System	System	System	-	-	-

platform, the relevant cell is blank, otherwise the task is supported in one of the interaction cycles as described below in **Section 3.4**. For some platforms, the function may not be available directly, but users could get the same result through other approaches. For example, *group moving* means to enter a new space as a group; in other words, all the users enter a space at the same time. In VRChat, a user can create a portal which allows all the users transport simultaneously, but in RecRoom, this mechanism is missing, so instead users would enter the room and then invite others to join them. Note that a task might be achievable in two different ways. For example, “Collaboration/2D” means that the participant could use either a Collaboration or 2D cycle. In the walkthrough the experimenter asked the questions appropriate for the cycle the participant opted for. Alternatively, a task might require two different styles of interaction, For example “System + 2D” implies that both 2D and Switch cycles are necessary to complete the tasks and thus the experimenter asked both sets of questions.

3.4 Interaction Cycles

Interaction cycles are used to model all the steps of interaction to predict behaviors and requirements for successful interaction (Nielsen, 1994). The cycles provide sets of fixed questions for inspecting the potential usability problems related to a particular action. Drawing on Tromp et al.’s work (Tromp et al., 2003), six interaction cycles were initially used.

- Normal task cycle 2D is used when a user is interacting with 2D interfaces (e.g., the menu, or pop-up windows). These are flat, or near flat, menus that are presented in-environment for the user to interact with (2D in **Table 2**).

- Normal task cycle 3D is used when a user is interacting with a 3D object in order to achieve a goal (3D in **Table 2**).
- Goal-directed exploration cycle is using when a user is searching for a certain target in the environment (Goal in **Table 2**).
- Exploratory browsing cycle is used when a user explores the system out of curiosity and seeking a greater understanding of the world (Explore in **Table 2**).
- Collaboration cycle is used when a user is interacting with other users. According to the different designs of tasks in the various platforms, different cycles should be linked. In self-walkthrough, the questions from six cycles are used and checked for their validity (Collaboration in **Table 2**).
- System initiative cycle is used when there are system prompts or events to take over the control from the user (Note that this cycle is replaced below.)

One of the authors did a full self-walkthrough of all platforms and all tasks from the left column of **Table 2** in order to complete the main body of that table which indicates which cycle each tasks corresponds to on each platform. This table was verified by the other author. For example, all platforms, except BigScreen and Spatial, allow text communication between friends. Text communication involves three actions: send texts, receive texts and check text board. Sending text and checking text board are done via the menu and link to cycle 2D. The main difference between platforms is the way in which the receiver of a text is notified. There are two types: one a pop-up floating window with text; the other is to notify the receiver via a notification and require the user to check the menu. For the floating window and menu, cycle 2D is thus applicable. The Supplemental Materials contains a more detailed discussion of major differences between

the platforms which are summarized in **Supplementary Tables 5–8**.

In that self-walkthrough we found that System Initiative Cycle was in-appropriate as within these social VR platforms there were no instances of the system taking over completely the display. This is probably because of the lack of control the user would have over their actions if the display was locked because the system is fully immersive. Thus this cycle was removed. We replaced this with a cycle System Alert (just “System” in **Table 2**), which represented important messages that the user might take notice of and react to. These messages are more akin to the notifications on a smart phone or desktop display as they are purely informational, non-modal and also do not take up much space. The walkthrough questions for this cycle are:

- Can user receive feedback about system status/changes?
- Is the system notification visible?
- Can the user notice the system notification?
- Can user understand the system information?
- Can the user decide what action to do next?
- Can user keep informed about system status?

Finally, *Switch* in **Table 2** indicates that the task might only be achieved outside the VR. This was not considered in the walkthrough as the interactions become very complicated, but we note instances of the action in the results. The full set of walkthrough questions for each of the six cycles is presented in **Supplementary Table 9**.

4 STUDY DESIGN

Having constructed and validated a walkthrough procedure that worked across the different platforms, we then moved to a user study. Notable here is that each participant tests multiple platforms and we are using the walkthrough results to compare between platforms. Another key feature is that the participants and experimenter might meet inside the social VR space to conduct the walkthrough.

4.1 Participants

Overall, 17 participants (8 males and 9 females, with an average age of 22.7) were recruited by advertising at University College London by online advertisements. Participants were assigned to groups of two or three according to their familiars. All the groups are based on the relationship of friends or a couple. None of them had prior VR experience. Some of them took part several times with different platforms, and all were incentivized with £10 (GBP) per hour payment for taking part. The study was approved by the UCL Research Ethics Committee.

4.2 Materials

A total of three Oculus Quests were used for the study with the latest versions, as of July 2020, of the five applications installed. Mozilla Hubs was accessed through on the on-board web browser. The devices were delivered to participants before every trial.

TABLE 3 | Arrangement of participants (P1-P17) and platforms. Note that R refers to the researcher (one of the authors).

Platform	Trial set 1	Trial set 2	Trial set 3
VRChat	P1,R	P4,P5	P16,P17,R
RecRoom	P1,R	P6,P7	P16,P17,R
AltSpace	P2,R	P8,P9	P16,P17,R
BigScreen	P2,R	P10,P11	P16,P17,R
Spatial	P3,R	P12,P13	P16,P17,R
Mozilla Hubs	P3,R	P14,P15	P16,P17,R

4.3 Procedure

To ensure a high quality of evaluation of different platforms, participants were arranged in groups to experience multiple platforms (**Table 3**). In the trials in Set 1, one participant met with the researcher to do a walkthrough of all the tasks. This was to check the validity of the questions and any operational issues working in pairs immersively. The participant and researcher thus were both represented as avatars in the virtual environments. The participant and researcher were located in different places in the real environment.

In the trials in Set 2, groups of two participants met on the platform. The researcher did not join through the software, but all three (two participants and researcher) were in the same place in the real environment. Thus the researcher could talk directly to the participants, not through the platform.

In the trials in Set 3 the researcher joins in with the interactive task with two others. Thus all three met on the social VR platform. The participants and researchers were in different real places, so voice communication was through the platform. In pilot trials, we attempted to coordinate three participants in trials, but this proved unwieldy. It worked well to have the researcher both act as a participant and ask the questions in Set 3.

Participants were required to take part in the Oculus Quest’s basic tutorial if they hadn’t had any experience of VR. They only needed to do this once. For every new platform, participants were required to complete the tutorial of the platform.

Participants, and in Sets 1 and 3, the researcher, met in private spaces in the platforms. In some platforms (Mozilla Hubs, Spatial), because of the way location works, it was essentially impossible for the participants to travel to a public space. In VRChat, AltSpace, RecRoom and Big Screen, the participants could navigate to public spaces. If they did so, the participants were instructed to travel back to a private space.

Participants were free to determine the order in which they complete the tasks under analysis (see **Section 3.1** and leftmost column of **Table 2**). Some sub-tasks obviously go together (e.g., identifying the others, and identifying the speaker). The experimenter chose the specific walkthrough questions to ask selected depending on the sub-task undertaken by the participant and how this mapped to an interaction cycles (refer to the columns in **Table 2**). The full list of questions in **Supplementary Table 9**). They also insured that the participants completed all the tasks. In general the experimenter asked the questions at the end of each task; but if the participant got stuck then they asked the relevant questions at the point in the cycle to prompt the user.

The first-person perspective view of each participant was recorded during the sessions using the Oculus Quest's built-in recording software. The answers to each question and the first-person recordings are the primary sources of analysis. Usability problems were identified by critical responses from the participants.

4.4 Interviews

In addition, in order to gain more insights about presence and co-presence, at the end of sessions, participants in the group trials were asked the following questions:

- When you move your body, do you feel the avatar in the system act the same as you?
- Does the system follow real world conventions?
- Do you want/have to move physically when you want to move in the platform?
- Do you feel you are in the environment? What's the level of your presence?
- When you face some problems or glitches, do you know what to do?
- When you are talking, do you know others are listening?
- When trying to manipulate the same objects, do you know who is interacting with it?
- When do you feel uncomfortable most?

5 RESULTS

Sessions varied in length between 30 and 90 min due to the needs of participants to do the Oculus Quest tutorial and/or the social VR platform's tutorial. Walkthroughs in Set 1 took an average of 30 min as they were one participant with the experimenter. Walkthroughs in Sets 2 and 3 took an average of 70 min as they involved two participants and had interactions with the experimenter. A total of 18 h of active walkthrough was recorded and then analyzed. This 18 h does not include recordings of the tutorials. **Supplementary Table 4** Gives a summary of the number of questions asked across the different cycles and tasks.

Since we have three trials of each platform and five participants experiencing each platform (see **Table 3**), problems reported by more than two participants in different groups are identified as usability problems. The logging and summarization of the problems from transcripts and videos was done by the first author. This initial list of problems was reviewed by both authors and summarized into the categories below. Comparing the usability problems of the six platforms, those problems reported in more than half of them are defined as common problems, while others specific in each platform are classified as particular problems.

5.1 Common Problems

We found none of the usability problems in the platforms to be severe enough to prevent progress or cause the termination of a session. The most common problems were confusion and hesitation over how to locate and activate features. Several problems were reported in all the platforms and came up

repeatedly. Some of them have no obvious solutions and raise interesting future research questions.

5.1.1 Communication Coordination

Verbal chat is typically restricted to people within the same room. While this is an obvious metaphor and platform implementation strategy, it meant participants were unable to communicate before they entered the same virtual room. In group trials, an extra instant communication tool had to be used for the participants to coordinate and keep track of each other. When communicating in a room, it was sometimes difficult to tell whether the intended audience is listening. For the platforms supporting text communication, participants were concerned whether others had read their messages, so participants tended to use another communication tool to inform others and then ensure they have received the text.

For a group of three, participants had difficulty in following threads. There was no indication to show one's speaking intentions, which led participants to talk over each other or unexpected silences.

Due to the environment complexity, the accuracy of pointing at objects was problematic. The Spatial platform enables all the pointing rays with the username visible to everyone, while others do not. However, the pointing spot was too small to recognize and the angle from which a user was observing would cause confusion.

5.1.2 Input Device Control and Blindness

Only the Oculus Quest controllers were used and each platform was designed with this controller in mind. Controller use is still a problem as there are many active controls (five buttons, a joystick and finger proximity sensors). Typically, the index finger should remain on the trigger button to select and the second finger is for the grab button. However, our participants would often hold the devices slightly incorrectly. The platforms endowed the other buttons with different functions. Almost all the participants felt confusion in using the controller, even when platforms provided a tutorial. Participants could not always conduct the right action the first time, especially when teleporting and interacting with objects. Most of the time, users would ask their collaborator in the same room for assistance. However, even with the verbal assistance and not being able to see the real controller in their hands, participants had difficulty finding the correct buttons quickly.

5.1.3 Invisible Menu in the Environment

Most functions, except for object manipulation and locomotion, are hidden in the menu. However, the menu is not normally visible meaning that participants found it hard to orient to the necessary task actions. This was especially true when creating a new room (Mozilla Hubs was the exception) and participants spent a lot of time exploring for the right action and would need to ask for hints. For the invitation task, some of the participants liked to look around the room to seek cues. However, participants were not able to get more information from the environment to assist them as the functionality was in the menu rather than any object in the environment. This conflict between functions



FIGURE 1 | Two people interacting with an earth model in Mozilla Hubs.

attached to world objects and functions in the world is sometimes referred to the conflict between diegetic or non-diegetic design (Salomoni et al., 2016). In addition, the menu in VR is private and customized to each user, and invisible and inoperable to other users. This increased the difficulty to communicate about the menu, for instance, when choosing a new environment to travel to.

5.1.4 Spatial Navigation Problems

Navigation is still a problem that still needs to be addressed in the platforms in which users are free to move about. Participants all faced the problem of themselves colliding with objects in the environments. When locomoting around the environment, users experienced some unexpected situations such as moving through walls and objects. In particular, avatars could overlap or disappear when coming close to others.

There are no facilities to record location, which means users had difficulty in re-entering the same location if they accidentally left. Unlike desktop-based virtual environments and desktop multiplayer games, the social VR platforms based on HMDs do not usually provide a map of the environment.

5.1.5 Joint Manipulation Problems

All the platforms enable users to interact with and grab some objects. After grabbing, often more related actions are available and depend on the object. For example, in RecRoom, users can grab a bottle of water and pour it; in Spatial, users can zoom into a 3D model. Once secured, an article is labeled as temporarily in private hands, though there are no indications of this to other users. This caused issues such as the object appearing to not be interactive, but this was just because another user was already holding it.

5.1.6 Lack of Object Identity

There were no facilities for identifying and locking shared objects or tools especially when the platforms uses a ray casting to control objects remotely. Sometimes, more than one user was able to select the same interactive object. However, participants did not

know who owned this object and who was interacting with it. For instance, **Figure 1** is a third-person view of two participants interacting with an earth model. While it appears that the person on the right was in control, it was actually the person on the left.

5.1.7 Lack of Mutual Awareness

When participants manipulated objects the results of their action was visible to the others. However, two participants were not allowed to manipulate the same object simultaneously. If two participants were trying to select the same object, a conflict occurred, yet none of the platforms supports user awareness of not being able to manipulate simultaneously. In Spatial, when asking participants to control the same object at the same time, they all thought they controlled the object but what they saw is only based on their own manipulation. In other platforms, participants had an illusion that the object was in their control.

5.1.8 Precise Control and Distance Limitation

Platforms need to make a choice about supporting interaction only within arms reach or remotely. Participants found it easy to grab the objects near to them. For remote objects a ray is often used, but this also has a maximum distance, and objects beyond the end of the ray are not grabable. Given that the ray casting can involve long distances, participants found it hard to select objects precisely, especially for the objects close to each other. **Figure 2** illustrates how the ray casting operates a name tag. However, an object at the same distance would not have been interactive adding further confusion.

Most platforms provide a pen for users to draw in the air or on a surface. Drawing was considered fun, but the shapes were hard to distinguish, especially if multiple drawings overlapped. After drawing, a challenging problem was how to erase it all because it is hard to select items precisely.

5.1.9 Bubble Mechanism Hinders Close Interactions

To avoid the overlapping of avatars, some platforms use a bubble mechanism to make the avatar transparent when in close proximity, though this also hinders interpersonal interactions. In trials, most of the participants would have liked to pat someone's head, but the bubble hindered this by making the



FIGURE 2 | Ray casting in AltSpaceVR.



FIGURE 3 | Handshaking in RecRoom.

other avatar disappear. As can be seen in **Figure 3**, in RecRoom, when participants shook hands, the other person could go invisible and it became hard to locate the hand.

5.1.10 Room Coordination

In the six social VR platforms, the process of traveling to other environments or rooms is done via the menu. Two of them (VRChat and AltSpace) enable a user to create a portal for groups of people, but they adopt different techniques to enter the portal; VRChat allows users to move into the portal while AltSpace requires a click and a confirmation. However, almost none of the participants used this function as they were not aware of how to activate it.

5.2 Particular Problems

Some particular usability problems that arise are unique to one platform.

5.2.1 RecRoom

RecRoom is a game center for friends, offering various group games and interpersonal interaction. Participants can find their friends in a public place called the Rec Center.

Emotion expression: Participants were able to express their emotions by choosing an emoji with the controller from a shortcut menu. Rec Room chose to display the emotion on the avatars' face. However, users found it hard to get feedback on whether or not they were taking the correct action. In this case, participants tended to get confirmation from others or by looking in the mirror.

Inconsistency of manipulation: When manipulating objects, all participants reported difficulty in putting things down as the system automatically applied a locking mechanism in game playing. Participants would grab something and then were unable to drop it by releasing the grab button until they pressed the 'B' button on the controller. However, participants would not know this and, in other environments, this was not the case.

5.2.2 AltSpaceVR

AltSpaceVR provides open events for everyone and enables meetups. It depends heavily on the friend mechanism for coordination since adding friends is the precondition of all the tasks.

Avatar representation and identity: To assist users in recognizing the person they encounter, all the platforms except AltSpaceVR provide a name tag on the top of the avatar's head. In AltSpaceVR, users need to point at the avatar, then the name with applicable actions will emerge. This design made things clear when there are many people in a room but made it difficult for participants to find their friends. Some participants recognized the wrong person as their friend when they had a similar avatar appearance. Emotional expression is displayed as a floating 2D emoji on avatar's head. Participants confirmed their emotional expression through others' responses and when others expressed emotions.

Delays and missing notifications: There was no system notification to tell the participants the status of the room and about any newcomers. In AltSpaceVR it is important to add friends since participants can only invite their online friends to a room. However, it is hard for participants to notice when there is a friend request as it has a long latency and it notified by a small red marker on the menu.

5.2.3 VRChat

VRChat is an open world and accessible from many devices. Users can customize their avatar and environment through the developing kit that the platform provides. However, because of the flexibility of the world, some rooms are of low fidelity and problematic.

Less presence of the avatar: The teleporting in VRChat uses a third-person view, which means the users can see their avatar walking in front of them and then the view changes. This design is easy for others to follow but weakens the sense of presence.



FIGURE 4 | Notification icon in VRChat.

Almost all the participants reported a feeling of controlling a figure in a game instead of their own embodiment.

Slight notification: The notification system in VRChat is relatively light as participants found it hard to get attention and take the right actions. All the messages, including friend requests and invitations, were flashing icons on the left bottom of the view, prompting the user to open the menu and check, but most participants tried to click the icon instead of open the menu (see **Figure 4**). Even when opening the menu, the message was indicated as only a friend portrait at the top of the menu, which was still hard to notice.

Lack of feedback from the menu interface: According to group trials, it was hardest to create a room in VRChat since there were no cues to follow and no feedback to indicate the right action has been taken. In addition, the interface of room details confused users and they found it difficult to decide what to do next.

5.2.4 BigScreen

BigScreen is a platform providing a virtual space for people to watch movies and TV together. Similar to AltSpaceVR, there are no notification alerts regarding room status and the entry of a newcomer. Since watching movies requires focus, there should not be too much interference, so there are only a few social functions provided on the platform. Friend lists are unsupported.

Room coordination: BigScreen uses codes to identify the rooms, but the codes are random and consist of numbers and letters, which are hard to remember and share. Participants needed to check several times to ensure they have the right code. Participants could also choose to watch the same TV in a room together, but they could not be sure that they enter the correct room before the program starts.

5.2.5 Spatial

Spatial is a virtual conference tool for group meetings, which is ideal for design reviews and presentations.

Web media problem-limited view and control: Spatial offers a web version of limited functions to log in, upload files and share the desktop with others. However, the camera view is not under user control and depends on which of the immersed users is speaking. Users without a headset are not able to manipulate or locate anything in the environment. For the VR users, the web users only exist through voice or video in the environment and it is difficult to build a face-to-face setting for web users.

Invitation transferring: Spatial uses email to send invitations which is a good practice to bridge the virtual world and reality. However, participants always got lost after they opened the invitation link in the email. For the first-time user, the link leads them to register, then they need to login to the website, and then pair their device. The two problems that arose were an inability to distinguish the link attributes before sharing and a lack of understanding of the teams concept.

Presentation difficulty: In the virtual environment of Spatial, there is a fixed board for sharing documents, which is presented as a wall for the environment. However, when participants faced the board and talked about the documents there, because it was very large it proved very difficult for them to notice what others were doing and their status because of the limited field of view.

5.2.6 Mozilla Hubs

As a web-based social VR platform, mobility is the biggest advantage but participants found that there was also a high risk of encountering glitches. For example, if using a pen in the environment to draw something in the air, participants found it hard to write in an exact place in the air as it wrote on the ground instead.

No room backup: The room used to meet up in the Mozilla Hubs is disposable since users cannot enter the same room the next time. If the room owner encountered a glitch or had to quit for some reason, the room became inaccessible to everyone.

Less presence in the environment with the avatar: The environment and the robotic avatar decreased the feeling of 'being there' with real people. In this environment, participants were less enthusiastic about exploring and playing. In addition, there were random usernames and avatars in the platform, since participants were allowed to use it without registering. This made it more difficult to identify friends.

5.3 Interviews

The responses to post walkthrough interviews also provide useful feedback on aspects of the different platforms.

When asked "When you move your body, do you feel the avatar in the system act the same as you?", one participant indicated:

When I see my avater self in the mirrors, I know how the avatar acts

and another

I think so, because when I try to use my hand to interact I could see my virtual hand in the environment.

This is interesting because the self-avatar of the systems are quite different in how they appear to the user. Several of the platforms include mirrors, indeed in RecRoom the user starts in front of a mirror in their own private room, where they can configure their avatar. Otherwise users typically see only their hands.

From the walkthrough, participants generally didn't have problems interacting with objects, but in response to "Does the system follow real world conventions?" one participant noted the unreality of the environment:

It is quite interesting that I can grab some things in VR and throw it. But the things are like a paper with no weight.

When asked "Do you want/have to move physically when you want to move in the platform?" two of the participants emphasized that there is a confusion with the controllers because the controller effects travel:

When you ask me to move, my first act was to move my body walk a bit. However, things did not change. That

makes me a bit confused until you told me to try my controllers.

The tutorial taught me how to use it, but it is hard to remember. Even when I use it in real sessions, I can't see the buttons which need more exploration.

Prior work in assessment of VR has often relied on questionnaires about presence, but when asked directly "Do you feel you are in the environment? What's the level of your presence?" participants made interesting comments about the form of the environment, rather than their engagement with it (note that these comments were not made in a comparative context):

In VRChat, it is quite low, like 2 or 1, especially in the environment created by others.

The environment is surreal in Mozilla Hubs.

When asked "When you are talking, do you know others are listening?", participants conveyed mixed feelings:

It's hard for me to judge. I know others are there virtually and I only could assume that they are listening. Sometimes I know it because they give me some response like looking at me, hand gestures and verbal interaction.

I noticed that people are looking at me when I talking to them.

These suggest that non-verbal communication can be successful, but that the systems aren't always conveying it convincingly.

We noted in the walkthrough that participants had problems with joint control of objects. When asked: "When trying to manipulate the same objects, do you know who is interacting with it?" responses included:

Not at all. In my mind, all the objects are controlled by me.

I could estimate it through the change of objects. For example, if it moved without my control, then it is controlled by others.

These again suggest that visibility of actions needs attention.

6 DISCUSSION

6.1 Coordinating Friends and Locations

Perhaps the key set of findings from the walkthrough were around how to support collaboration within a social or work context not bounded by the VR experience itself. Here we expand on some of these points because they are very challenging for these platforms. We recommend that future inspections or evaluations consider these as primary issues.

6.1.1 Friend Mechanisms

VRChat, AltSpaceVR and RecRoom all provide a friend mechanism, allowing users to add friends, send friend invitations and exchange text messages. AltSpaceVR provides more details about the location of online friends and the ability to teleport to a friend's space. There are already four choices of finding friends on the different platforms:

- Using a code to search. This is a similar process to that used on online social platforms except that unlike on a mobile phone, say, you cannot show the other person the code as you enter it because the local view of your VR display is private.
- Using friend's username to search. This needs it to be spelled out and it is unlikely to be common with a name on another platform.
- Adding through face-to-face virtually, which requires people to meet in the same room on the VR platform and identify each other. This difficult for the first-time user: it needs a series of informal confirmation steps (e.g., recognizing the voice of the person) inside the VR or needs more coordination outside the platform.
- Importing friends from other platforms such as Oculus friends or Facebook friends. This is easy but is restricted to the existing group of friends that have already tried the platform and for whom you have IDs on the social platforms.

Clearly this is a challenge to these platforms. They are conceived and built as standalone applications, but they live within a broader ecosystem that involves many different social platforms.

6.1.2 Invitation Mechanisms

Invitations to meet involve both identifying the person and the location. In the platforms studied, there are three ways to navigate people into the same room: 1) share code or link; 2) send an invitation directly to online friends in the platform and 3) ask friends to use the 'go-to' function. The results show that sending invitations to online friends is the quickest of these three methods, though the ability to add friends to each other and an instant alert would be useful additional functions. Obviously this is complicated if you are not friends on the platform, but Spatial and Hubs allow links to be shared external to their platform.

6.2 Emotion and Gesture

Non-verbal communication such as emotion and gesture is significant in social VR (Moustafa and Steed, 2018; Zibrek and McDonnell, 2019). In our study, while the platforms supported simple gestures such as pointing, thumbs-up and hand waving, users had low expectations around interpersonal interactions. Rec Room was notable for having a specific interaction for high fives and handshakes. McVeigh-Schultz et al. (2019) noted those emotions and interactions as social catalysts and we believe that better, more nuanced expressions in VR could enhance engagement and improve the user experience. We note that there is a lot of

activity on capturing more facial expression and body movement (e.g., (Hickson et al., 2019)) but then the issue may then be asymmetry between users with and without these capabilities, or trust that the expressions are veridical rather than simulated.

6.3 Social Protection and Policy

We found, as would be expected, that users wanted sufficient personal space in the social VR environment. A personal space bubble was used on most platforms chose to retain a suitable distance between avatars, to prevent them being overlapped by other users, and to stop more than one avatar from occupying the same place. Violations of the bubbles are seen as being uncomfortable, and we should note that personal space violation is a key source of harassment in social VR (Blackwell et al., 2019a). This provides a challenge to governance and monitoring of these platforms (Blackwell et al., 2019b). The friend and invitation mechanisms mentioned above do give users ability to filter people and create private locations, but these are not very flexible mechanisms in many social situations, such as, say meeting new people in an organized large-scale meeting venue.

6.4 Other Design Choice Considerations

Some less critical design observations are:

1. The name tag and representative avatars are effective in identifying others. A photo-generated avatar was only considered by our users to be acceptable in business scenarios where users are familiar with the others; in a more open world, such avatars could create social awkwardness especially if users were free to choose or edit the pictures.
2. Others' actions, as well as interacting with the menu, should be visible and recognisable to everyone. For example, RecRoom and VRChat use a lightened screen to represent the personal menu to others. This allows the others to understand what the user was doing while protecting privacy.
3. Some of the tasks in some platforms are not observing basic usability features of the interaction cycle, such as giving instant feedback. For example interacting with the 3D objects in the environment, in some situations is giving feedback only through movement, but in a shared environment there are more channels for feedback, such as conveying success to the person who picks the object, and making sure that a second person grabbing the object gets appropriate feedback.
4. Eyes gaze of an avatar could create an illusion of focus. All the platforms assume that what the avatar is looking at is the same as the user, but there is less clarity around who is listening. One participant in AltSpaceVR, in answer to the question 'do you know others are listening to you?' replied to the affirmative, saying that he sees the eyes of others moving and always look at him. These were, in fact, simulated gaze directions so the user had been mis-led.

Additionally, after comparing usability evaluation, there are some routes for novel design as they are relatively under explored:

1. Verbal communication is essential, but the room-based model may not be ideal. We suggest exploring how groups might communicate between rooms: either when they are joining or moving, so as to avoid drop-outs and confusion.
2. Amalgamating all functions into menus does slow down users. While some platforms offer short-cuts to access emotion functions, this is not standard. We suggest that this is an area where some split of functionality would be appropriate, with separate menus for self-modifying actions and world-modifying actions.
3. While the metaphor of the social VR is that a person is in the environment, not a set of devices, participants do spend quite a lot of time talking about the controllers and how to effect actions. We might consider some guidelines about how to optionally represent the controllers to participants so that everyone can talk about them.
4. Teleporting is the most common approach to locomotion in virtual reality. However, for others in the space, it is disruptive to notice someone disappear suddenly. Currently platforms have a variety of ways of making teleportation more understandable to the user making the teleport, so some visualization for others might also be useful.
5. An effective notification system should be developed. This might be resolved at a operating system level, in a similar way that notifications on desktops, mobile devices or game services are handled.
6. Group navigation is a problem. It has been noted as a problem in CVEs for a while (Tromp et al., 1998). Portals are a solution in some situations for switching rooms, but difficult to use in general situations. For movement within visible range techniques are being developed (Weissker et al., 2019), but no platform had a good solution as yet.
7. Some of the platforms and world designs make good use of in-world (diegetic) interfaces. We expect that the design space here will be very interesting to explore.
8. Some platforms offer a shared board for people to write on, but it is hard to write on. Others provide free-form line sketching. Effective tools that draw on more experience in gesture-based and pen-based interfaces are urgently needed.

6.5 Reflection on Walkthrough Method

We used walkthrough as a method to get formative feedback about a set of social VR systems. This isn't the common use of this class of technique which is traditionally used to get expert or user feedback on an interface in development. However we found that it highlighted many usability issues, and that many of these were common across platforms or posed challenges to the way in which these social VR are constructed and posed.

One aspect of our work is that the usability issues are broad and not deep. There is a lot of detail to the walkthrough about specific problems, but we are not the developers of these platforms, so we were not interested in very specific usability feedback. We thus noted the problem and didn't attempt to do deep inspections of repeatability or severity of any specific issue. As noted, no issue was severe in that it crashed a system or stopped a task from being completed. The shallowness could be seen as a limitation, but as we have seen, a social VR platform incorporates a broad set of

interaction types and functionalities. Each area deserves detailed study on its own (e.g., emotion activation, see (Tanenbaum et al., 2020)), so we see our work, and this style of formative inspection, as being effective at raising new issues.

We did work with users that are new to the systems. They did have time to get familiar with tutorials. Naive users are a legitimate target evaluation population, as learnability is an important aspect of the usability of social VR platforms, especially as they are relatively novel and these platforms are expanding their user base quite rapidly as VR devices become cheaper and more available. However, usability walkthrough is typically done with expert users acting as users. More expert users might have learnt different strategies and be very fluent with a particular interface. This might be more amenable to a longitudinal approach (e.g., see (Moustafa and Steed, 2018)). Experts who were trained in usability inspection might also be able to give much more detailed feedback on specific usability issues. Thus one limitation of our results is that of the platforms tested, we can't know what proportion of usability issues has been discovered. However, our focus in this paper was more around uncovering common problems across platforms, and demonstrating that the guided group walkthrough method had merit.

Future use of our method would need to consider the specific objectives of the study it was being employed in. Within the iterative development of an application, one rule of thumb is that five experts should review the application (Nielsen and Landauer, 1993). While this might not reveal all the usability issues, the reviewing process is expensive and this level is sufficient to give good feedback to develop another iteration. If the objective of the study were to be more formally comparative of two or more platforms, we would suggest having a very detailed task breakdown, and being explicit about verifying the task breakdown is fair to each platform in use. While our tasks were very generic, if the goal of the study had a stronger focus on a specific task type (e.g., conduction of a brain-storming meeting), then the platforms might provide quite different tools to use.

Finally, we noted that having the experimenter included as an avatar changed the interaction with participants. Having the experimenter physically co-located with the participants does allow them to help with interaction issues such as controller use. However, if the participants are more experienced, this should not be so much of an issue. Having the experimenter in the social VR experience does enable remote study and thus might be important for effective use of the method. One technical aspect that we did not explore, but would be very useful, is remote streaming of screen views so that a remote experimenter can observe what the participant(s) are seeing. While the experimenter might see the participant's avatar and many of its action, interactions with system user interface components such as menus are not typically visible to other users. This then suggests that a mixed model of remote and co-located studies would be advantageous.

7 CONCLUSION

The overall aim of this study was to compare social VR experiences on different platforms and identify usability problems. These social VR platforms are relatively new, but they have to support some key features around group coordination. Thus we expected that there would be common problems across the platforms due to the nature of the medium, and also platform-specific issues. Exploring these common and specific problems would highlight interesting design issues for the platforms and challenges for researchers in the area.

Thus we prepared a representative task analysis of social groups formation and activity within social VR. For each of six social VR platforms, we analyzed the operation of the individual tasks and sub-tasks and then mapped them to one of several interaction cycles (2D, 3D, etc.) that had a different set of usability heuristic question. We then ran a set of guided group walkthroughs with participants, where participant activity was recorded, and an experimenter asked the participants the appropriate questions for the task that they were undertaking.

The walkthrough results gathered some common usability problems in social VR platforms such as communication coordination, spatial navigation and joint manipulation. We also identified some issues specific to particular platforms. We identified the key problems of coordinating friends and locations, and conveying gesture and emotion. Overall, we believe that this complements recent work that have inventoried some of the design choices of recent social VR platforms.

One limitation of our study is that because we worked with commercial social VR experiences, we were not able to record sensor data from the devices nor log data from the social VR platform itself to provide diagnostic feedback that could be supplied to the designers or engineers to rectify the problems. It was not the explicit intention of our platform to find specific issues that would need such data, but future use of a walkthrough method might consider recording additional data in order to reconstruct incidents. We further note that is recordings can be made in a documented format, they can be a useful resource for other researchers (e.g. see (Murgia et al., 2008)).

Nevertheless, our guided group walkthrough method proved effective in this context. The experimenter interacted with representative users through the social VR platform itself. Future work might consider whether having the experimenter within the social VR is desirable or whether they should sit outside the system. We note that in our situation having the experimenter within the social VR as an avatar solved logistical problems of running the study remotely, as this would otherwise have needed a real-time observation in another manner. We further note that we anyway needed a background text channel to help participants rendezvous with the experimenter in the social VR spaces. Overall we believe that the strategy of guided group walkthrough method is a useful method a way to uncover the problems users are encountering in social VR platforms.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by UCL Research Ethics Committee. The participants provided their written informed consent to participate in this study.

REFERENCES

- Argelaguet, F., Hoyet, L., Trico, M., and Lecuyer, A. (2016). "The Role of Interaction in Virtual Embodiment: Effects of the Virtual Hand representation," in IEEE Virtual Reality (VR). 3–10, Greenville, SC, March 19–23, 2016 (IEEE), 2375–5334.
- Bias, R. G. (1994). "The Pluralistic Usability Walkthrough: Coordinated Empathies," in *Usability Inspection Methods*. New York, NY: John Wiley and Sons Inc, 63–76.
- Bigscreen (2021). Bigscreen Inc. Available at: <https://www.bigscreenvr.com/> (Accessed April 14, 2021).
- Biocca, F., Harms, C., and Burgoon, J. K. (2003). Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence Teleoperators Virtual Environ.* 12, 456–480. doi:10.1162/105474603322761270
- Blackwell, L., Ellison, N., Elliott-Deflo, N., and Schwartz, R. (2019a). Harassment in Social Virtual Reality: Challenges for Platform Governance. *Proc. ACM Hum. Comput. Interact.* 3, 1–25. doi:10.1145/3359202
- Blackwell, L., Ellison, N., Elliott-Deflo, N., and Schwartz, R. (2019b). "Harassment in Social Vr: Implications for Design," in IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, March 23–27, 2019 (IEEE), 854–855.
- Blanchard, C., Burgess, S., Harvill, Y., Lanier, J., Lasko, A., Oberman, M., et al. (1990). "Reality Built for Two: a Virtual Reality Tool," in Proceedings of the 1990 Symposium on Interactive 3D Graphics, Snowbird, UT, March 1990. (New York, NY, USA: Association for Computing Machinery), 35–36. doi:10.1145/91394.91409
- Bowman, D. A., Gabbard, J. L., and Hix, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence Teleoperators Virtual Environ.* 11, 404–424. doi:10.1162/105474602760204309
- Churchill, E. F., and Snowdon, D. (1998). Environments: An Introductory Review of Issues and Systems. *Virtual Reality* 3, 3–15. doi:10.1007/bf01409793
- Damer, B., Judson, J., Dove, J., Illustrator-DiPaola, S., Illustrator-Ebtekar, A., Illustrator-McGehee, S., et al. (1997). *Avatars!: Exploring and Building Virtual Worlds on the Internet*. Berkeley, CA: Peachpit Press.
- Fabri, M., Moore, D. J., and Hobbs, D. J. (1999). "The Emotional Avatar: Non-verbal Communication between Inhabitants of Collaborative Virtual Environments," in *Gesture-Based Communication in Human-Computer Interaction*. Editors A. Braffort, R. Gherbi, S. Gibet, D. Teil, and J. Richardson (Springer Berlin Heidelberg: Lecture Notes in Computer Science), 269–273. doi:10.1007/3-540-46616-9_24
- Freeman, G., Zamanifard, S., Maloney, D., and Adkins, A. (2020). My Body, My Avatar: How People Perceive Their Avatars in Social Virtual reality. Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. *CHI EA 20*, 1–8. doi:10.1145/3334480.3382923
- Gabbard, J. L. (1997). A Taxonomy of Usability Characteristics in Virtual Environments. Master's thesis. Va. Tech.
- Geszten, D., Komlódi, A., Hercegf, K., Hámornik, B., Young, A., Koles, M., et al. (2018). A Content-Analysis Approach for Exploring Usability Problems in a Collaborative Virtual Environment. *Acta Polytechnica Hungarica* 15, 67–88. doi:10.12700/APH.15.5.2018.5
- Gonzalez-Franco, M., and Peck, T. C. (2018). Avatar Embodiment. Towards a Standardized Questionnaire. *Front. Robotics AI* 5, 74. doi:10.3389/frobt.2018.00074
- Hai, W., Jain, N., Wydra, A., Thalmann, N. M., and Thalmann, D. (2018). "Increasing the Feeling of Social Presence by Incorporating Realistic Interactions in Multi-Party VrCASA," in Proceedings of the 31st International Conference on Computer Animation and Social Agents, Beijing, China, May 2018. (New York, NY, USA: Association for Computing Machinery), 7–10.
- Hartson, H. R., Andre, T. S., and Williges, R. C. (2001). Criteria for Evaluating Usability Evaluation Methods. *Int. J. Human Comput. Interact.* 13, 373–410. doi:10.1207/s15327590ijhc1304_03
- Hickson, S., Kwatra, V., Dufour, N., Sud, A., and Essa, I. (2019). "Eyemotion: Classifying Facial Expressions in VR Using Eye-Tracking Cameras," in Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019, Osaka, Japan, March 7–11, 2019 (IEEE), 1626–1635.
- Hindmarsh, J., Fraser, M., Benford, S., Greenhalgh, C., and Heath, C. (2000). Object-Focused Interaction in Collaborative Virtual Environments. *ACM Trans. Comput. Human Interact.* 7, 477–509. doi:10.1145/365058.365088
- Hollingsed, T., and Novick, D. G. (2007). "Usability Inspection Methods after 15 Years of Research and Practice SIGDOC '07," in Proceedings of the 25th Annual ACM International Conference on Design of Communication, El Paso, TX, October 2007 (New York, NY, USA: Association for Computing Machinery), 249–255.
- Jadhav, D., Bhutkar, G., and Mehta, V. (2013). "Usability Evaluation of Messenger Applications for Android Phones Using Cognitive Walkthrough APCHI '13," in Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction, Bangalore, India, September 2013. (New York, NY, USA: Association for Computing Machinery), 9–18. doi:10.1145/2525194.2525202
- Jonas, M., Said, S., Yu, D., Aiello, C., Furlo, N., and Zytka, D. (2019). "Towards a Taxonomy of Social VR Application Design," in Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts CHI PLAY '19 Extended Abstracts, Barcelona, Spain, October 2019. (New York, NY, USA: Association for Computing Machinery), 437–444.
- Kalawsky, R. S. (1999). VRUSE—a Computerised Diagnostic Tool: for Usability Evaluation of Virtual/synthetic Environment Systems. *Appl. Ergon.* 30, 11–25. doi:10.1016/s0003-6870(98)00047-7
- Kilteni, K., Groten, R., and Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence Teleoperators Virtual Environ.* 21, 373–387. doi:10.1162/pres_a_00124
- Kolesnichenko, A., McVeigh-Schultz, J., and Isbister, K. (2019). "Understanding Emerging Design Practices for Avatar Systems in the Commercial Social Vr Ecology," in Proceedings of the 2019 on Designing Interactive Systems Conference DIS 19, San Diego, CA, June 2019. (New York, NY, USA: Association for Computing Machinery), 241–252.
- Latoschik, M. E., Roth, D., Achenbach, J., Waltemate, T., and Botsch, M. (2017). "The Effect of Avatar Realism in Immersive Social Virtual Realities VRST," in Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology 17, Gothenburg, Sweden, November 2017. (New York, NY, USA: Association for Computing Machinery), 1–10.

AUTHOR CONTRIBUTIONS

QL Developed the materials for the study, ran the walkthrough sessions, wrote up the method and results. AS Proposed the project, supervised the work and co-wrote the paper.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frvir.2021.668181/full#supplementary-material>

- LaViola, J., Jr, Kruijff, E., McMahan, R., Bowman, D., and Poupyrev, I. (2017). *3D User Interfaces: Theory and Practice*. Boston: Addison-Wesley.
- Lewis, C., Polson, P. G., Wharton, C., and Riemann, J. (1990). "Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-And-Use Interfaces," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '90, Seattle, WA, March 1990. (New York, NY, USA: Association for Computing Machinery), 235–242. doi:10.1145/97243.97279
- Mahatody, T., Sagar, M., and Kolski, C. (2010). State of the Art on the Cognitive Walkthrough Method, its Variants and Evolutions. *Int. J. Human Comput. Interact.* 26, 741–785. doi:10.1080/10447311003781409
- Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. (2015). Changing Bodies Changes Minds: Owning Another Body Affects Social Cognition. *Trends Cogn. Sci.* 19, 6–12. doi:10.1016/j.tics.2014.11.001
- Mantovani, G. (1995). Virtual Reality as a Communication Environment: Consensual Hallucination, Fiction, and Possible Selves. *Hum. Relations* 48, 669–683. doi:10.1177/001872679504800604
- Marsh, T. (1999). "Evaluation of Virtual Reality Systems for Usability CHI EA '99," in CHI '99 Extended Abstracts on Human Factors in Computing Systems, Pittsburgh, PA, May 1999. (New York, NY, USA: Association for Computing Machinery), 61–62.
- McVeigh-Schultz, J., Kolesnichenko, A., and Isbister, K. (2019). "Shaping Prosocial Interaction in Vr: An Emerging Design Framework," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 19, Glasgow, United Kingdom, May 2019. (New York, NY, USA: Association for Computing Machinery), 1–12.
- Microsoft (2020). AltspaceVR. Available at: <https://altvr.com/> (Accessed April 14, 2021).
- Moustafa, F., and Steed, A. (2018). "A Longitudinal Study of Small Group Interaction in Social Virtual Reality," in Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology VRST '18, Tokyo, Japan, November 2018. (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3281505.3281527
- Mozilla Corporation (2021). Mozilla Hubs. Available at: <https://hubs.mozilla.com> (Accessed April 14, 2021).
- Murgia, A., Wolff, R., Steptoe, W., Sharkey, P., Roberts, D., Guimaraes, E., et al. (2008). "A Tool for Replay and Analysis of Gaze-Enhanced Multiparty Sessions Captured in Immersive Collaborative Environments," in 12th IEEE/ACM International Symposium on Distributed Simulation and Real-Time Applications, Vancouver, BC, October 27–29, 2008, 252–258. doi:10.1109/DS-RT.2008.25
- Nielsen, J., and Landauer, T. K. (1993). "A Mathematical Model of the Finding of Usability Problems CHI," in Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, CHI'93, Amsterdam, Netherlands, May 1993. (New York, NY, USA: Association for Computing Machinery), 206–213. doi:10.1145/169059.169166
- Nielsen, J. (1994). "Usability Inspection Methods," in Conference Companion on Human Factors in Computing Systems, CHI'94, Boston, MA, April 24–28, 1994. (New York, NY, USA: Association for Computing Machinery), 413–414.
- Pan, Y., and Steed, A. (2017). The Impact of Self-Avatars on Trust and Collaboration in Shared Virtual Environments. *PLOS ONE* 12, e0189078. doi:10.1371/journal.pone.0189078
- Pinelle, D., and Gutwin, C. (2002). "Groupware Walkthrough: Adding Context to Groupware Usability Evaluation," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '02, Minneapolis, MN, April 2002. (New York, NY, USA: Association for Computing Machinery), 455–462.
- Rec. Room (2021). Rec. Room. Available at: <https://recroom.com> (Accessed April 14, 2021).
- Salomoni, P., Prandi, C., Rocchetti, M., Casanova, L., and Marchetti, L. (2016). "Assessing the Efficacy of a Diegetic Game Interface with Oculus Rift," in 13th IEEE Annual Consumer Communications Networking Conference (CCNC), 387–392, Las Vegas, NV, January 9–12, 2016 (IEEE), 2331–9860.
- Schroeder, R. (2010). *Being There Together: Social Interaction in Shared Virtual Environments*. Oxford, United Kingdom: Oxford University Press.
- Schulz, R. (2020). Comprehensive List of Social VR Platforms and Virtual Worlds. Available at: <https://ryanschultz.com/list-of-social-vr-virtual-worlds/> (Accessed April 14, 2021).
- Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence Teleoperators Virtual Environ.* 1, 120–126. doi:10.1162/pres.1992.1.1.120
- Slater, M. (2009). Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Trans. R. Soc. B: Biol. Sci.* 364, 3549–3557. doi:10.1098/rstb.2009.0138
- Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. V. (2008). Towards a Digital Body: the Virtual Arm Illusion. *Front. Hum. Neurosci.* 2, 6. doi:10.3389/neuro.09.006.2008
- Spatial Systems (2021). Spatial Systems Inc. Available at: <https://spatial.io/> (Accessed April 14, 2021).
- Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., and Graeber, D. A. (2003). Usability Engineering of Virtual Environments (VEs): Identifying Multiple Criteria that Drive Effective VE System Design. *Int. J. Human Comput. Stud.* 58, 447–481. doi:10.1016/s1071-5819(03)00015-6
- Stone, V. E. (1993). Social Interaction and Social Development in Virtual Environments. *Presence Teleoperators Virtual Environ.* 2, 153–161. doi:10.1162/pres.1993.2.2.153
- Sutcliffe, A. G., and Kaur, K. D. (2000). Evaluating the Usability of Virtual Reality User Interfaces. *Behav. Inf. Technology* 19, 415–426. doi:10.1080/014492900750052679
- Tanenbaum, T. J., Hartoonian, N., and Bryan, J. (2020). "How Do I Make This Thing Smile?": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems CHI '20, Honolulu, HI, April 2020. (New York, NY, USA: Association for Computing Machinery), 1–13.
- Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., and Frécon, E. (1998). Small Group Behavior Experiments in the Coven Project. *IEEE Computer Graphics Appl.* 18, 53–63. doi:10.1109/38.734980
- Tromp, J. G., Steed, A., and Wilson, J. R. (2003). Systematic Usability Evaluation and Design Issues for Collaborative Virtual Environments. *Presence Teleoperators Virtual Environ.* 12, 241–267. doi:10.1162/105474603765879512
- Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., et al. (1999). "Walking >Walking-In-Place >Flying, in Virtual Environments," in Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques '99, Los Angeles, CA, August 1999. (New York, NY, USA: SIGGRAPH: ACM Press/Addison-Wesley Publishing Co.), 359–364.
- VRChat Inc (2021). VRChat. Available at: <https://www.vrchat.com/> (Accessed April 14, 2021).
- Weissker, T., Kulik, A., and Froehlich, B. (2019). "Multi-Ray Jumping: Comprehensible Group Navigation for Collocated Users in Immersive Virtual Reality," in IEEE Conference On Virtual Reality And 3D User Interfaces (VR) 136–144, Osaka, Japan, March 23–27, 2019 (IEEE), 2642–5254.
- Witmer, B. G., and Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence Teleoperators Virtual Environ.* 7, 225–240. doi:10.1162/105474698565686
- Yee, N., Bailenson, J. N., Urbanek, M., Chang, F., and Merget, D. (2007). The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. *CyberPsychology Behav.* 10, 115–121. doi:10.1089/cpb.2006.9984
- Zibrek, K., and McDonnell, R. (2019). "Social Presence and Place Illusion Are Affected by Photorealism in Embodied Vr," in Motion, Interaction and Games MIG '19, Newcastle upon Tyne, United Kingdom, October 2019. (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3359566.3360064

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Liu and Steed. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.