



User Experience in Collaborative Extended Reality: Overview Study

Huyen Nguyen^{1,2(✉)} and Tomasz Bednarz^{2,3}

- ¹ Université Paris-Saclay, CNRS, LIMSI, VENISE team, 91400 Orsay, France
huyen.nguyen@limsi.fr
- ² EPICentre, University of New South Wales, Sydney, NSW 2021, Australia
t.bednarz@unsw.edu.au
- ³ CSIRO Data61, Australian Technology Park, Eveleigh, NSW 2015, Australia

Abstract. Recent trends in Extended Reality technologies, including Virtual Reality and Mixed Reality, indicate that the future infrastructure will be distributed and collaborative, where end-users as well as experts meet, communicate, learn, interact with each other, and coordinate their activities using a globally shared network and mediated environments. The integration of new display devices has largely changed how users interact with the system and how those activities, in turn, change their perception and experience. Although a considerable amount of research has already been done in the fields of computer-supported collaborative work, human-computer interaction, extended reality, cognitive psychology, perception, and social sciences, there is still no in-depth review to determine the current state of research on multiple-user-experience-centred design at the intersection of these domains. This paper aims to present an overview of research work on coexperience and analyses important aspects of human factors to be considered to enhance collaboration and user interaction in collaborative extended reality platforms, including: (i) presence-related factors, (ii) group dynamics and collaboration patterns, (iii) avatars and embodied agents, (iv) nonverbal communication, (v) group size, and (vi) awareness of physical and virtual world. Finally, this paper identifies research gaps and suggests key directions for future research considerations in this multidisciplinary research domain.

Keywords: User experience · Coexperience · Virtual Reality · Augmented Reality · Mixed Reality · eXtended Reality · Collaboration · Interaction design

1 Introduction

Advances in eXtended Reality (XR) technologies, which is a term referring to Virtual Reality (VR) and Mixed Reality (MR) (Augmented Reality (AR) and Augmented Virtuality) within the Milgram's reality-virtuality continuum [70],

have dramatically changed the human-machine interactions mediated by computers and wearable devices. These technologies give Computer-Supported Collaborative Work (CSCW) the possibility to integrate various elements into a shared world, including heterogeneous user interfaces, data structures, information models, and graphical representations of users themselves. For instance, several overviews on collaboration in MR can be found in [12,61]. The integration of multiple devices and interaction modalities has largely changed how users interact with data and with other users. Using XR systems, the human experience, behaviour, and cognitive performance are an immensely important topic across other domains of Human-Computer Interaction (HCI) design, cognitive psychology, perception and others, particularly to make the most effective use of such systems. Therefore, appropriate standard multi-sensory stimuli interaction design and exchange mechanisms are needed to facilitate the full potential of the interaction between humans, data and artefacts, XR platforms, and the physical world. Furthermore, with recent breakthroughs in AR and great effort in bringing this technology to larger public, there is a need to merge the physical world with the virtual world while preserving the presence, copresence and the sense of collaboration between users using different modalities. In our opinion, it will not be long before the XR will become a platform of choice not only for complex task solving such as scientific data analysis, modelling, simulation, but also for public use such as education, social networking, video games, online custom services, and entertainment. In education, for example, Johnson-Glenberg et al. have argued the importance of collaborative MR environment to learning on the motivation, social cohesion, cognitive development, and cognitive elaboration perspectives [57].

Scientists and developers in HCI, design, and human behaviour research have been working on different factors of User Experience (UX) and how to quantify it, e.g., [112]. As defined in [49], user experience is indeed a complex and dynamic concept which involves a wide range of perspectives from user's internal states (e.g., motivation, emotions, expectations), to system settings (e.g., complexity, usability, functionality, purpose) and interaction context (e.g., environment, organisational/social setting, meaningfulness of activities). On the other hand, Battarbee and Koskinen proposed a taxonomy of existing approaches and considered UX under the three main perspectives: *measuring*, *empathic*, and *pragmatist* approach [11]. They also introduced the *coexperience* concept which explores "how the meanings of individual experiences emerge and change as they become part of social interaction". Sharing the same interest in coexperience, we approach the UX concept from another angle in the collaborative XR context. We focus on the general characteristics and features of multiple-user-experience-centred approach in collaborative XR systems in order to interpret and apply them into the design process. More specifically, we are interested in different factors relating to coexperience and co-interaction, including, but not limited to, presence, copresence, social presence, social effects, group collaboration patterns, embodiment, and so forth. Many existing works on these factors have been done within a short period of trials and experiments thanks to the capacity

and flexibility to replicate and control environments that easily fit experimental designs. Immersive projection technology and head-mounted displays (HMDs) are often the most used systems in UX studies and their performance is generally evaluated against existing desktop systems [90]. Therefore, with the recent advances in XR technology, especially in AR, we believe that different aspects of UX in collaborative XR platforms needs to be reviewed and reassessed. In this paper, we provide an overview of research conducted on UX in collaborative XR systems, especially in shared virtual or augmented environments. Our objective is to provide an introduction to researchers to this multidisciplinary domain and present opportunities for future research directions.

2 Context and Scope

In order to situate our study on UX in collaborative XR systems in the current related work, we have conducted a preliminary analysis of the research publications in the multidisciplinary domain of XR, CSCW and UX. Specifically, we used the Citation Network dataset¹ version 12 published on April 9, 2020 for the analysis. This dataset is constructed from DBLP², ACM³, MAG⁴, amongst other sources to provide a comprehensive list of research papers in major computer science journals and proceedings. This latest version contains almost five million publications and more than 45 million citation relationships. The sheer amount of data collected in this full dataset begs for some preprocessing steps before we could visualise it in the form of graph. After the data retrieval step, a parser has been used to transform its JSON original format into CSV format with only few fields of interest from each paper, including identification number, title, list of authors, year, and list of field of study (FOS). This dataset was then “standardised” by reformatting each word, removing punctuations and escape sequences, and converting all the characters into lower case. We extracted a smaller subset of this data by using several FOS that reflect the joint domains of interest of XR, CSCW and UX for this study. For instance, ‘collaborative virtual environment’, ‘augmented reality’, ‘virtual reality’, ‘augmented virtuality’, ‘immersive technology’, ‘user experience design’, ‘user experience evaluation’ and other relevant FOS were selected from the full list of available FOS of the dataset. Any paper that contains at least one of these FOS is picked from the original dataset. As a result, the subset has been reduced to 50,662 papers which are associated with 19,792 FOS. Since each paper is linked to a set of FOS, we proposed to analyse the dataset using seven FOS categories that represent main aspects to be considered in this study. Each category contains many FOS so only few examples are listed as follows:

- Extended reality: ‘virtual reality’, ‘augmented reality’, ‘mixed reality’, ‘3d interaction’, ‘immersion (virtual reality)’, ‘virtual learning environment’

¹ <https://www.aminer.org/citation>.

² <https://dblp.uni-trier.de/>.

³ <https://dl.acm.org/>.

⁴ <https://aka.ms/msracad>.

- User experience: ‘user experience design’, ‘user modeling’, ‘user-centered design’, ‘quality of experience’, ‘human factors and ergonomics’
- Communication: ‘gesture’, ‘gesture recognition’, ‘eye tracking’, ‘gaze’, ‘natural interaction’, ‘facial expression’, ‘communication skills’
- Collaboration: ‘collaborative virtual environment’, ‘computer-supported collaborative work’, ‘collaborative learning’, ‘virtual classroom’
- Emotion: ‘uncanny valley’, ‘anticipation’, ‘enthusiasm’, ‘surprise’, ‘happiness’, ‘emotional expression’, ‘confusion’, ‘pleasure’, ‘curiosity’
- Psychology: ‘social psychology’, ‘cognitive psychology’, ‘sociology’, ‘cognitive science’, ‘mental health’, ‘exposure therapy’, ‘cognitive walkthrough’
- Others: ‘situation awareness’, ‘spatial contextual awareness’, ‘perception’, ‘personality’, ‘sense of presence’, ‘sensation’, ‘personal space’

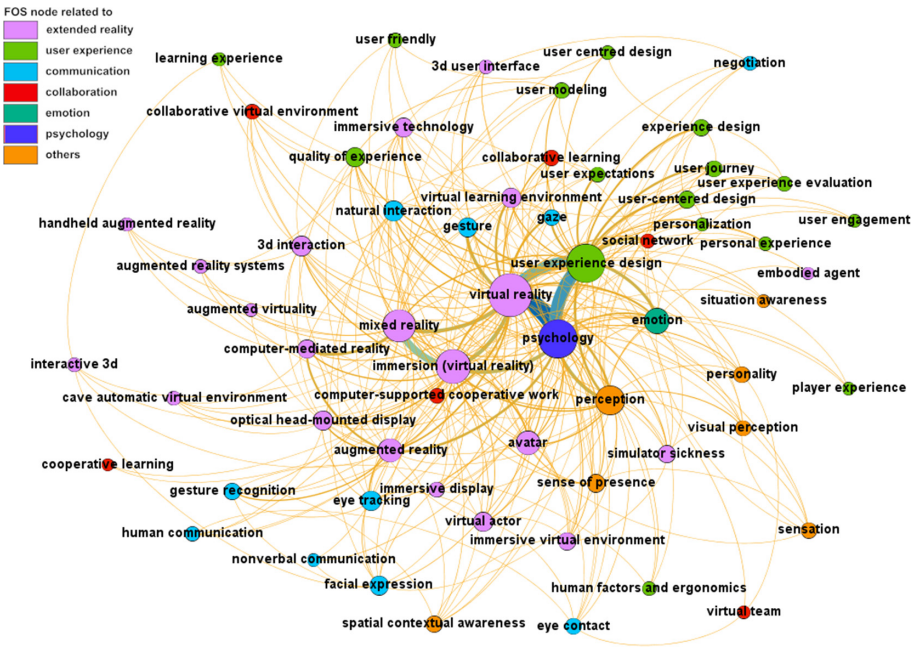


Fig. 1. An undirected graph built from a subset of the citation network dataset which focuses on the multidisciplinary domain of XR, CSCW, and UX. Except the two ‘psychology’ and ‘emotion’ nodes, each node represents a FOS of interest. The FOS are categorised in seven groups and represented in distinct colours. The size of each node represents the number of occurrences of its FOS. Each edge connects two FOS nodes when a publication is associated with these two FOS. The number of co-occurrences of two linked FOS is used to weight the width and adapt colour (from orange to green and blue) of each edge. (Color figure online)

Using Gephi⁵ software and Force Atlas graph layout algorithm [10], an undirected graph was built representing the number of papers that connect these FOS categories (see Fig. 1). To simplify the graph due to the limited visualisation space, the FOS of the two Psychology and Emotion categories have been generalised to create two ‘psychology’ and ‘emotion’ nodes, respectively. Based on this full graph and its subgraphs extracted using Gephi, we have come to some general observations regarding the relationship between XR, CSCW, and UX as follows:

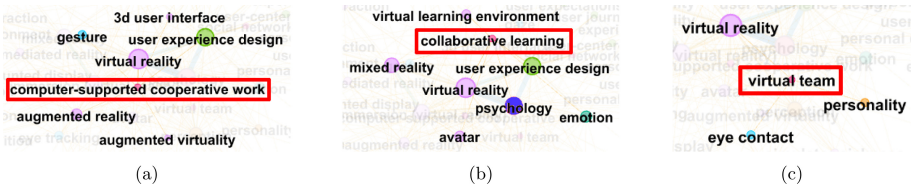


Fig. 2. The connection of the three FOS in Collaboration category: (a) ‘computer-supported cooperative work’, (b) ‘collaborative learning’, and (c) ‘virtual team’ to other FOS. All the FOS present in each figure have a direct connection with the FOS of interest (in red box). The layout of the graph has been slightly adjusted for the readability purpose, which is similarly applied in Fig. 3.

- There are strong connections in research between ‘virtual reality’, ‘user experience design’, and ‘psychology’ as demonstrated in Fig. 1. However, collaboration aspect is only weakly presented in the existing literature in general.
- The ‘computer-supported collaborative work’ and ‘collaborative virtual environment’ are only linked to ‘user experience design’, ‘quality of experience’, ‘gesture’ and the FOS of XR field in general (Fig. 2a). However, the ‘collaborative learning’ shows a more divers correlation with ‘avatar’, ‘emotion’, ‘psychology’, ‘user experience design’, and ‘mixed reality’, ‘virtual reality’ and ‘virtual learning environment’ (Fig. 2b). In addition, the ‘virtual team’ is connected solely to ‘virtual reality’, ‘eye contact’, and ‘personality’ (Fig. 2c). These results demonstrate a growing interest in the application of collaborative XR environment in education and training and its effectiveness on learners in terms of psychology and behavioural health care. These observations also lead us to believe that there is a need to study more closely the effect of collaborative XR environment on user experience and perception and vice versa, how to improve users’ experience when they work with others in a more general immersive context.
- Avatars have been largely studied relating to many domains of XR, UX, psychology, human emotion, communication, amongst others (Fig. 3a). This confirms the important role of the use of avatar and embodied agents in the context of our study.

⁵ <https://gephi.org/>.

- The ‘nonverbal communication’ and ‘natural interaction’ are closely linked with XR and UX fields, including ‘virtual reality’, ‘augmented reality’, ‘user experience design’, ‘user expectations’, and others (Fig. 3b). They are also characterised by ‘gesture’, ‘facial expression’, ‘gaze’, and ‘eye tracking’. Surprisingly, from the graph, there is little connection found between the FOS of the Communication category with those of the Collaboration group.
- Similarly, we cannot find the strong connection between the ‘sense of presence’ and the Collaboration category (Fig. 3c). Besides being linked to XR domain, the presence aspect is often studied in relation to psychology and human emotion, and interestingly, to some subtopics of communication in ‘facial expression’ and ‘negotiation’ as well.
- In the same situation, the FOS of the Collaboration category are not present in the list of FOS related to ‘spatial contextual awareness’ (Fig. 3d). In addition to FOS of XR field, the FOS on awareness also connects directly to psychology and ‘user experience design’. It is interesting to point out that this FOS also has a connection with ‘gesture recognition’.

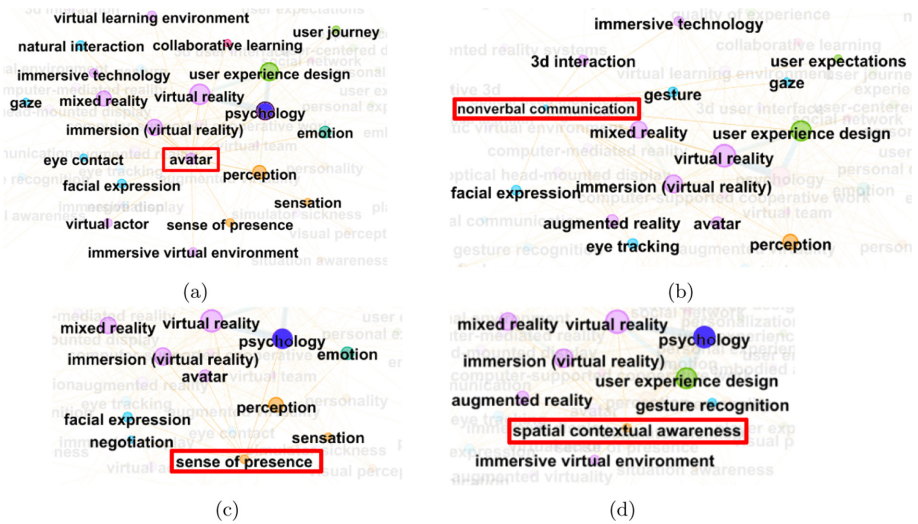


Fig. 3. The connection of (a) ‘avatar’, (b) ‘nonverbal communication’, (c) ‘sense of presence’, and (d) ‘spatial contextual awareness’ to other FOS.

There are several limits to the above observations that we take into account. Firstly, the graph was built from the Citation Network dataset which contains FOS generated using an automated keyword extraction algorithm with hierarchical topic modelling and natural language processing [93]. It is likely that errors can be accumulated starting from the extraction algorithm to the graph generation, which might limit the accuracy of the result. Secondly, these observations are not intended to be exhaustive. Several topics of the UX such as ‘user

friendly’, ‘user expectations’, ‘user journey’, ‘experience design’ have not been analysed. We consider these observations only as the first guidelines to help us identify important aspects of the related work in this multidisciplinary domain.

In this study, we will focus on the five factors that stand out from the preliminary analysis, including presence-related factors, group dynamics and collaboration patterns with virtual teams, avatars and embodied agents, nonverbal communication, and awareness of physical and virtual world in spatial contextual awareness. We add the group size as another factor to be considered as well. In Sect. 3 we begin by presenting these factors relating to UX in collaborative XR environments, focusing on shared virtual worlds.

3 User Experience in Collaborative Extended Reality Platforms

It has been confirmed that XR technologies applied to collaborative user interfaces help to enhance communication and support seamless functional and cognitive workflows between users [17]. However, the influence of technology on UX and how they behave within XR systems still leave a lot to be explored because most of existing studies have been conducted only in collaborative VR systems. Based on the preliminary study in Sect. 2, we present a deep analysis of the six following aspects considered for coexperience-centred collaborative immersive design. These aspects are seemingly independent and isolated from one another. In contrast, they are the key facets that construct the coexperience concept in collaborative XR systems. Research opportunities arising from this review will be summarised in Sect. 4.

3.1 Presence, Copresence and Social Presence

Presence, also known as physical presence or telepresence, has been one of the most studied research topics in VR and psychology. It is an ultimate goal of all the VR systems to initiate and maintain an individual’s sense of “being there” or “being in a virtual place” to make them believe or feel that they exist within the virtual world [50, 51, 96]. It involves the subconscious and conscious processes of being in and interacting with the virtual world: from automatic reactions to spatial and visual cues and triggers at the low intuitive level of perception, to more complex mental models of virtual spaces to create the illusion of place [19]. IJsselsteijn et al. in [56] emphasised on the importance of high quality mediated environments in terms of fidelity of sensory information, match between sensors and display, contents, and user characteristics to create the sense of “being there”. Slater & Wilbur argued that what to be expected when users feel a sense of presence within a virtual world is that their behaviours are consistent with those that would have occurred in the real world in similar situations [96]. Schuemie et al. have produced an overview on how to measure presence using subjective questionnaires and objective measures on behavioural and physiological responses [92]. In this study, we put more focus on social

presence and copresence, and how presence influences these two aspects when users interact with each other.

Social presence, as defined by Heeter, is the degree to which users believe that they are with other human beings and interact with them [50]. This definition has been expanded by Biocca & Harms in which social presence is “the moment-to-moment awareness of co-presence of a mediated body and the sense of accessibility of the other being’s psychological, emotional, and intentional states” [18]. It is argued that the social presence reflects the actual presence of others, the implied presence of them, or the imagined presence conveyed through sensory information transmission in mediated environments [1]. Therefore, the capacity of XR systems or technology is important in social presence to provide high fidelity of communication cues including proximity and orientation of others, physical appearance, facial expressions, gaze and mutual gaze, postures and gestures, verbal signals. Unfortunately, new technologies in XR have not been able to fully satisfy these requirements yet.

Another similar term which is often mentioned in social psychological research for virtual environments is copresence. Copresence, as summarised by Schroeder in [89], is the sense of “being there together” and acting with other users at the same time. Copresence puts more focus on the individual’s feelings of being part of a group and being capable of perceiving others [95]. In other words, a mutual awareness between individuals on the existence of each other is emphasised in copresence measures [24]. Schroeder considered copresence within collaborative virtual environments based on what activities users do together [90]. Compared to social presence, which relies on the quality of the mediated environment and users’ perception of it, copresence reflects more psychological interactions between them [89]. Also, studies on copresence need to consider users’ experience when they do things together, and not only when they are just immersed together in the virtual world. Schroeder in [89] separated three types of study on users’ experience with others: short-term interaction when users collaborate to perform tasks, which requires attention and mutual awareness and is measured mainly on collaborative task performance; long-term interaction for socialising and entertainment via web-based virtual environments, which measures persistence of characters, of groups and of environment, social rules and convention, and effect of virtual world on real life; and the influence of the long-term use of immersive systems on performing short-term tasks.

Considering that many factors can influence UX in collaborative XR environments, the dynamic relationship amongst presence, social presence, and copresence needs to be studied together to understand how presence affects UX in general. Copresence and social presence are often considered as sub mental models of presence. Empirically, it is found that presence and copresence are positively correlated [89, 95, 109], and the same positive correlation occurs between presence and social presence [106]. However, other studies (e.g., [5, 26]) show that the relationship between presence and copresence is not definitively correlated or well defined. Schroeder in [91] proposed the concept of a connected presence cube. It maps presence, copresence and the extent of individual connected

presence to three dimensions of a cube representing the end-state of users' experience in shared virtual environments. He argued that the level of presence and copresence will be affected by the medium used to create the virtual world for users to feel a sense of connectedness such as desktop-, projection-, or HMD-based systems. Bulu in [24] suggested that all the three aspects will directly affect the satisfaction of users in immersive environments and they are all closely related in shared virtual environments. What has not been studied yet is how the different device settings of distant users can affect the UX and in particular the individual sense of presence, copresence, and social presence. Especially, XR technology can change the existing social psychology studies in the domain of user interaction and experience in real-and-virtual combined environments.

3.2 Group Dynamics and Collaboration Patterns

In this section, we look into the dynamics of how users work in groups and the collaboration patterns that users explicitly or implicitly employ. In social sciences, group dynamics studies human behaviours within a social group (intra-group) or between social groups (intergroup) [29, 105]. However, in collaborative XR systems, especially for remote collaboration, mediated environments change the way research findings in group dynamics are applied. Users working together over such environment are often considered as members of a virtual team. Virtual teams are “teams whose members use technology to varying degrees in working across locational, temporal, and relational boundaries to accomplish an interdependent task” [69]. The task performance of a virtual team is partly decided by how well workload is distributed, managed, and coordinated amongst team's members at group's level and partly by “the extent to which team members use virtual tools to coordinate and execute team processes” [58] at individual's level. It is important, therefore, to study how the collaboration occurs at individuals' level and group dynamics in fully or partly immersive systems.

Considering the processes of how each individual joins in groups or subgroups and how groups are formed over time, four models of change and continuity in group structure have been described [4].

- The first model that depicts stages in which different group structural patterns are formed is called *life cycle*. The Tuckman's four-stage model [111] is its known representative which summarises different stages of group development: *forming* for groups to identify the interpersonal and task behaviours and to establish dependency relationships with leaders, other group members, or predefined standards; *storming* for individuals to resolve interpersonal issues regarding to group influence and task requirements; *norming* to develop new standards for groups and to adopt new roles; and *performing* for groups to finalise interpersonal structure for task activities.
- The second model is *robust equilibrium*, which defines how a group's structure evolves through a short period of fluctuation followed by a stable state [28].
- Another developmental model is *punctuated equilibrium*, which indicates that groups develop through processes of sudden formation, maintenance, and

revision for performance by taking into account timing and mechanisms of change relating to the groups' context [46].

- The last model *adaptive response* describes groups' active changes to adapt to current task [99], technology [55] and environment [83] situations.

In the context of collaborative immersive systems, the Tuckman's four-stage model of groups [111] is often employed in designing communication and navigation mechanisms for users travelling in large-scale virtual world to be aware of other members' activities while performing collaborative tasks [32]. However, in our opinion, all the developmental models described above can be applied and reevaluated more extensively in novel mediated environments, which constitute challenges in creating effective workspace for virtual teams.

While group dynamics study how groups evolve over time under different situational factors, collaboration pattern research, on the other hand, looks into relationships between collaborators within groups and how they can adapt their behaviours to a collaborative task. Several studies in various domains have theorised different patterns and taxonomies of patterns of collaboration. For instance, in a collaboration systems for architectural designers, Caneparo [27] has explored the group coordination mechanism through four cases: *hierarchy order* when a leader of the group establishes the task's outlines and evaluates members' suggestions and performance, *individual initiative* when each member has their own freedom and acts independently, *participation* when members follow a working consensus build from discussion and negotiation, and *collaboration* when the group works on an agreed design solution after comparison and consensus. In the context of collaborative e-learning, Wasson and Mørch [117] have identified collaboration patterns occurring amongst students, teachers, and learning facilitators. The patterns consist of: *adaptation* when students working in groups in order to solve a common problem learn and adapt to others' behaviours, *coordinated desynchronisation* when group members coordinate activities after they have identified their common goal, *constructive commenting* when members give comments, and *informal language* when the relationship between group members become more intimate and is measured by the informal language they use.

In addition, another paradigm for collaboration patterns in product designing process was proposed in [65]. It considers four possible scenarios that can occur in group collaboration patterns: *peer-to-peer* when each member of the group contributes equally, *leader-member* when the leader of the group contributes more than other members, *complementary* when subgroups are formed to solve a portion of the task and their contributions are joined at the final stage, and *competitive* when subgroups are formed to compete with other subgroups by approaching the task from different angles.

Amongst all the four patterns defined above, *leader-member* collaboration pattern is one of the most studied topics. From social sciences' perspective, leadership is determined by traits and personality qualities inherent within certain individuals of a group [15]. Leadership skills, therefore, are often gradually developed outside and also within a group setting with or without the involvement of other group members. Competent leaders can help to build solid groups

to work productively. However, in many situations, the effectiveness of a group is decided not by the skills of the leader alone but also by the multilaterally shared responsibility in leader-member relationship. Depending on types of collaboration tasks, leader role can be implicitly or explicitly designated. When there is no predefined collaboration structure amongst members of a group, leadership can be regarded and evaluated through the contributions of each member to a shared collaborative task ('division of labor') and/or the act of taking charge by doing most of the talking ('talkativeness'), suggesting ideas, and giving instructions [5].

In real-life face-to-face circumstances, the location of an individual where they sit or stand can create direct assumptions from others about their leadership role [3, 52, 116]. However, in collaborative XR environments with limited access to non-verbal communication cues, different approaches have been employed by group members to determine or establish the leader-member relationship. Being virtually inhabited in virtual worlds, users are often represented and interact with others through '*avatars*' (see Sect. 3.3). Therefore, these avatars can have significant effect on others' perception about social behaviours and can determine collaboration mechanism between members. Yee and Bailenson in [119] studied the effect of height of users' avatars on their negotiation behaviour. This behaviour is a dominant personality trait of people with leadership skills because it is often associated with confidence, high self-esteem and ultimately leadership capability [103]. By isolating other factors which can affect the leadership behaviour in real world, such as age, gender, physical appearance, it shows that the impersonating tall avatar as self-representation of users can significantly increase their confidence in negotiation tasks. Additionally, other study [47] reports that the relative locations of the avatar representation of remote users within collaborative immersive environments should be appropriately chosen to make them appear in virtually equal size to improve their task performance, especially when they follow *peer-to-peer* collaboration pattern.

In shared virtual environments, users having advantages in computational performance, especially in level of immersion, are likely to emerge as leaders. Several studies [5, 94, 95, 98] report that without even being aware of others' working systems, users who were fully immersed were likely to be perceived as leaders and were rated high on talkativeness scores. In a more recent study, Pan et al. [77] have studied how two users collaborate in four different settings: AR to AR, AR to VR, AR to VR with virtual body, and AR to desktop. The results show that interactions in 3D could facilitate the emergence of leadership pattern, and that the more asymmetry in immersion level between collaborators, the stronger effect of leadership with users using AR interface of high level of immersion and situational awareness. However, if all the users share the same system capacity and are equally immersed, the leader role is often decided by the one who actively takes in the role of task navigator and manager [5].

As it is demonstrated with XR technology and its advantages in psychological therapies, the long-term effect of being confident in immersive environments compared to the sense of confidence and leadership skills in real life still needs to be extensively evaluated. In more general context, leadership skills are mostly

determined by personality traits of each individual and can be also attained by training. Therefore, the influence of personality on leadership in immersive systems needs to be studied to verify the correlations between leadership pattern and immersion levels using XR technology. For instance, in the experiment conducted by Slater [94], a questionnaire on Interaction Anxiousness Scale [64] was employed to measure participants' social anxiety, which inversely correlates to the degree of leadership. The results of the experiment have confirmed this special correlation between social anxiety, immersion, and leadership scores.

3.3 Avatars and Embodied Agents

Digital representations of users are an important factor to be considered while designing any collaborative XR platform. They help users to develop a sense of social connection with others, to be aware of others' presence and activities, and to have visual elements to focus on when they communicate with. Those representations can be categorised into: *avatars*, *embodied agents*, and *hybrid forms* [37]. The main difference between them is the control behind the representation. Avatar is a self-representation of a user who participates in the collaboration session in real time [6, 8, 37, 84]. Embodied agents, on the other hand, are controlled by computer algorithms to appear anthropomorphically and behave similarly like a human being. They are, therefore, defined as 'acting entities', whose behaviours are rendered based on simulation and Artificial Intelligence (AI) [35, 84]. An embodied agent has to be incorporated with four main capabilities in an adaptive functionality to be able to interact with humans in real-time: *perception*, *interpretation*, *reasoning*, and *autonomous responses* towards predefined goals [9]. Finally, hybrid combination of avatar and intelligent agent [86] is often employed in collaborative XR environments when the real presence and participation of users are not always guaranteed [44, 45] or when the use of AI algorithms helps to free user from fine-grained manipulations of avatars. In this section, we explore the usefulness of these virtual representations from two perspectives: how the use of avatars affect perception of users themselves (i.e., self-perception), and how users perceive others, either real-time collaborators or intelligent agents, through their visual representations.

Self-perception via Avatars. Generally, avatars represent people on social media and entertainment platforms such as online chat, video games, networking sites and online virtual worlds (e.g., Second Life⁶). Avatars, in a certain way, can be considered as a projection of users or an external self representation. In the immersive context, users can choose (passively or actively) how to represent themselves within limited options proposed by systems. Their representation can, in turn, influence their performance in executing tasks, communicating, as well as reflecting and perceiving of self independent of how other people perceive them. There are three types of avatar that can be employed: *authentic*, *modified* or *augmented*, and *non-anthropomorphic* or *novel* representation forms.

⁶ <https://secondlife.com/>.

The objective of providing *authentic* avatars is to guarantee high visual fidelity and behavioural authenticity of digital representations [115]. Researchers have tried to incorporate human physical capabilities in expressing nonverbal cues during conversations into digital models, giving avatars more faithful replication and realistic expressions and behaviours. There still are, however, several issues in designing and using avatars, including identity, awareness of current states, availability and degree of presence, gesture and facial expressions [14]. In a collaborative AR system, the self-presentation as an avatar besides the real body can potentially affect body ownership and self-localisation [85].

Modified or *augmented* representations of users are often used in evaluating the self-perception of people through the lens of their avatars. Yee & Bailenson have studied the Proteus effect, a hypothesis on the conformity of people's behaviours to their self-representations [119]. They have discovered that high level of attractiveness of avatar models can make people shorten their interpersonal distance [48]. Users can feel more intimate and open with others, and even the height of avatar can increase their confidence in a negotiation task. These results confirm the self-perception theory proposed by Bem on the dissimilarity in perception between the physical self and the digital modified self-representation [13]. Similarly, positive communication experience for users could be obtained by enhancing the smiling expressions of users through their avatars [76]. Furthermore, the negative effect of over-sexualised representations of women on sexual objectification and rape myth acceptance in virtual platforms has been also studied [38].

Non-anthropomorphic avatar approach represents users in a non-biological human form. This capability in mapping non-linearly the user's body with avatar's can facilitate novel form of interactions and manipulations that are not readily supported in conventional platforms. A concept of homuncular flexibility explores the idea of modifying representations of people to see how they can learn to control new form of avatars with extra limbs [62]. This concept has been further developed in extending avatar with a flexible tail attached to its coccyx [101], and alternating the visuomotor and visuotactile feedback of users' fingers via a six-finger illusion [53]. Verhulst et al. have studied how being embodied in an obese virtual body can help to change people's shopping behaviour [113]. The substitute for physical bodies with virtual ones is often measured by users' senses of ownership (i.e. perception of virtual parts of avatar as their own) and agency (i.e. perception of controlling these new forms). However, the extension of one's virtual body in collaborative context has not been extensively studied yet and it will be an important future research direction.

Perception of Others and Social Influence. Many researchers have studied the aspect of how users perceive others via avatars or visual representations and how that perception will influence social presence. Recent research has explored the potential of AI agents and social actors on the improvement of the social presence and perception of individuals within immersive environments [20]. For instance, the study conducted by Nowak & Biocca finds that people respond

socially to human and embodied agents alike in virtual world [75]. High level of copresence and social presence is also recorded when people interact with avatars of low anthropomorphic representations compared to realistic anthropomorphic images of the others, indicating a complex relationship between avatar representations and expectations from users when seeing them.

There are two main theoretical models that explain social influence of avatars and embodied agents on the social behaviour of human interlocutors. The first theoretical model by Nass and Moon in [72] concluded that if there are enough social cues in conversations, people will apply the same rules in real-life social interactions to interactions with agents even though they are aware that the experience is not real. Recently, this model has been revised, evaluated and confirmed [43, 84]. Blascovich et al. in [21], on the other hand, predicted that social influence within virtual environments will be decided by two additive factors (behavioural realism and social presence) and two moderating factors (self-relevance and target response system). They also argued that the social influence of a real person behind an avatar will always be higher than an embodied agent, and that the effect of an agent on social influence will depend on its behaviour realism. The hypothesis that avatars are more influential than agents on the social influence scale was confirmed in the research done by Fox et al. [37].

When integrating embodied agents into a collaborative scenario, many requirements are established to satisfy natural interaction with real-time users, which include life-like behaviours in conversations, responsiveness in a dynamic and unscripted environment, plausibility to create a sufficient illusion from users, and interpretable behaviour to allow users to interpret their responses [107]. For conversational agents, several frameworks for conversational interaction between an agent and a human user have been developed. For instance, FMTB (Functions, Modalities, Timing, Behaviours) conversational framework [30] supports conversational behaviours and actions via several modalities of communication such as hand gestures, facial expressions, eye gaze, etc. SmartBody [107] is another framework facilitating creation of animated conversational agents in real-time from hierarchically connected animation controllers. In general, besides the benefits of having automated agents as always-present interactive characters in virtual environments such as video games or online custom services, embodied agents can help to increase the experience of co-presence in shared environments, especially on social networking platforms [16]. Furthermore, it is argued that embodied agents may help people emote freely and reveal more sensitive information compared to conversational situations with real human users. For instance, perceived virtual human can help patients in clinical interviews disclose more sensitive information, hence overcoming the barrier between real and virtual actors behind mediated avatars [68].

Avatars play an important role in reinforcing the perception of others and social influence in collaborative environments. The effect of time and stage of the collaborative task on how users interact with others through avatars has been studied [97]. It is argued that when the collaboration time is short and users work together for the first time, they normally do not inquire about their

partners' avatars. The appearance of avatars get more attention when they collaborate for a longer period and the physical appearance of people behind the avatars becomes a topic of interest. Furthermore, the way people treat others' avatar varies from social discomfort and embarrassment when the avatars are in their interpersonal zone or overlapped in a desktop-based shared virtual environment [73,94] to unawareness and disinterest when they go through others' avatar while focusing on performing their task in a immersion projection technology system [97].

In conclusion, the effectiveness of avatar and embodied agents largely depends on their behaviour and appearance realism, and how they are used in different situations. Realism factor is often highly demanded in developing collaborative XR frameworks. However, there is also downside of the realism. Bailenson et al. in [8] found that people emote their feelings more freely when their avatar does not capture and express those emotions. In addition, the Uncanny Valley [71] predicts that negative experience can be evoked in human when robot appears and behaves too close to human-likeness. The same principle can be applied in the case of virtual characters or embodied agents. The study in [108] demonstrates that exaggerated facial expression via magnitude of mouth movements during speech to express different emotions can affect the uncanny for characters.

3.4 Nonverbal Communication

Verbal and nonverbal communication are considered absolutely essential in collaborative systems, whether they are designed for task solving, social networking or entertainment [59]. In problem solving systems, besides the main goal of helping users to convey information and keep in contact with others, communication channels provide means for them to understand the task, negotiate shared workload, form strategies, and be aware of what has been done and what is being done [74]. In general, there are several modalities that are available in 3D shared environments such as auditory channels, embodiment and nonverbal communication, text and 2D/3D annotation, and so forth. Additionally, they can be used explicitly or implicitly by remote users. Cassell et al. in [30] have distinguished between behaviour for propositional purposes and for interaction purposes of conversation. According to the authors, propositional purposes can be obtained through meaningful speech, hand gestures, and intonation to convey, complement, or elaborate upon the information being communicated. On the other hand, interactional functions serve to indicate the current state of the conversation and can include nonverbal cues such as head nods, raising hands, or eye gaze for conversation invitation, speaking turn-taking, feedback, breaking away behaviour in conversations. These two activities often occur simultaneously when speakers and listeners continuously monitor each other's behaviour and hence be able to contribute to the conversation depending on the course of conversation established through information delivered and decoded. In this section, we focus on the nonverbal communication channel for synchronous collaboration, how it has been supported in collaborative XR platforms, and how it can effect the performance of communication amongst users.

Complementing to auditory channels, nonverbal communication, or bodily communication, is defined as another means used by one person to influence others. According to Argyle [2], in face-to-face conversations, many nonverbal communication modalities are subtly employed at the same time including facial expression, gestures, eye gaze, bodily movements and contact, spatial behaviour, and nonverbal vocalisations. Nonverbal signals can be provided intentional or unconscious, and in many cases they can be the mixture of those two. There are mainly five functional types of nonverbal communications including *expressing emotions, communicating interpersonal attitudes, accompanying and supporting speech, self-presenting, and rituals*. In other words, nonverbal communication is multidimensional and multifunctional when several modalities (e.g., postures, gestures, eye gazes) can serve different functional types simultaneously [16].

Considering the important roles of nonverbal communication in collaborative XR environments, it is essential to capture nonverbal behaviour of users and replicate it, either faithfully or strategically, to other users. Avatars can be effectively used as a medium to transfer nonverbal cues if the users' body is being tracked partially or completely. If the avatars cannot fully represent the body and/or facial movements of users, they would have to learn to adapt to the missing nonverbal communication channel and convey their activities through verbal explanations [97]. In the case of lack of tracking system, nonverbal communication cues such as gestures or facial expressions can be preprogrammed and triggered via a text chat window during the interaction in a desktop-based virtual environment [110]. Amongst many modalities of nonverbal behaviour that is tracked and rendered in real time, head orientation and eye gaze are considered subtle but critical in providing bidirectional signals for monitoring and synchronising actions. Several studies have been conducted on the impact of eye gaze on communication [41, 42, 100]. The results show that even without eyelid movement and blinking behaviours implemented, representing users' eye gaze on their avatars in real time could improve the interaction between remote users and their collaborative task performance. Compared to static eye or simulated eye gaze integrated on avatars, using tracked eye gaze can help users to indicate and capture accurately focus of current attention, inform and estimate next actions, and effectively communicate. Furthermore, in one-to-many conversations, eye gaze can also be transformed and augmented so that the eye gaze of the speaker is rendered individually to each listener so that they would have an impression that the speaker is gazing at them only [7].

Another nonverbal cues that get attention from researchers are facial expressions, bodily movements, postures and gestures. Thanks to recent advances in real-time facial motion capture technology (e.g., Dynamixyz⁷, Faceware⁸, Facerig⁹) and 3D modelling, capturing facial expressions of users and rendering them realistically have become largely applicable. In a recent research work, Oh et al. investigated the enhanced smiling expression on communication experience [76].

⁷ <http://www.dynamixyz.com/>.

⁸ <http://facewaretech.com/products/software/>.

⁹ <https://facerig.com/>.

The users' smile is recorded and strategically rendered through their avatar. And when the participants' smile is enhanced, it is found that those participants themselves experienced stronger social presence compared to the faithful rendering condition. Another approach has explored three visual transformations for eye contact, joint attention identified by head direction, and grouping based on proxemic behavior to augment social behaviour by extending the physical communication condition into the virtual world [86]. Similarly, many approaches that communicate bodily movements, postures and gestures have been proposed. For instance, there are remote embodiment cues to improved awareness in a desktop-based virtual environment [39], hand movements of remote users via virtual hand shadows [88], remote user's head position, face direction, and hand poses for users using MR platform [80]. Recently, Pan et al. [78] integrated the foot tracking which allows users to see their full body in the shared VR environment, even though its impact on interaction, embodiment and presence is still subtle.

3.5 Does Group Size Matter? Collaboration and Social Interaction in Dyads, Triads, and Large Groups

The impact of different group size on collaboration mechanism, communication, and social interaction between users, especially remote users, in XR environments has not been extensively examined in the literature. Moreover, partially due to the limits of connection bandwidth and the large amount of data that needs to be transferred over the network to ensure a smooth collaboration, face-to-face or dyadic collaboration gets most of the attention from researchers. Since the nature of collaboration techniques in communication and interaction changes according to dyadic, triadic, and large group, we discuss in this section current research trends that have been explored for collaborative XR systems.

Communication patterns and group size have not been a highlighted topic and only limited research has considered the effects of group size on collaboration. In social sciences, it is concluded that increased group size decreases verbally interacting groups [99], individual contribution, perceived responsibility, involvement [63], and ideas generated per person [40]. Burgoon et al. in [25] have determined the limited number of members of a small group participating in a task without affecting interactivity and communication patterns. However, this limit depends on collaboration scenarios (co-located vs. remote) as well as the affordability of technology in supporting interdependent, contingent, participative, and synchronous interaction and communication between users. A theoretical model has been developed depicting the negative influence of group size and positive effect of social presence on the quality of communication within small groups regarding its appropriateness, richness, openness, and accuracy [67]. From the results of the experiment with 3-person and 6-person groups, it is argued that compared to 6-person groups, 3-person groups would obtain better communication in terms of appropriateness, openness, and accuracy.

In regard to dyadic interactions supported in collaborative virtual platforms, many behavioural model and interaction modes have been designed for face-to-face collaboration. Gaze and mutual gaze are the most important factors to be

considered in the nonverbal-behaviour-supported platforms. Indeed, Argyle & Cook have analysed closely the relationship between mutual gaze and conversation progress between two interlocutors [3]. Therefore, gaze behaviour has been strongly supported in collaborative virtual systems. For instance, an eye gaze model for dyadic interaction in shared virtual environments has been proposed as part of the support for avatar realism within negotiation scenarios [114]. Avatar realism and nonverbal communication in face-to-face social interactions have also been largely studied, which can be augmented or enhanced to improve user experience in dyadic interaction such as verbal and nonverbal communication, copresence, emotion recognition, and so forth [8, 41] (see Sect. 3.3 and 3.4). Furthermore, several social norms such as the gender, degree of intimacy, interpersonal distance, turn-talking in online virtual environments have also been studied [120]. More specifically in the context of cooperative manipulation and task solving, others factors such as concurrency control, collaborative manipulation mechanisms need to be taken into consideration. Regarding collaborative manipulation techniques, there are two main categories allowing users to concurrently and synchronously manipulating shared artefacts: splitting the degrees of freedom of the manipulated objects [36], and combining concurrent access to the same artefacts [87]. It is important to note that these two approaches for cooperative manipulation tasks do not limit to only two users but can be extended to multiple collaborators work jointly at the same time. The concurrency control at a higher level has been further investigated for peer-to-peer virtual environments [66]. Through a concurrency control hierarchy, three methods have been proposed to control sudden changes in closely-coupled, object-focuses tasks, which include Change It ('rollback' mechanism for simple shared object property changes without broadcasting updates), Grab It ('transaction-lock' mechanism for exclusive shared object property changes or deletions with broadcasting updates), and Build It ('intention-preservation' mechanism for shared object structure changes in highly dynamic environments).

Collaboration and communication within triad groups and small-size groups bear similar characteristics as in the dyadic groups in terms of cooperative manipulation and concurrency control. However, as there are more members participating in the session, social presence and interaction may change according to the nature of the collaborative and individual tasks as well as each member's roles. For instance, users' behaviour has been studied when they perform a task of puzzle solving in small groups of three people (one HMD and two desktop displays) within a shared virtual environment, and compared to their own behaviour when they continue doing the same task in the real world [95]. During the experiment, the experimenters also asked one member of the group to follow and observe another member without letting them know about it. The results in regard to the silent observation set-up show that shared VR platforms have the capacity to evoke emotional responses such as discomfort and embarrassment, even through simple avatars. Another experiment has been conducted by Steed et al. [98] in which leadership, presence, copresence, social presence, and accord between group members have been investigated within small groups of strangers

carrying out collaborative tasks. An overview done by Schroeder [89] lists several factors that need to be considered in order to improve user experience within small groups working together on short-term tasks. Those factors mostly serve synchronous collaboration purposes such as shared focus of attention, mutual awareness, and collaborative task performance.

Finally, regarding collaboration and interaction mechanism within large groups such as social networking or online virtual worlds for entertainment, researchers take a different approach in trying to understand how individuals within these groups form their relationships and adapt to the virtual environment over a long period; how being exposed to virtual worlds can affect their life in the real world; and which social rules are preserved or changed within the worlds of no boundaries [89]. Many parts of these research questions are still left unanswered and require extensive research effort in multiple disciplines. We discuss in this section some early works that have been performed to measure some social responses from an individual viewpoint when a user is interacting with or in front of a big group of others. In social psychology, Zajonc [121] and Taylor et al. [104] have reviewed and analysed the effect of performing a task in the presence of others on the user's performance which depends mostly on the difficulty level of the task and how the user has mastered it in advance. These analyses have been theorised into the concepts of social facilitation and inhibition. To apply and measure these concepts into the collaborative virtual environment, Hoyt et al. [54] have sought to replicate these effects and measure them in a study with participants performed a mastered and a non-mastered task, either alone or in the presence of a virtual human audience which was led to believe that they could be avatars or embodied agents. Their experiment has confirmed the social inhibition theory that performing a novel task in front of avatars can impair users' performance on subordinate responses. Furthermore, the behaviour of members in a big audience (e.g., eye contact, individual facial expressions, gesture, posture, and behavioural pattern between themselves) can also be registered as empirical design basis, which, in turn, can be useful to stimulate the users' experience in a virtual human audience [82].

These aspects relating to social responses mediated by XR technology are summarised in an attempt to identify what features that collaborative immersive systems can provide to make users experience and enjoy their time in immersive world and maximise their potential in using this world for different purposes.

3.6 Physical and Virtual World: How to Increase the Awareness

VR can forge a great sense of immersion in users when their senses (visual, auditory, and others) are replaced by synthesised digital channels. Different to presence experience, immersion is measured by objective technology-related factors such as field of view, field of regard, display resolution, head-based rendering, frame rate, and degree of interactivity [23]. Therefore, the more immersed users are in virtual environment, the more successful the system is in terms of isolating users from their physical world, increasing their perception of self-inclusion and self-movement [118]. However, since users still move in the physical world,

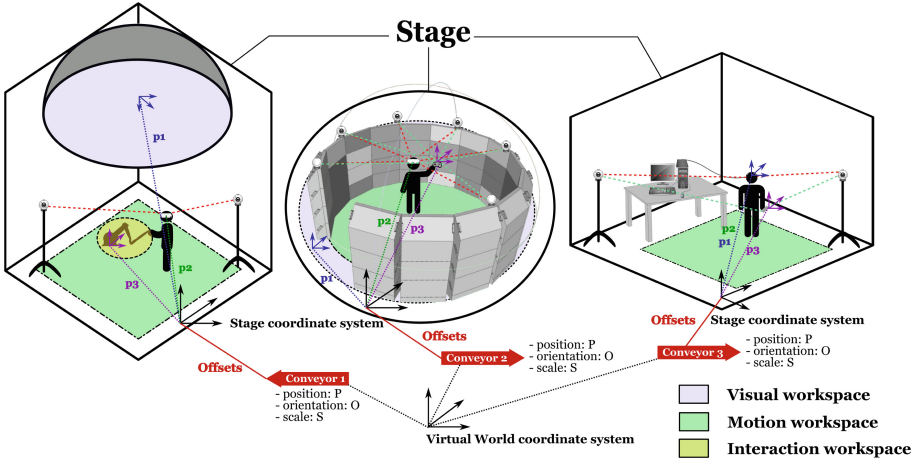


Fig. 4. IIVC model representing an abstraction of users’ physical environment including a conveyor, a stage and its workspaces for each user. It was adapted from [33,34] for a collaborative XR platform which includes hemispherical dome (left), 340-degree panoramic projection (middle), and HMD (right) systems [22]

any mismatches between physical and virtual world can break the illusion and even endanger them physically due to collision with physical objects in their surrounding area. In this section, we explore the idea of how to help users to be aware of the physical world while working in the virtual one without losing their immersion, presence and experience, and how to communicate the differences in hardware capabilities to remote collaborators.

The awareness of the physical environment with its constraints and limitations is essential when users are fully immersed. Steed et al. in [97] have pointed out several problems when the physical and virtual world do not align in projection-based systems such as the use of the non-tracked hand or the collision with the wall which to avoid users have to use their hand to feel the wall. Duval et al. in [34] have proposed the model of IIVC (Immersive Interactive Virtual Cabin) to encapsulate an abstraction of users’ physical environment and represent it in the virtual world. The IIVC comprises of three main components: workspace (3D space depicting physical area in which the user can move around or limits of physical devices), stage (virtual description of the user’ real environment), and conveyor (integration frame of the stage into the virtual world). Figure 4 shows in details the adapted IIVC model for a collaborative XR platform of hemispherical dome, panoramic projection, and HMD system, in which each system has a conveyor carries its stage and its workspaces. This model is useful in helping developers to precisely define the physical world’s parameters and integrate them into the virtual world. For instance, in order to enhance the awareness of the user when they work in a CAVE-like system and to prevent them from colliding with its front display screen, a 3D grid which becomes clearer and sharper when the user gets closer to the screen or physical boundaries [74].

By sharing the model or configurations of the user's physical space to others, it helps them to be aware of the working condition of others and can thus predict their possible limits and constraints. Explicit representations of field of view or grasping range are examples of how to communicate users' interaction abilities to others [39]. In asymmetric collaborative virtual environments, there also is a potential desynchronisation problem in coordinating activities between users in real time with different settings and viewpoints, which requires a mutual awareness to be established [31]. Piumsomboon et al. proposed and evaluated the effects of sharing awareness cues (field of view frustum, eye-gaze ray, head-gaze ray) on user performance in a collaborative MR system [81]. And in co-located shared VR environments, a research on mutual awareness has been conducted by Lacoche et al. in [60] for the collaboration between co-located users immersed via HMDs when they share the same physical space but navigate independently in the virtual world. Thus, there is a potential discrepancy in the perception and awareness of the physical and virtual world for co-located users. The authors have proposed and compared three approaches including Extended Grid (a grid cylinder representing physical location of others), Ghost Avatar (an avatar of the HMD model and its two controllers), and Safe Navigation Floor (a rendering of the physical floor with colors marking safe areas and collision zones where the others occupy). It is argued that these representations can also be used for co-located users even when they do not share the same virtual space, or for real static and dynamic objects if their position and occupied space can be tracked in the physical environment.

To conclude, the awareness in collaborative XR environment and of the difference between the physical and virtual environment is always essential to coordinate a group's activities no matter what the nature of the collaborative work is. Awareness of many factors and activities going on in the collaborative session can help to reduce errors and increase efficiency of the group effort. Despite the importance in facilitating a process to obtain the awareness, there are still many factors that have not been fully explored. For instance, Steed et al. have pointed out that when users interact with each other, they expect that others would grasp the context of their interaction and communication via gestures and bodily movements as well as their viewpoints implicitly [97]. Another factor to be considered to increase the presence of users in collaborative virtual environments is the discrepancy between physical moving in the real world and virtual traveling using different metaphors such as 'flying' [79] or 'jumping' [102], which can cause directional disorientation in spatial awareness. Finally, with the emerging technology in AR, the model that represents the physical environment in the virtual world needs to be revised to adapt to new collaborative XR platforms.

4 Research Opportunities

As coexperience in collaborative extended reality is a transdisciplinary research topic at the intersection of human-computer interaction, extended reality, computer-supported collaborative work, cognitive psychology, perception, and

social sciences, it is still challenging to fully identify all the pertinent research opportunities. Based on the analysis of several important aspects of human factors outlined in the previous section, we aim to encapsulate main directions in a non-exclusive list in this section for future research projects.

Presence, Copresence, and Social Presence. In the new context of XR platforms, these three factors relating to presence can be re-explored and assessed as XR technology, especially AR, has changed the nature of communication approach for remote collaboration. In the near future, XR technology will be able to provide high fidelity communication cues including virtual proximity and orientation of others, physical appearance, facial expressions, gaze and mutual gaze, postures and gestures, as well as verbal signals. However, there is still a lack of (explicit and implicit) exchange and integration mechanisms of these communication channels and representations of communicational cues in mixed real-virtual environments. Furthermore, dynamic relationship between presence, social presence, copresence and UX needs to be studied further in order to determine decisive factors to be considered when designing user-centred MR systems.

Asynchronous Collaboration. This mode of collaboration will become prevalent once the use of XR in collaborative work expands in the future. Asynchronous collaboration of distant users can affect UX and particularly will influence individual sense of presence and social connection. XR technology will, therefore, change the methodologies of social psychology studies in the domain of user interaction and UX over virtual and augmented environments.

Long-Term Effects of Collaborative Extended Reality on User Experience. From the aspect of using immersive environments in social life, long-term effects of being immersed in such environments and working together with others over a distance on individual personality traits such as social anxiety and leadership skills can become an interesting research undertaking between social scientists, cognitive psychologists, and computer researchers. Moreover, considering collaboration and interaction mechanisms within large group of users when XR platforms are used for social networking, online virtual worlds and entertainment, it is important to understand how individuals within these groups form their relationships and adapt to the mediated environment over a long period, how being exposed to virtual worlds can affect their life in the real world, and which social rules are preserved or changed within the virtual worlds of no boundaries [89]. Many parts of these research questions are still left unanswered and require extensive research effort combined from multiple disciplines.

Group Dynamics and Collaboration Patterns. Task performance of a whole group is partly decided by how well workload is distributed, managed, and coordinated amongst members and partly by how well each member uses tools to coordinate and execute tasks. Therefore, it is an important topic to study how the collaboration occurs at individuals' level and group dynamics in collaborative mixed immersive systems. Moreover, studies can be carried out to measure collaboration performance and competition within a group and between groups.

Virtual Representations of Self and of Others. The use of virtual bodies of users as the representation of self on computer-mediated environments can change their sense of ownership (i.e. perception that extends or modifies virtual parts of avatar as their own) and agency (i.e. perception of controlling these new forms). Accurate representations of users within XR environments extracted from all the tracking systems have the potential to be able to render highly realistic models to facilitate real-time face-to-face interaction and communication between users. The extension of one's virtual body in collaborative context can be extensively studied in the near future for a more complete understanding of how each user perceives and experiences within XR environments. For instance, in a collaborative AR world, a research question can be how the representation of both physical and virtual body can affect the user's self-perception and self-localisation. In addition, another aspect of using virtual representations of avatars and embodied agents with their behavioural realism and appearance realism can be broadly studied on the measures of collaborative tasks performance and UX.

Merging of Physical and Virtual Worlds. The integration of physical world into the virtual world and how the virtual world manifests itself in the physical world need to be revised to adapt to collaborative XR systems within which multiple users have their own hardware capabilities and may not be aware of the differences between them. Since more and more XR devices have been marketed to larger public, it would be necessary to study how mediated environments created by HMD, smart glasses, projection-based or CAVE-like systems can be perceived individually by each user and how these differences in display and interaction devices will affect users' roles in the whole collaboration process.

5 Conclusion

Multiple user experience or coexperience in collaborative extended reality environments is an important topic that requires synergistic research collaboration amongst cognitive psychologists and social scientists, human-computer interaction researchers and designers, extended reality (virtual reality and augmented reality) scientists and developers, data scientists, amongst others. The future outcomes of this research will facilitate greatly the interaction between humans and computer-generated worlds through multi-sensory stimuli interaction design and exchange mechanisms. We present in this paper several main aspects of coexperience in collaborative extended reality environments including presence-related factors, group dynamics and collaboration patterns, avatars and embodied agents, nonverbal communication, group size, awareness of physical and virtual world as the initiative to review the current state of the art of this multi-disciplinary research domain. Many future research opportunities are outlined in the previous section that could be of interest to researchers and scientists in different fields. There are still many unexplored topics in this multi-discipline domain and great research effort, resources, and collaboration need to be initiated to solve these challenges as collaboration between users, especially remotely located users, is technologically challenging in providing seamlessly transferring communication, manipulation and task execution process.

References

1. Allport, G.: *The Historical Background of Modern Social Psychology*, vol. I, 3rd edn., pp. 1–46. Addison-Wesley, Boston (1985)
2. Argyle, M.: *Bodily Communication*, 2nd edn. International Universities Press Inc. (1988)
3. Argyle, M., Cook, M.: *Gaze and Mutual Gaze*. Cambridge University Press, New York (1976)
4. Arrow, H.: Stability, bistability, and instability in small group influence patterns. *J. Pers. Soc. Psychol.* **72**(1), 75–85 (1997)
5. Axelsson, A.S., Abelin, A., Heldal, I., Schroeder, R., Widestrom, J.: Cubes in the cube: a comparison of a puzzle-solving task in a virtual and a real environment. *CyberPsychol. Behav.* **4**(2), 279–286 (2001)
6. Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., Turk, M.: Transformed social interaction: decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoper. Virtual Environ.* **13**(4), 428–441 (2004)
7. Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., Turk, M.: Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Hum. Commun. Res.* **31**(4), 511–537 (2005)
8. Bailenson, J.N., Yee, N., Merget, D., Schroeder, R.: The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoper. Virtual Environ.* **15**(4), 359–372 (2006)
9. Balakrishnan, K., Honavar, V.: *Evolutionary and Neural Synthesis of Intelligent Agents*, pp. 1–26. MIT Press, Cambridge (2001). Book section 1
10. Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: *Proceedings of the 2009 AAI Conference on Weblogs and Social Media*, pp. 361–362 (2009)
11. Battarbee, K., Koskinen, I.: Co-experience: user experience as interaction. *CoDesign* **1**(1), 5–18 (2005)
12. de Belen, R.A.J., Nguyen, H., Filonik, D., Del Favero, D., Bednarz, T.: A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electron. Electr. Eng.* **3**, 181 (2019)
13. Bem, D.J.: *Self-perception Theory*, vol. 6, pp. 1–62. Academic Press, Cambridge (1972)
14. Benford, S., Bowers, J., Fahlen, L.E., Greenhalgh, C., Snowdon, D.: User embodiment in collaborative virtual environments. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 242–249. ACM Press/Addison-Wesley Publishing Co. (1995). 223935
15. Benne, K.D., Sheats, P.: Functional roles of group members. *J. Soc. Issues* **4**(2), 41–49 (1948)
16. Bente, G., Rüggenberg, S., Krämer, N.C., Eschenburg, F.: Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. *Hum. Commun. Res.* **34**(2), 287–318 (2008)
17. Billinghurst, M., Kato, H.: Collaborative mixed reality. In: *Proceedings of the First International Symposium on Mixed Reality (ISMR)*, pp. 261–284 (1999)
18. Biocca, F., Harms, C.: Defining and measuring social presence: contribution to the networked minds theory and measure. In: *Proceedings of the 5th International Workshop on Presence*, pp. 7–36 (2002)

19. Biocca, F., Harms, C., Burgoon, J.K.: Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoper. Virtual Environ.* **12**(5), 456–480 (2003)
20. Blascovich, J.: Social influence within immersive virtual environments. In: Schroeder, R. (ed.) *The Social Life of Avatars. Computer Supported Cooperative Work*, pp. 127–145. Springer, London (2002). https://doi.org/10.1007/978-1-4471-0277-9_8
21. Blascovich, J., Loomis, J., Beall, A.C., Swinth, K.R., Hoyt, C.L., Bailenson, J.N.: Immersive virtual environment technology as a methodological tool for social psychology. *Psychol. Inq.* **13**(2), 103–124 (2002)
22. Bourke, C., Bednarz, T.: Multi-modal high-end visualization system. In: *Proceedings of the 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry, VRCAI 2019*, pp. 1–2. Association for Computing Machinery (2019)
23. Bowman, D., McMahan, R.: Virtual reality: how much immersion is enough? *Computer* **40**(7), 36–43 (2007)
24. Bulu, S.T.: Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Comput. Educ.* **58**(1), 154–161 (2012)
25. Burgoon, J.K., Bonito, J.A., Ramirez, J.A., Dunbar, N.E., Kam, K., Fischer, J.: Testing the interactivity principle: effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. *J. Commun.* **52**(3), 657–677 (2002)
26. Bystrom, K.E., Barfield, W.: Collaborative task performance for learning using a virtual environment. *Presence: Teleoper. Virtual Environ.* **8**(4), 435–448 (1999)
27. Caneparo, L.: Coordinative virtual space for architectural design. In: *Proceedings of the 6th International Conference on Computer-Aided Architectural Design Futures. CAAD Futures* (1995)
28. Carley, K.: A theory of group stability. *Am. Sociol. Rev.* **56**(3), 331–354 (1991)
29. Cartwright, D., Zander, A.: *Group Dynamics*, 3rd edn. Harper + Row, Oxford (1968)
30. Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., Yan, H.: Human conversation as a system framework: designing embodied conversational agents, pp. 29–63. MIT Press (2000)
31. Chenechal, M.L., Chalme, S., Duval, T., Royan, J., Gouranton, V., Arnaldi, B.: Toward an enhanced mutual awareness in asymmetric CVE. In: *Proceedings of the 2015 International Conference on Collaboration Technologies and Systems (CTS)*, pp. 233–240 (2015)
32. Dodds, T.J., Ruddle, R.A.: Mobile group dynamics in large-scale collaborative virtual environments. In: *Proceedings of the 2008 IEEE Virtual Reality Conference*, pp. 59–66 (2008)
33. Duval, T., Nguyen, H., Fleury, C., Chauffaut, A., Dumont, G., Gouranton, V.: Embedding the features of the users’ physical environments to improve the feeling of presence in collaborative virtual environments. In: *Proceedings of the IEEE 3rd International Conference on Cognitive Infocommunications*, pp. 243–248. CogInfoCom (2012)
34. Duval, T., Nguyen, T.T.H., Fleury, C., Chauffaut, A., Dumont, G., Gouranton, V.: Improving awareness for 3D virtual collaboration by embedding the features of users’ physical environments and by augmenting interaction tools with cognitive feedback cues. *J. Multimodal User Interfaces* **8**(2), 187–197 (2014)
35. Erickson, T.: *Designing agents as if people mattered*, pp. 79–96. MIT Press (1997)

36. Fleury, C., Duval, T., Gouranton, V., Steed, A.: Evaluation of remote collaborative manipulation for scientific data analysis. In: Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology, pp. 129–136. ACM (2012). 2407361
37. Fox, J., Ahn, S.J., Janssen, J.H., Yeykelis, L., Segovia, K.Y., Bailenson, J.N.: Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. *Hum. Comput. Interact.* **30**(5), 401–432 (2015)
38. Fox, J., Bailenson, J.N., Tricase, L.: The embodiment of sexualized virtual selves: the proteus effect and experiences of self-objectification via avatars. *Comput. Hum. Behav.* **29**(3), 930–938 (2013)
39. Fraser, M., Benford, S., Hindmarsh, J., Heath, C.: Supporting awareness and interaction through collaborative virtual interfaces. In: Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology, pp. 27–36. ACM (1999). 322580
40. Gallupe, R.B., et al.: Electronic brainstorming and group size. *Acad. Manag. J.* **35**(2), 350–369 (1992)
41. Garau, M., Slater, M., V., Brogni, A., Steed, A., Sasse, M.A.: The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 529–536. ACM (2003). 642703
42. Garau, M., Slater, M., Bee, S., Sasse, M.A.: The impact of eye gaze on communication using humanoid avatars. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 309–316. ACM (2001). 365121
43. Garau, M., Slater, M., Pertaub, D.P., Razzaque, S.: The responses of people to virtual humans in an immersive virtual environment. *Presence: Teleoper. Virtual Environ. - Spec. Issue: Collabor. Inf. Vis. Environ.* **14**(1), 104–116 (2005)
44. Gerhard, M., Moore, D.J., Hobbs, D.J.: Continuous presence in collaborative virtual environments: towards a hybrid avatar-agent model for user representation. In: de Antonio, A., Aylett, R., Ballin, D. (eds.) IVA 2001. LNCS (LNAI), vol. 2190, pp. 137–155. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44812-8_12
45. Gerhard, M., Moore, D., Hobbs, D.: Embodiment and copresence in collaborative interfaces. *Int. J. Hum. Comput. Stud.* **61**(4), 453–480 (2004)
46. Gersick, C.J.G.: Time and transition in work teams: toward a new model of group development. *Acad. Manag. J.* **31**(1), 9–41 (1988)
47. Goebbels, G., Lalioti, V.: Co-presence and co-working in distributed collaborative virtual environments. In: Proceedings of the 1st International Conference on Computer Graphics, Virtual Reality and Visualisation, pp. 109–114. ACM (2001). 513891
48. Hall, E.T.: *The Hidden Dimension*. Doubleday, Garden City (1966)
49. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. *Behav. Inf. Technol.* **25**(2), 91–97 (2006)
50. Heeter, C.: Being there: the subjective experience of presence. *Presence: Teleoper. Virtual Environ.* **1**(2), 262–271 (1992)
51. Held, R.M., Durlach, N.I.: Telepresence. *Presence: Teleoper. Virtual Environ.* **1**(1), 109–112 (1992)
52. Howells, L.T., Becker, S.W.: Seating arrangement and leadership emergence. *J. Abnorm. Soc. Psychol.* **64**(2), 148–150 (1962)
53. Hoyet, L., Argelaguet, F., Nicole, C., Lécuyer, A.: "Wow! I have six fingers!": would you accept structural changes of your hand in VR? *Front. Robot. AI* **3**(27), 1–12 (2016)

54. Hoyt, C.L., Blascovich, J., Swinth, K.R.: Social inhibition in immersive virtual environments. *Presence: Teleoper. Virtual Environ.* **12**(2), 183–195 (2003)
55. Hulin, C.L., Roznowski, M.: Organizational technologies: effects on organizations' characteristics and individuals' responses. *Res. Organ. Behav.* **7**, 39–85 (1985)
56. IJsselsteijn, W.A., Ridder, H.d., Freeman, J., Avons, S.E.: Presence: concept, determinants, and measurement. In: *Proceedings of the Electronic Imaging conference*, vol. 3959, p. 10. SPIE (2000)
57. Johnson-Glenberg, M.C., Birchfield, D.A., Tolentino, L., Koziupa, T.: Collaborative embodied learning in mixed reality motion-capture environments: two science studies. *J. Educ. Psychol.* **106**(1), 86–104 (2014)
58. Kirkman, B.L., Mathieu, J.E.: The dimensions and antecedents of team virtuality. *J. Manag.* **31**(5), 700–718 (2005)
59. Knapp, M.L., Hall, J.A., Horgan, T.G.: *Nonverbal Communication in Human Interaction*, 8th edn. Wadsworth CENGAGE Learning (2013)
60. Lacoche, J., Pallamin, N., Boggini, T., Royan, J.: Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In: *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–9. ACM (2017). 3139142
61. Ladwig, P., Geiger, C.: A literature review on collaboration in mixed reality. In: Auer, M.E., Langmann, R. (eds.) *REV 2018. LNNS*, vol. 47, pp. 591–600. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-95678-7_65
62. Lanier, J.: *Homuncular flexibility* (2006)
63. Latane, B., Wolf, S.: The social impact of majorities and minorities. *Psychol. Rev.* **88**(5), 438–453 (1981)
64. Leary, M.R.: Social anxiousness: the construct and its measurement. *J. Pers. Assess.* **47**(1), 66–75 (1983)
65. Linebarger, J.M., Janneck, C.D., Kessler, G.D.: Leaving the world behind: supporting group collaboration patterns in a shared virtual environment for product design. *Presence: Teleoper. Virtual Environ.* **14**(6), 697–719 (2005)
66. Linebarger, J.M., Kessler, G.D.: Concurrency control mechanisms for closely coupled collaboration in multithreaded peer-to-peer virtual environments. *Presence: Teleoper. Virtual Environ.* **13**(3), 296–314 (2004)
67. Lowry, P.B., Roberts, T.L., Romano, N.C., Cheney, P.D., Hightower, R.T.: The impact of group size and social presence on small-group communication: does computer-mediated communication make a difference? *Small Group Res.* **37**(6), 631–661 (2006)
68. Lucas, G.M., Gratch, J., King, A., Morency, L.P.: It's only a computer: virtual humans increase willingness to disclose. *Comput. Hum. Behav.* **37**, 94–100 (2014)
69. Martins, L.L., Gilson, L.L., Maynard, M.T.: Virtual teams: what do we know and where do we go from here? *J. Manag.* **30**(6), 805–835 (2004)
70. Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: a class of displays on the reality-virtuality continuum. In: *Photonics for Industrial Applications*, vol. 2351, pp. 282–292. SPIE (1995)
71. Mori, M.: Bukimi no tani (the uncanny valley). *Energy* **7**, 33–35 (1970)
72. Nass, C., Moon, Y.: Machines and mindlessness: social responses to computers. *J. Soc. Issues* **56**(1), 81–103 (2000)
73. Nassiri, N., Powell, N., Moore, D.: Avatar gender and personal space invasion anxiety level in desktop collaborative virtual environments. *Virtual Reality* **8**(2), 107–117 (2004)

74. Nguyen, T.T.H., Duval, T.: A survey of communication and awareness in collaborative virtual environments. In: Proceedings of the 2014 International Workshop on Collaborative Virtual Environments, pp. 1–8. 3DCVE (2014)
75. Nowak, K.L., Biocca, F.: The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoper. Virtual Environ.* **12**(5), 481–494 (2003)
76. Oh, S.Y., Bailenson, J., Krämer, N., Li, B.: Let the avatar brighten your smile: effects of enhancing facial expressions in virtual environments. *PLOS One* **11**(9), e0161794 (2016)
77. Pan, Y., Sinclair, D., Mitchell, K.: Empowerment and embodiment for collaborative mixed reality systems. *Comput. Animat. Virtual Worlds* **29**(3–4), 1–11 (2018)
78. Pan, Y., Steed, A.: How foot tracking matters: the impact of an animated self-avatar on interaction, embodiment and presence in shared virtual environments. *Front. Robot. AI* **6**, 1–13 (2019)
79. Pausch, R., Burnette, T., Brockway, D., Weiblen, M.E.: Navigation and locomotion in virtual worlds via flight into hand-held miniatures. In: Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, pp. 399–400. ACM (1995). 218495
80. Piumsomboon, T., Day, A., Ens, B., Lee, Y., Lee, G., Billingham, M.: Exploring enhancements for remote mixed reality collaboration. In: Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, pp. 1–5. ACM. (2017). 3139200
81. Piumsomboon, T., Dey, A., Ens, B., Lee, G., Billingham, M.: The effects of sharing awareness cues in collaborative mixed reality. *Front. Robot. AI* **6**, 1–18 (2019)
82. Poeschl, S., Doering, N.: Designing virtual audiences for fear of public speaking training - an observation study on realistic nonverbal behavior. *Stud. Health Technol. Inf.* **181**, 218–222 (2012)
83. Poole, M.S., Roth, J.: Decision development in small groups V test of a contingency model. *Hum. Commun. Res.* **15**(4), 549–589 (1989)
84. von der Putten, A.M., Kramer, N.C., Gratch, J., Kang, S.H.: “it doesn't matter what you are!” explaining social effects of agents and avatars. *Comput. Hum. Behav.* **26**(6), 1641–1650 (2010)
85. Rosa, N.: Player/avatar body relations in multimodal augmented reality games. In: Proceedings of the 18th ACM International Conference on Multimodal Interaction, ICMI 2016, pp. 550–553 (2016)
86. Roth, D., Kleinbeck, C., Feigl, T., Mutschler, C., Latoschik, M.E.: Beyond replication: augmenting social behaviors in multi-user virtual realities. In: Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 215–222 (2018)
87. Ruddle, R.A., Savage, J.C.D., Jones, D.M.: Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments. *J. ACM Trans. Comput.-Hum. Interact. (TOCHI)* **9**(4), 285–308 (2002)
88. Sakong, K., Nam, T.J.: Supporting telepresence by visual and physical cues in distributed 3D collaborative design environments. In: Proceedings of CHI 2006 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2006, pp. 1283–1288 (2006)
89. Schroeder, R.: Copresence and interaction in virtual environments: an overview of the range of issues. In: Proceedings of the 5th International Workshop on Presence, pp. 274–295 (2002)

90. Schroeder, R.: social interaction in virtual environments: key issues, common themes, and a framework for research. In: Schroeder, R. (ed.) *The Social Life of Avatars*. Computer Supported Cooperative Work, pp. 1–18. Springer, London (2002). https://doi.org/10.1007/978-1-4471-0277-9_1
91. Schroeder, R.: Being there together and the future of connected presence. *Presence: Teleoper. Virtual Environ.* **15**(4), 438–454 (2006)
92. Schuemie, M.J., van der Straaten, P., Krijn, M., van der Mast, C.A.: Research on presence in virtual reality: a survey. *CyberPsychol. Behav.* **4**(2), 183–201 (2001)
93. Shen, Z., Ma, H., Wang, K.: A web-scale system for scientific knowledge exploration. In: *Proceedings of the 2018 Meeting of the Association for Computational Linguistics*, pp. 87–92 (2018)
94. Slater, M.: Real people meeting virtually real people – a review of some experiments in shared virtual environments. *BT Technol. J.* **17**(1), 120–127 (1999)
95. Slater, M., Sadagic, A., Usoh, M., Schroeder, R.: Small-group behavior in a virtual and real environment: a comparative study. *Presence: Teleoper. Virtual Environ.* **9**(1), 37–51 (2000)
96. Slater, M., Wilbur, S.: A framework for immersive virtual environments (five): speculations on the role of presence in virtual environments. *Presence: Teleoper. Virtual Environ.* **6**(6), 603–616 (1997)
97. Steed, A., Roberts, D., Schroeder, R., Heldal, I.: Interaction between users of immersion projection technology systems. In: *Proceedings of the 11th International Conference on Human Computer Interaction (HCI)*, pp. 1–10 (2005)
98. Steed, A., Slater, M., Sadagic, A., Bullock, A., Tromp, J.: Leadership and collaboration in shared virtual environments. In: *Proceedings of the IEEE Virtual Reality*, pp. 112–115 (1999)
99. Steiner, I.D.: *Group Process and Productivity (Social Psychological Monograph)*. Academic Press, New York (1972)
100. Steptoe, W., et al.: Eye tracking for avatar eye gaze control during object-focused multiparty interaction in immersive collaborative virtual environments. In: *Proceedings of the 2009 IEEE Virtual Reality Conference*, pp. 83–90 (2009)
101. Steptoe, W., Steed, A., Slater, M.: Human tails: ownership and control of extended humanoid avatars. *IEEE Trans. Vis. Comput. Graph. (TVCG)* **19**(4), 583–590 (2013)
102. Stoakley, R., Conway, M.J., Pausch, R.: Virtual reality on a WIM: interactive worlds in miniature. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 265–272. ACM Press/Addison-Wesley Publishing Co. (1995). 223938
103. Stogdill, R.M.: Personal factors associated with leadership: a survey of the literature. *J. Psychol.* **25**(1), 35–71 (1948)
104. Taylor, S.E., Peplau, L.A., Sears, D.O.: *Social Psychology*, 12th edn. Pearson, London (2005)
105. Thibaut, J.W., Kelley, H.H.: *The Social Psychology of Groups*, 1st edn. Routledge, New York (1959)
106. Thie, S., Wijk, J.V.: A general theory on presence: experimental evaluation of social virtual presence in a decision making task. In: *Proceedings of the Presence in Shared Virtual Environments Workshop*, pp. 53–63 (1998)
107. Thiebaut, M., Marsella, S., Marshall, A.N., Kallmann, M.: SmartBody: behavior realization for embodied conversational agents. In: *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems*, vol. 1. pp. 151–158. International Foundation for Autonomous Agents and Multiagent Systems (2008). 1402409

108. Tinwell, A., Grimshaw, M., Abdel-Nabi, D.: Effect of emotion and articulation of speech on the uncanny valley in virtual characters. In: D’Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) *ACII 2011*. LNCS, vol. 6975, pp. 557–566. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-24571-8_69
109. Tromp, J., Bullock, A., Steed, A., Sadagic, A., Slater, M., Frecon, E.: Small group behaviour experiments in the coven project. *IEEE Comput. Graph. Appl.* **18**(6), 53–63 (1998)
110. Tsiatsos, T., Terzidou, T.: Supporting collaborative learning processes in CVEs by augmenting student avatars, with nonverbal communication features. In: *Proceedings of the 10th IEEE International Conference on Advanced Learning Technologies*, pp. 578–580 (2010)
111. Tuckman, B.W.: Developmental sequence in small groups. *Psychol. Bull.* **63**(6), 384–399 (1965)
112. Tullis, T., Albert, B.: *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*, 2nd edn. Elsevier Inc., Amsterdam (2013)
113. Verhulst, A., Normand, J.M., Lombart, C., Sugimoto, M., Moreau, G.: Influence of being embodied in an obese virtual body on shopping behavior and products perception in VR. *Front. Robot. AI* **5**, 1–20 (2018)
114. Vinayagamorthy, V., Garau, M., Steed, A., Slater, M.: An eye gaze model for dyadic interaction in an immersive virtual environment: practice and experience. *Comput. Graph. Forum* **23**(1), 1–11 (2004)
115. Vinayagamorthy, V., Steed, A., Slater, M.: Building characters: lessons drawn from virtual environments. In: *Proceedings of the Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop*, pp. 119–126 (2005)
116. Ward, C.D.: Seating arrangement and leadership emergence in small discussion groups. *J. Soc. Psychol.* **74**(1), 83–90 (1968)
117. Wasson, B., Mørch, A.I.: Identifying collaboration patterns in collaborative telelearning scenarios. *J. Educ. Technol. Soc.* **3**(3), 237–248 (2000)
118. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. *Presence: Teleoper. Virtual Environ.* **7**(3), 225–240 (1998)
119. Yee, N., Bailenson, J.: The proteus effect: the effect of transformed self-representation on behavior. *Hum. Commun. Res.* **33**(3), 271–290 (2007)
120. Yee, N., Bailenson, J.N., Urbanek, M., Chang, F., Merget, D.: The unbearable likeness of being digital: the persistence of nonverbal social norms in online virtual environments. *CyberPsychol. Behav.* **10**(1), 115–121 (2007)
121. Zajonc, R.B.: Social facilitation. *Science* **149**(3681), 269–274 (1965)