Exploring How Role and Background Influence Through Analysis of Spatial Dialogue in Collaborative Problem-Solving Games

Cigdem Uz-Bilgin\textsuperscript{1} \& Meredith Thompson\textsuperscript{1} \& Melat Anteneh\textsuperscript{1}

Published online: 14 September 2020
© Springer Nature B.V. 2020

Abstract
This study examines how different roles and background knowledge transform players’ dyadic conversations into spatial dialogues in a virtual cellular biology game. Cellverse is a collaborative virtual reality (VR) game designed to teach cell biology. Players work in pairs, assuming the role of either a Navigator, with reference material and a global view through a tablet, or an Explorer, with a more detailed interactive view of the cell through a VR headset and hand controllers. The game is designed so players must collaborate in order to complete the game. Our results show that roles influenced their reference perspectives at a level of statistical significance. Furthermore, players with high prior knowledge tried to reduce their partner’s mental effort by giving spatial information from their point of view, thus producing fewer occurrences of spatial unawareness. Results of this study suggest that designers can build in different roles and leverage different background knowledge to prompt effective partnerships during collaborative games.

Keywords Spatial dialogue \& Collaborative games \& Virtual reality \& Prior knowledge \& Science education

Introduction
Collaborative problem-solving (CPS) is an essential skill in education and the workplace, yet CPS skills are challenging to develop and to assess (Fiore et al. 2017). Virtual simulations may provide an avenue for both developing and measuring twenty-first century skills such as CPS. Immersive virtual reality (VR) has the potential to create a shared environment that could enable users to collaborate and evaluate their performances (Ens et al. 2019). An important step in CPS is being able to understand the problem and establish a common language (Duncan and West 2018; Sawyer 2017). This study sets up a CPS scenario where pairs of players collaboratively explore a shared environment from two different viewpoints: an Explorer embedded within a cell and a Navigator who has a more global, yet less detailed view. Having clearly defined roles and a range of expertise helps foster “positive interdependence” among team members (Johnson et al. 1991) where in successful completion of the activity requires a joint effort (Weber and Kim 2015). Players must find ways to describe the strange and unfamiliar environment of a human lung cell to a partner when neither partner has a full understanding of what the other one can see. In this study, we focus on how these players establish a shared understanding through the conversations about the environment during the VR-based game.

VR is a simulated environment that creates realistic experiences by providing users with regular sensory feedback (Johnson et al. 2016). VR has been used for education and training in a wide variety of fields including medicine (Abe et al. 2018), education (Freina and Ott 2015), gaming, and entertainment (Liszio and Masuch 2016), and to assist in spatial learning (León et al. 2018; Tascón et al. 2017). A properly designed VR environment can be capable of evoking a sense of presence and immersion in users interacting with and within it (Sherman and Craig 2018). Spatial presence, the perception of existing within a space, and spatial immersion, the feeling of being physically present in a space, are important contributors to users’ enjoyment of VR environments (Shafer et al. 2019). When implemented correctly, VR environments can generate levels of engagement that surpass those of traditional screen-based content (Zaman et al. 2015). The degree to which users are immersed in the environment and feel that

\textsuperscript{1} The Education Arcade, Massachusetts Institute of Technology, Cambridge, MA, USA

\textsuperscript{2} Cigdem Uz-Bilgin
uze@cigdem@gmail.com; cigdemuz@mit.edu
they are physically present depends on several factors including how realistically users are represented in the environment (e.g., avatars) and what equipment is used to interact with the environment (e.g., head-mounted displays) (Seibert and Shafer 2017; Zaman et al. 2015). The majority of VR setups utilize a headset and some sort of handheld controller used to navigate within and interact with the environment (Hahn 2017).

Witmer and Singer (1998) divide spatial presence in VR into three contributing (control, sensory, and realism) and one detracting (distraction) components. Although spatial presence is typically measured using subjective measures such as scales and questionnaires, researchers have also developed some performance and psychological ways to measure spatial presence (Laari et al. 2015). Minimizing distractions and maximizing users’ ability to selectively focus on the environment has been found to improve spatial presence (Hite et al. 2019; Tussyadiah et al. 2017). Hite et al. (2019) found that spatial acuity, spatial rotation ability, and understanding of angular geometry all contributed to students’ feelings of presence during a VR science lesson. Coxon et al. (2016) found a correlation between spatial presence and self-reported visuo-spatial imagery as measured by the MEC-SPQ’s visual spatial imagery scale.

Many of the same strategies used when navigating in the real world can be applied to virtual environments. Easily navigable spaces, virtual or otherwise, are designed in such a way that facilitates users’ spatial orientation, the ability to determine one’s position and heading towards a destination in an environment (Pietropaolo and Crusio 2012). The inability to locate oneself in an environment or recognize landmarks, also known as spatial disorientation or spatial unawareness, can be assessed in VR settings (Kober et al. 2013). Similar to CPS, effective collaboration in a VR setting requires users to establish a shared perspective either physically or verbally (Pouliquen-Lardy et al. 2016). However, many natural physical motions that are normally used to indicate directions or objects (e.g., pointing) are not easily transferred through hand controllers into a virtual environment, resulting in a loss of information (Giusti et al. 2012).

Spatial dialogue is information communicated by collaborators to establish a common mental representation of the environment (Pouliquen-Lardy et al. 2016). Spatial dialogue can also communicate explicit spatial information such as spatial presence, spatial unawareness, spatial orientation, navigation, and viewpoint information. Zaman et al. (2015) found that participants performing a collaborative VR task communicated better when they were able to establish a shared perspective with their partner. The study also showed that uttered phrases were typically categorized as being either egocentric (from the viewpoint of the speaker) or exocentric (from a viewpoint other than that of the speaker). Pouliquen-Lardy et al. (2016) further divided these classifiers into five codes: “neutral” (independent of viewpoint), “ego-centered” (egocentric), “addressee-centered” (from the viewpoint of the listener), “object-centered” (from the viewpoint of a reference object), and “other-centered” (from the viewpoint of more than one reference). In their study, participants were split into groups of two, each with one “manipulator” and one “guide” in order to observe remote collaboration in VR. Manipulators used significantly more ego-centered language while guides used primarily addressee-centered language that required their partners to make more mental rotations. Utilizing different viewpoints requires different levels of cognitive effort (Pouliquen-Lardy et al. 2016; Schöber 1996). Giving and receiving spatial information that is not ego-centered takes more time and requires higher mental workload, especially in the case of addressee-centered representation (Pouliquen-Lardy et al. 2016; Schöber 1996). Speakers who consider the information and speak from the perspective of the listener during communication minimize the overall mental workload for listeners (Duran et al. 2011; Pouliquen-Lardy et al. 2016; Schöber 1996).

Through this study, we sought to understand how individuals collaborating in a cross-platform (VR to tablet) collaborative game communicate spatial information by examining their dialogue during collaboration. This is a novel study that explores spatial dialogues in a VR collaborative educational cell biology game that is developed so players with different roles must collaborate in order to complete the game. This study shows that a collaborative VR game might be an effective way to understand individuals’ spatial information processing. The study aims to examine how players discuss the cell as a spatial domain. Moreover, this study aims to fill the gap in the literature by exploring how prior content knowledge impacts players’ dialogue during gameplay. Specifically, the present study poses the following research questions:

1. How do individuals communicate spatial information during a role playing cross-platform collaborative game?
2. How does an individuals’ role and prior knowledge in biology impact their dialogue during gameplay?

Method

Participants

The participants were 8 pairs of individuals (16 students total) while playing the game. The audience for the game is high school students, however, in this initial study researchers sampled slightly below (middle school students) and slightly above (recent high school graduates) the intended audience. This divergent sampling strategy is well suited to an exploratory study (Bickman and Rog
as it shows a wide range of responses to the game. The participants included four middle school (MS) students, four high school (HS) students, and eight students in their first semester of a Biotechnology Workforce Program (BWP). The BWP program is a post-high school workforce development program that prepares students to work in biotechnology laboratories. The middle school students had taken 1 year of life science; the high school students had taken life science in middle school and high school biology, and the BWP participants had taken high school biology and were focused on biology in their program. The focus of the game was understanding the cellular environment. Background knowledge of participants was assessed by having students draw a cell and describe their drawing, and asking them about their past and current biology courses. In our research, we have collected hundreds of drawings of cells, and have found that these drawings and interviews are the best way to determine a holistic view of individuals’ level of understanding and recall of cells (Wang et al., 2019). After these analyses, researchers included 8 BWP students with high prior knowledge and 8 middle/high school students with low prior knowledge in this study. Students were paired according to their prior knowledge and their relationship level. Table 1 and Table 2 contain participants’ gender, relationship level, VR experience, prior knowledge, and other demographic information about participants.

The Game Cellverse

Goal of the Game and Background Cellverse is a collaborative VR game designed to teach cell biology, as shown in Fig. 1. The goal of the game is to learn about cells by looking for clues in the cell in order to diagnose the type of cystic fibrosis in the cell. Cystic fibrosis is a genetic disease where individuals have a malformed protein called cystic fibrosis transmembrane conductance regulator (CFTR). Malformed CFTR prohibits the exchange of ions across the cell membrane in special types of lung cells called “ionocytes” (Montoro et al. 2018); this faulty ion exchange prevents tiny hairs called cilia from moving back and forth to sweep mucus out of the lungs. This same CFTR malfunction can be caused by a few different genetic mutations, which manifest in the CFTR proteins either as truncated proteins, misfolded proteins, too few proteins on the membrane, or no proteins at all (CF Foundation 2019). During the game, players review background information to figure out what types of clues they should search for in the environment, and then search for those clues collaboratively to reach the game goal of diagnosing the cell.

Roles and Resources Cellverse is a collaborative game purposefully designed for two players. The roles of Navigator and Explorer are linked to the mode of technology used in the game (Wang et al., 2019). The Explorer views the virtual cell through a VR HMD. VR is well suited for providing a detailed, interactive view of the cell, and allowing the player to develop spatial awareness for the cell. The Explorer view includes limited text, in the form of “just in time” information about the organelle being selected as shown in Fig. 2. The Navigator views the virtual cell through a tablet, which allows the player to have a holistic view of the cell and to access text and image based reference material about CFTR, about organelles, and about possible therapies to address the patient’s CF type, as shown in Fig. 3. The only way to achieve the goal of diagnosing the cell is for the pair to share information with each other so they can understand what clues to look for (Navigator shares with Explorer) and to find them in the virtual cell (Explorer shares with the Navigator), verify that the clues match the CF type (both players jointly), and select an appropriate therapy (both players jointly). Cellverse requires players to learn or discover the location of organelles, their connections and navigation and orientation to understand them all during gameplay. Distributing information across two platforms,
VR and tablet, requires players to understand how their information is similar to or different from different perspectives of the environment and communicate spatial information.

There are also two scales in the game: nano view and macro view (see Figs. 4 and 5). Cellverse allows students to shift between these two scales, enabling users to compare spatial relationships of the objects.

**Study Procedure**

Players were informed that the objective of Cellverse was to work together to figure out what is wrong with the cell and that they had 40 min of time to play the game. Participants decided who would take on the role of the Explorer and who would take on the role of the Navigator. The pair discussed who wanted to be in VR and who wanted to use the tablet, and came to the decision without influence from the researcher. They did not switch the roles while playing the game, but at the end of the game they had the opportunity to try different roles and technologies. Instructions on the VR headset, controllers, tablet, and key points of the game were introduced before the game started. After they were set up with their respective technology (headset and controllers or tablet), each player learned basic features associated with their view through an individual tutorial. The tutorial introduces the players to the capabilities they have in the game. The player in virtual reality begins in a vesicle, which is a small fluid-filled sac or vacuole. The vesicle is a less dense environment containing only some RNA and fluid, this allows the player to focus on learning how to use the hand controllers and move around the cell. The player on the tablet does a tutorial that introduces the player to the menus available to them: information about cystic fibrosis, information about organelles, and the ability to pinch to zoom out and the ability to rotate the cell. After completing the tutorials, the players started the game. They played the game side by side and were able to talk with one another throughout the experience (see Fig. 1). A member of the research team provided technical support and answered technical questions during gameplay, but players were not given extra information about the game content. The teams played until they either finished the game or after 40 min of gameplay. During gameplay, players were recorded with a video camera and lapel microphones. Dialogues between team members were transcribed and analyzed.

**Data Analysis**

The unit of analysis for this study was the approximately 40 min of dialogue between the two players. Conversations (dialogue between the two players) were transcribed by the research team while watching videos. Each researcher watched two videos and transcribed, then checked two transcripts by watching the videos and listening to the conversations that were transcribed by the other researchers. Researchers read the transcripts and highlighted the segments related to spatial context. Codes were developed from the literature (emic) and from a review of the patterns of dialogue in the transcriptions (emic) (Miles and Huberman 1994). The

![Fig. 1 A pair of players during gameplay, showing the Navigator viewing the cell through a tablet, and the Explorer with a VR HMD and hand controllers](image-url)
category of navigation direction codes (emic) were based in part on those from the study conducted by Pouliquen-Lardy et al. (2016). These codes were related to giving navigation direction: neutral, ego-centered, addressee-centered, and object-centered. The emic codes were developed from reviewing the data in light of the research questions. Codes included navigation question, spatial reference, spatial unawareness, spatial orientation, and reference to a specific viewpoint. These codes were compiled into a codebook that included information about the codes, their description, and example utterances that fit in these codes. Researchers followed the following steps to create the codebook (Hruschka et al. 2004):

1) Coders 1 and 2 developed a codebook together based on initial review of the transcriptions and the literature.
2) Coders 1 and 2 coded the same transcript from one gameplay session separately.
3) Coder 1 and coder 2 discussed and came to a consensus on problematic codes. For example, coder 2 used “spatial orientation” code at a time when a player gets closer to an organelle or an object in the cell; after discussion, coders also decided to code “spatial orientation” at a time when players turn around and look organelles from different perspectives. After coders came to a consensus, they modified the codebook.
4) Coders 1 and 2 coded the remaining transcripts separately (7 conversations).

5) A random subset of transcripts were chosen (4 conversations out of 7 conversations) and checked in terms of consistency between the coders. This subset showed an inter-coder reliability percentage of 92%.

The codebook that describes the codes and gives examples of utterances is below (also see Table 3):

a) Navigation directions: utterances that are about giving directions while navigating in VR:
   i) Neutral: Directions are independent of viewpoint (e.g., “Go to the ribosome”).
   ii) Ego-centered: Directions are from the speaker’s viewpoint (e.g., “The ER is in front of me”).
   iii) Addressee-centered: Directions are from the listener’s viewpoint (e.g., “Nucleus is on your right”).
   iv) Object-centered: Directions are from the viewpoint of a reference object (e.g., “Look at the organelle that is in front of the ER”).

b) Navigation question: Utterances that are questions related to navigation (for example: “Where am I going now?”, “Where should I go?”)

c) Spatial reference: Utterances that are about referencing absolute location or relative location in the environment (when they find out where they are, when they identify organelles present in the environment, when they are aware of objects surrounding them, etc., e.g., “Now I see, Golgi Body, here it is”).

d) Spatial unawareness: Utterances that are about the inability to locate oneself in the environment. (e.g., “Where am I? What is surrounding me?”).

e) Spatial orientation: Utterances that are about the ability to determine one’s position and where they are heading in the environment (getting closer to the organelles, turning around and looking from different perspectives, etc. e.g., “I am zooming in, I am getting closer”).

f) Viewpoint: Utterances that are about either the Navigator’s and Explorer’s point of view, indicating that one or the other player recognizes that they have different views of the cell or that they can see different things (e.g., “Do you see me? Do you see where am I pointing?”).

Transcripts of conversations were imported into spreadsheets with one row for each utterance per player, and a separate column for each code. Each utterance was coded for every code either by marking a “1” in the code column if the code was present, or a “0” in the code column if the code was absent. This data format allowed us to analyze the frequencies and the relationships between codes in the conversation and also to model the conversation as a network of interrelated ideas through Epistemic Network Analysis (ENA) (Shaffer 2017; Shaffer et al. 2016). “A network consists of nodes (objects or ideas) and relationships between nodes (ties or connections), in ENA, the nodes are represented as the coded data (i.e., stages and themes) and the relationship between nodes indicate when two codes occur within the same segmentation of time” (Fisher et al. 2016, p. 1). Each code column is represented by a node on the network graph. During ENA, the researcher establishes a moving frame, consisting of a set of rows of utterances, and tracks how ideas are connected during that segment of the conversation. This enables the analysis to consider not only the topics of conversation but also the number of times the topics are discussed. We defined conversations as all lines of data aligned with a single value of speaker type (Navigator vs. Explorer) subsetted by pairs (collaborators). For example, one conversation consisted of all the lines aligned with one Navigator and Explorer pair.

We first conducted a chi-square test for independence to investigate the association between spatial codes and background knowledge (middle/high vs. BWP students) or role of the player (Navigator or Explorer). This test is used to
Table 3 Codes and example utterances

| Code               | Example utterances                                                                 |
|--------------------|-----------------------------------------------------------------------------------|
| Spatial reference  | So all those red things are ribosomes. The blue things are transfer RNA. Let us see what else. |
| Spatial unawareness| I have no idea where I am.                                                          |
| Viewpoint          | Oh look! I can see you there!                                                       |
| Navigation         | Neutral: I will go to the ribosomes                                                |
| Directions         | Ego-centered: I am in front of the ER.                                              |
|                    | Address-se-centered: It is just in front of you.                                   |
| Spatial orientation| I am getting closer! I am in Nano.                                                  |
| Navigation         | Where should I go?                                                                 |
| questions          | Maybe I should go there?                                                            |

compare observed frequencies in each categorical variable (background knowledge and role of the player) (Pallant 2007). Moreover, we compared spatial utterances by applying ENA to our data. We used ENA to further investigate the relationship between the players’ background knowledge and role and the way they discussed spatial ideas during the conversation. Our ENA model included the following spatial codes: Spatial Reference, Spatial Unawareness, Spatial Orientation, Neutral, Ego-centered, Address-se-centered, and Viewpoint. We removed the codes with insufficient frequency (less than 5) for ENA analysis (object-centered, other-centered), because their relationships were not depicted graphically.

Results

Chi-Square Test for Independence Results

A total of 1418 utterances were produced by Navigators (707 utterances) and Explorers (711 utterances). There was no significant difference in the number of utterances between roles. Of those 1418 utterances, 1030 were coded using the codebook. Spatial reference was the code with the highest frequency among other codes (58%) (see Fig. 6), because spatial reference was used as a top-level code to identify any time when players talked about the game environment, thus engaging in spatial dialogue. The second most stated utterance type was viewpoint (13%), marking when one of the players acknowledged they had a different view of the information than their partner. The code most often attributed to directions was “Neutral,” with a frequency of 8%. Utterances coded as “neutral” were about giving directions (neutral, ego-centered, address-se-centered, and object-centered) while navigating in VR. Collaborators almost never used object-centered directions. Navigators produced significantly more neutral ($\chi^2 (1, 1030) = 12.91, p < .05$) and address-se-centered ($\chi^2 (1, 1030) = 9.77, p < .05$) utterances. However, Explorers produced significantly more utterances about spatial unawareness ($\chi^2 (1, 1030) = 4.58, p < .05$) and ego-centered utterances ($\chi^2 (1, 1030) = 13.24, p < .05$). In other codes, there was not any significant difference according to roles.

More total utterances were produced by BWP students (734 utterances) than MS/HS students (690 utterances). MS/HS students produced significantly more spatial unawareness utterances ($\chi^2 (1, 1030) = 10.32, p < .05$). On the other hand, BWP students produced significantly more address-se-centered ($\chi^2 (1, 1030) = 3.83, p < .05$) and viewpoint ($\chi^2 (1, 1030) = 41.4, p < .05$) utterances than MS/HS students (see Fig. 7).

Epistemic Network Analysis Results

Navigator vs. Explorer

ENA network graphs show how frequently the code occurs by the size of the node and by the co-occurrences of codes in the reference frame through linear connections. Spatial reference (SR) is at the center of both graphs because all of these codes include spatial reference. The axes indicate the highest percent of variance explained by a dimension. In the present study, Neutral and spatial unawareness codes are in the positive $Y$ direction, so conversations that are located more positively contain more neutral references or more discussions of feeling lost. Students who have higher frequency of co-occurrences of the codes “neutral” and “spatial unawareness” would have strong connections located in the positive $Y$ direction. ENA explains 34.5% of the variance in coding co-occurrences along the $X$-axis and 20.7% of the variance on the $Y$-axis. Along the $X$-axis, a Mann-Whitney test showed that the overall network of the Explorer ($Mdn = -0.99, N = 8$) was statistically significantly different from Navigator ($Mdn = 1.06, N = 8 U = 4.00, p = .05, r = .88$). Individual network diagrams explain what codes are causing this shift; the individual diagrams are shown in Fig. 8.

The network diagram for the Navigator shows stronger ties between the codes SR-Viewpoint and SR-Neutral and a weak connection between SR-Spatial unawareness. The network diagram for Explorer has connections that show a presence of ties between the codes SR-Viewpoint and SR-Spatial unawareness, and a weak connection between SR-Neutral and SR-Ego-centered. Viewpoint is a common and frequently used code both for Explorers and Navigators, because they tried to figure out whether they were seeing the same things, or if they could see each other in the environment during the conversation. A sample excerpt from a conversation is included below, with codes in parentheses.
NAV1: OK it seems like it works now. OK, so you are still in the nucleus, but I’m going to try and light up the Golgi body. Do you see that? (Viewpoint and SR)
EXP1: Yup.
NAV1: Alright so now I’m going to light up the ER and ribosomes. Should be purple.
EXP1: Yep.
NAV1: Cool. Ummmm. OK I see now. So now I’m going to have you [try]… (SR)

According to chi-square analysis, Explorers produced more ego-centered utterances, indicating a preference for their own perspective. The ENA network diagram also showed a connection between ego-centered and spatial reference in utterances by Explorers but not Navigators. In the dialogue below, the Explorer is telling his position from his own perspective using an ego-centered utterance, and at the end of the conversation he describes seeing proteins “floating around” him, which indicates spatial referencing.

EXP2: No, I think I’m inside the ribosome. (Ego-centered)
NAV2: You’re in the ribosome? Can you see proteins?
EXP2: They’re the blue things floating around.

The difference in diagrams between the Navigator and Explorer can be explained by SR-ego-centered, “SR-neutral,” and “SR-spatial unawareness” utterances. The network graphs demonstrate that Explorers have stronger connections between spatial reference and spatial unawareness than Navigators. This finding shows that students in the VR environment talked about spatial unawareness and spatial reference together. As shown in the dialogue below, while attempting to identify objects in VR, the Explorer initially references spatial information in the environment even while having the feeling of being lost.

The difference in diagrams between the Navigator and Explorer can be explained by SR-ego-centered, “SR-neutral,” and “SR-spatial unawareness” utterances. The network graphs demonstrate that Explorers have stronger connections between spatial reference and spatial unawareness than Navigators. This finding shows that students in the VR environment talked about spatial unawareness and spatial reference together. As shown in the dialogue below, while attempting to identify objects in VR, the Explorer initially references spatial information in the environment even while having the feeling of being lost.
NAV2: Yeah.
EXP2: Yup, I can see them. (Spatial reference)

Network diagrams showed that Navigators had stronger connections between a neutral description of the environment and spatial reference than Explorers. Navigators preferred not to take a perspective while directing Explorers.

NAV3: Go to the protein. (Neutral)
EXP3: So what’s this? It looks like a giant clump of it. Maybe it has to do with something. [Explorer is looking up and focusing on giant clump]
NAV3: So it does not it’s supposed to be, like, kind of neatly folded. It’s just...
EXP3: like all over the place. And look at these - look at them - [they] are like flying around. Some of them aren’t. And so it looks like these are all in clumps. (Spatial reference)
NAV3: Yeah.

MS/HS Students vs. BWP Students

To explore how the player’s background may influence their spatial dialogue, we also created a network graph for each group of students, one for MS/HS students and one for BWP students. When MS/HS students and BWP students were compared, ENA explained 25.7% of the total variance in coding co-occurrences along the X-axis and 32.3% of the total variance on the Y-axis. Along the X-axis, a Mann-Whitney test showed that the overall network for BWP ($\text{MdN} = 2.40, N = 4$) was statistically significantly different from the overall network for MS/HS ($\text{MdN} = −2.24, N = 4 U = 15.00, p = 05, r = −0.87$), suggesting that players’ biology background had an effect on dialogue. The individual network diagrams are shown in Fig. 9.

The network diagram for the BWP students shows a presence of ties to the codes for SR-Neutral, SR-Viewpoint, SR-Spatial orientation and weak connections between SR-Addressee-centered and SR-Ego-centered. The diagram for middle/high school shows a presence of ties to the codes for SR-Neutral, SR-Spatial unawareness, SR-Viewpoint (strong), and also connections to the spatial orientation and Ego-centered(weak).

When these diagrams were compared, BWP students had stronger connections between SR-Spatial orientation and SR-Addressee-centered direction than middle/high school students had. In other words, BWP students used different types of points of views (addressee-centered, ego-centered, and neutral) while discussing objects in the environment with their partners. In the example dialogue of BWP students, the Navigator is trying to help the Explorer by taking his perspective (coded as addressee-centered). This was a pattern among the BWP students; they tried to understand the Explorer’s viewpoint in order to support their partners while looking for organelles or clues in the environment.

BWP Students:
NAV1: OK so looks like you are in……you are in the nucleus again. So that’s kind of where I wanted you to be. So I want you to look at that. I’m going to turn some of the proteins on. Look at the Golgi apparatus in front
of you. Look at the proteins and Golgi body. (Addressee-centered)
EXP1: Yup. So those are the Golgis, those are the rough ER. I see. (Spatial Reference)
NAV1: OK, do you see [anything] looking around that seems off to you?
EXP1: Ummmmmm... I see well there’s a bunch of like little ribosomes in the.... (Spatial reference)
NAV1: Purple or pink [referring to organelle colors]
EXP1: In the purple.
NAV1: OK I’ll turn the Golgi body off so you are not confused and so you can see better.

On the other hand, MS/HS school students had stronger connections between spatial reference and spatial unawareness than BWP students. MS and HS students expressed unawareness of their location, rather than using different points of view to support navigation processes of each other. As seen in the example below, both Navigator and Explorer are having difficulty understanding the virtual environment, thus showing spatial unawareness.

MS/HS Students:
EXP4: What, it’s disappearing. (Spatial unawareness)
NAV4: What are you trying to find? [NAV looks at EXP’s screen]
EXP4: I do not know. (Spatial unawareness)
NAV4: Okay
EXP4: Okay, I’m just like floating here. I do not know what I’m supposed to do. Oh, woah. Well. I clicked something. (Spatial unawareness)

Discussion

Participants who played Cellverse faced a few challenges. They had to understand the dynamic and novel environment of a three dimensional cell. They had to decipher an ill-defined problem space and had to determine the resources available to them in order to determine what clues were associated with the type of cystic fibrosis (the role of the Navigator) and seek the clues in the cellular environment (the role of the Explorer) and communicate this information in order to diagnose the cell. Each of these challenges required the players to develop a way of communicating about a novel environment. In observing how players approached these challenges, we found that different roles and different levels of biology knowledge influenced their conversations during the game.

Players discussed the environment from a number of perspectives. When dialogues between pairs were analyzed, we found utterances that could be categorized as spatial reference, spatial unawareness, spatial orientation, navigation directions (neutral, ego-centered, addressee-centered, and object-centered), viewpoint, and navigation questions. While the codes navigation direction (neutral, ego-centered, addressee-centered, and object-centered) already existed in the literature, spatial reference, spatial awareness, spatial unawareness, and spatial orientation were novel codes that were found out in this study. Viewpoint and spatial reference were the most frequently used utterances both by Explorers and Navigators. While playing the game, both Navigators and Explorers tried to figured out whether they were seeing the same things, or if they could see each other in VR, which is consistent with the results of Giusti et al. (2012). We did not provide players with any clues before the game began, as a result, collaborators spent the first part of their gameplay figuring out their viewpoints. The different information provided to the players created a necessity for collaboration, resulting in positive interdependence (Johnson et al. 1991). After understanding their partner’s viewpoint, taking different perspectives while referencing objects in a space is important for navigation and wayfinding (Miniaci and De Leonibus 2018). We found that the reference perspective used by Explorers and Navigators did vary significantly. Explorers used more egocentered references, referring to their own perspective, which is not surprising as the HMD put the Explorer at the center of the action. On the other hand, Navigators used more addressee-centered and neutral references to help Explorers navigate and find clues in the game, which reflects the more global viewpoint of the cell provided on the tablet. Pouliquen-Lardy et al.’s (2016) study also found that the player giving directions uses more addressee-centered and neutral references, while the participant that moving the virtual objects according to the instructions used more ego-centered references. Although the roles were similar in these two studies, the technology and perspectives were different than in this study. In our study, the Navigator views the virtual cell through a tablet, which allows the player to have a holistic view of the cell, and the Explorer views the virtual cell through a VR HMD. VR is well suited for providing a detailed, interactive view of the cell from an ego-centered perspective. However in Pouliquen-Lardy et al.’s (Pouliquen-Lardy et al. 2016) study, both the guide (Navigator) and manipulator (Explorer) were in VR and they had the same perspectives with different reference materials. Regardless of these differences, task distribution results in similar findings for both our study and Pouliquen-Lardy et al.’s (Pouliquen-Lardy et al. 2016) study that Navigators/guides are better able to describe the environment from the perspective of the Explorers/manipulators, but Explorers/manipulators are not as able to speak in terms of the Navigator/guide perspective. This study supports the idea that role or task distribution is an important factor that affects the processing of spatial information. Giving spatial information that is addressee-centered requires taking the point of view of the other player, which
requires higher mental workload and takes more time than speaking from the player’s own perspective (Michelon and Zacks 2006; Pouliquen-Lardy et al. 2016; Schober 1996). Zaman et al. (2015) also explored how pairs used spatial references in a natural and embodied way in a shared-perspective VR. They found that directions were given from the perspectives of the speaker and listener significantly more often than directions from their partner’s perspectives (object-centered and other-centered). In this study, collaborators almost never used object-centered or other-centered directions. This makes sense, because talking from the viewpoint of others requires more perspective changes and leads to a higher mental workload for both members of the team (Pouliquen-Lardy et al. 2016).

The study found evidence that players felt spatial presence during the game in an authentic way: by reviewing dialogues during gameplay. Prior research on the potential benefits of VR environments have focused on how VR experiences can create a sense of spatial presence in users. A number of studies have explored spatial presence using questionnaires and self-report surveys (Cheng and Tsai 2019; Seibert and Shafer 2017). Moreover, interviews (Garau et al. 2008) and think-aloud protocols (Turner et al. 2003) have also been used for measuring spatial presence, and researchers are still thinking about different ways of measuring spatial presence (Laarni et al. 2015). Moreover, interviews (Garau et al. 2008) and think-aloud protocols (Turner et al. 2003) have also been used for measuring spatial presence, and researchers are still thinking about different ways of measuring spatial presence (Laarni et al. 2015). Pouliquen-Lardy et al. (2016) asked participants about their feeling of being present in the environment to find out their levels of spatial presence. The study did not find any difference between collaborators. Instead of using scales/questionnaires or interviews, we analyzed utterances that players produced that give clues about their feeling of spatial presence. This strategy is more authentic than questionnaires and self-report surveys because we are examining participants’ actions in the moment, rather than asking for their perception of their actions after they have finished the game. Examining conversations during gameplay revealed a number of spatial references, providing evidence that players were developing a sense of spatial presence within the environment. Additionally, we found that the role each player took influenced the way they made spatial references. Players in VR used more ego-centered utterances, suggesting that the VR technology became an integrated part of the speaker’s viewpoint (Riva and Waterworth 2014) and supporting past findings that feelings of presence can be enhanced with head-mounted displays (Bruder et al. 2009). According to Wirth et al. (2007), spatial presence occurs when the player accepts the media as his/her primary ego reference frame, because their sensation of being located in the environment is connected to the environment. Media, user characteristics, and activities in the environment are all associated with the feeling of being physically situated within the environment (Laarni et al. 2015). In the present study, the HMD that isolates the player from the real environment and the ability to interact with the objects via handheld controllers are both media-specific factors that might affect players’ feeling of “being there”. Interacting with organelles and searching for clues by exploring the virtual environment might also result in feelings of spatial presence. As the interactive aspects of technology improve and the barrier between the technology and reality becomes more seamless, we should see larger gains in spatial presence.

**Fig. 9** Individual network diagrams for BWP and middle/high school students.
presence resulting from experiences in VR (Regenbrecht and Schubert 2002).

We also found that background knowledge influenced how players described the environment. When BWP and MS/HS students were compared, BWP students produced more addressee-centered and viewpoint utterances. On the other hand, MS/HS students produced more utterances related to spatial unawareness than BWP students did. BWP students with higher prior knowledge about cell biology, the main topic of the game, tried to minimize the overall effort of their collaborators. Familiarity with the game topic might give players a conceptual framework of what to expect in the environment, thereby reducing the effort required to give spatial information from different points of view. In other words, participants with more pre-knowledge do not have to focus as hard on the game and can take on the additional mental workload of performing the mental rotations necessary for non-ego-centered dialogue.

Previous studies found that participants’ learning and levels of self-efficacy increased when they knew the names and main concepts before performing activities in VR (Meyer et al. 2019). In the present study, BWP students might have had the opportunity to focus on navigating and finding clues to complete the game instead of dealing with new concepts in VR than the MS/HS students. Additionally, MS/HS students produced more utterances related to spatial unawareness than BWP students did. In this respect, familiarity to the game topic might impact players’ feeling of being there and the type of spatial information they shared.

Conclusion

Cellverse is a collaborative problem-solving game that gives the opportunity to players to establish a shared understanding of the spatial representation of a virtual cell as a three-dimensional environment through their discussion during the game. Players made frequent spatial references during the game while trying to find organelles, exploring the location of organelles, and giving directions to each other in order to navigate in the environment. We found that different roles and different levels of biology knowledge influenced how they used spatial references in their conversations during the game.

This study makes a twofold contribution to the body of literature. First, this study shows that a collaborative VR game might be an authentic way to understand individuals’ spatial information processing. The dialogues we observed aligned with prior research on spatial dialogue, thus we were able to examine how players discussed the cell as a spatial domain. Analyzing dialogue between collaborators with epistemic network analysis gave clues about how pairs shared spatial information. This method does not require recalling experiences, one of the disadvantages of interviews, and avoids distracting players with think-alouds during gameplay. Observing interactions between partners as they experienced the game provides an authentic view into their thinking and actions “in the moment.” This method of studying spatial understanding through examining dialogue during collaborative problem-solving could be useful for future research.

Second, findings from this study can guide designers as they think about how to scaffold collaborative cross-platform experiences. Learning how to communicate with others who have different levels of knowledge and different roles is essential to successful collaborative problem-solving. Navigators with low prior knowledge might be supported with additional materials or aids to help them guide the Explorers in VR. For instance, an initial spatial representation of the game environment might be given to Navigators to make them use more addressee-centered perspectives for Explorers. This representation might result in better performance in navigation, more efficient searches for organelles or clues that they need during problem-solving in VR, and lower mental workload for Explorers. Explorers could be supported with a map during the game, in order to see where they are. This spatial representation of a cell might be crucial for the players with low prior knowledge who do not have adequate ideas about where the organelles are or the general structure of a cell.

Our results demonstrate that the player’s role (task and technology distribution) and familiarity with the game topic all have an impact on the context of spatial information shared during gameplay. In the future, different factors can be studied in order to develop a better understanding of spatial processing between collaborators. Gameplay experience, familiarity with VR technology, and spatial abilities including mental rotation ability, spatial reasoning, or object location memory might impact learners’ performances. Moreover, Cellverse is a cell biology game that you need to learn or discover the location of organelles, their connections and requires navigation and orientation to understand them all during gameplay. In the future, different games with different contexts might be used that require both spatial information and content knowledge processing. These games will help us further explore and understand the important domain of collaborative problem-solving.

Funding This material is based upon work supported by Oculus Education.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethics Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee. Research on human subjects has been approved by the Institutional Review Board of Massachusetts Institute of Technology.
Consent to Participate  Informed consent was obtained from all individual participants included in the study.

References

Abe, T., Raison, N., Shinohara, N., Khan, M.S., Ahmed, K., & Dasgupta, P. (2018). The effect of visual-spatial ability on the learning of robot-assisted surgical skills. Journal of Surgical Education, 75(2), 458–464. https://doi.org/10.1016/j.jsure.2017.08.017.

Bickman, L., & Rog, D. J. (2008). The SAGE handbook of applied social research methods. Sage publications.

Bruder, G., Steinicke, F., Rothaus, K., & Hinrichs, K. (2009). Enhancing presence in head-mounted display environments by visual body feedback using head-mounted cameras. In 2009 International Conference on CyberWorlds (pp. 43–50). IEEE.

CF Foundation. (retrieved, July 2019) “Basics of the CFTR protein.” CF Foundation. www.cf.org/research/research-into-the-disease/Restore-CFTR-Function/Basics-of-the-CFTR-Protein/.

Cheng, K.-H., & Tsai, C.-C. (2019). A case study of immersive virtual field trips in an elementary classroom: Students’ learning experience and teacher-student interaction behaviors. Computers & Education, 140, 103600. https://doi.org/10.1016/j.compedu.2019.103600.

Coxon, M., Kelly, N., & Page, S. (2016). Individual differences in virtual reality: Are spatial presence and spatial ability linked? Virtual Reality, 20(4), 203–212. https://doi.org/10.1007/s10055-016-0292-x.

Duncan, J., & West, R. E. (2018). Conceptualizing group flow: A framework. Educational Research and Reviews, 13(1), 1–11.

Durán, J.F. (2017). Virtual reality in education: State of the art and perspectives. In The International Scientific Conference eLearning and Software for Education (Vol. 1, p. 133). “Carol I” National Defence University.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., … & Hinrichs, K. (2009). Enhancing presence in head-mounted display environments by visual body feedback using head-mounted cameras. In 2009 International Conference on CyberWorlds (pp. 43–50). IEEE.

Financial support: This research was supported by a grant from the National Science Foundation (DRL-1403630) to Dr. John D. Jenkins.

Fisher, K. Q., Hirshfield, L., Siebert-Evenstone, A., Arastoopour, G., & Koretsky, M. (2016). Network analysis of interactions between students and an instructor during design meetings. In Proceedings of the American society for engineering education. ASEE.

Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., … & Souilé, H. (2017). Collaborative problem solving: Considerations for the national assessment of educational progress. National Center for Educational Statistics.

Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. In The International Scientific Conference eLearning and Software for Education (Vol. 1, p. 133). “Carol I” National Defence University.

Garau, M., Friedman, D., Ritter Widenfeld, H., Antley, A., Brogni, A., & Slater, M. (2008). Temporal and spatial variations in presence: Qualitative analysis of interviews from an experiment on breaks in presence. Presence, 17(3), 293–309.

Giusti, L., Xerxes, K., Schladow, A., Wallen, N., Zane, F., & Casalegno, F. (2012). Workspace configurations: Setting the stage for remote collaboration on physical tasks. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (pp. 351–360). ACM.

Hahn, J. F. (2017). Virtual reality library environments. American Library Association.

Hite, R. L., Jones, M. G., Childers, G. M., Ennes, M., Chesnutt, K., Pereyra, M., & Cayton, E. (2019). Investigating potential relationships between adolescents’ cognitive development and perceptions of presence in 3D, haptic-enabled, virtual reality science instruction. Journal of Science Education and Technology, 28(3), 265-284. https://doi.org/10.1007/s10956-018-9764-y

Hruschka, D. J., Schwartz, D., St. John, D. C., Picone-Decaro, E., Jenkins, R. A., & Carey, J. W. (2004). Reliability in coding opened-ended data: Lessons learned from HIV behavioral research. Field Methods, 16(3), 307–331.

Johnson, D. W., Johnson, R. T., Ortiz, A. E., & Stamey, M. (1991). The impact of positive goal and resource interdependence on achievement, interaction, and attitudes. J Gen Psychol, 118(4), 341–347.

Johnson, L., Adams Becker, S., Cummings, M., Estrada, V., Freeman, A., & Hall, C. (2016). NMC horizon report: 2016 higher Education Edition. Austin: The New Media Consortium.

Kober, S. E., Wood, G., Hofer, D., Kreuzig, W., Kiefer, M., & Neuper, C. (2013). Virtual reality in neurologic rehabilitation of spatial disorientation. Journal of NeuroEngineering and Rehabilitation, 10(1), 17.

Laarni, J., Ravaja, N., Saari, T., Böcking, S., Hartmann, T., & Schramm, H. (2015). Ways to measure spatial presence: Review and future directions. In Immersed in Media (pp. 139–185). Cham: Springer.

León, I., Tascón, L., Orléats-Pareja, J.I., & Cimadevilla, J.M. (2018). Virtual reality assessment of walking and non-walking space in men and women with virtual reality-based tasks. PLoS One, 13(10) https://doi.org/10.1371/journal.pone.0204995.

Liszio, S., & Masuch, M. (2016). Designing shared virtual reality gaming experiences in local multi-platform games. In International Conference on Entertainment Computing (pp. 235–240). Cham: Springer.

Meyer, O. A., Ondahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. Computers & Education, 140, 103603.

Michelon, P., & Zacks, J. M. (2006). Two kinds of visual perspective taking. Percept Psychophys, 68(2), 327–337.

Miles, M. A., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook (pp. 50–72). Thousand Oaks: SAGE Publications.

Miniaci, M. C., & De Leonibus, E. (2018). Missing the egocentric spatial reference: A blank on the map. F1000Research, 7.

Montoro, D. T., Haber, A. L., Biton, M., Vinarsky, V., Lin, B., Birket, S. E., … et al. (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature, 560(7718), 319–324.

Pallant, J. (2007). A step by step guide to data analysis using SPSS for windows. SPSS Survival manual, McGraw-Hill Education (UK).

Pietropaolo, S., & Crussi, W. E. (2012). Learning spatial orientation. In NMC horizon report: 2016 higher Education Edition (pp. 210–210). Austin: The New Media Consortium.

Pouliquen-Lardy, L., Milleville-Pennel, I., Guillaume, F., & Mars, F. (2016). Remote collaboration in virtual reality: Asymmetrical effects of task distribution on spatial processing and mental workload. Virtual Reality, 20(4), 213–220.

Regenbrecht, H., & Schubert, T. E. (2002). Real and illusory interactions enhance presence in virtual environments. Presence: Teleoperators and Virtual Environments, 11(4), 425–434.

Riva, G., & Watervorth, J. A. (2014). Being present in a virtual world (pp. 205–221). The Oxford Handbook of Virtuality.

Sawyer, K. (2017). Group genius: The creative power of collaboration. Basic books.

Schobre, M. F. (1996). Addressee-and object-centered frames of reference in spatial descriptions. In American Association for Artificial Intelligence, Working Notes of the 1996 AAAI Spring Symposium on Cognitive and Computational Models of Spatial Representation, (47), 92–100.

Seibert, J., & Shaffer, D. M. (2017). Control mapping in virtual reality: Effects on spatial presence and controller naturalness. Virtual Reality, 22(1), 79–88. https://doi.org/10.1007/s10055-017-0316-1.

Shaffer, D. M., Carbonara, C. P., & Korpi, M. F. (2019). Factors affecting enjoyment of virtual reality games: A comparison involving
consumer-grade virtual reality technology. *Games for Health Journal*, 8(1), 15–23.

Sherman, W. R., & Craig, A. B. (2018). Understanding virtual reality: Interface, application, and design. Morgan Kaufmann.

Tascón, L., García-Moreno, L. M., & Cimadevilla, J. M. (2017). Almeria spatial memory recognition test (ASMRT): Gender differences emerged in a new passive spatial task. *Neuroscience Letters, 651*, 188–191. https://doi.org/10.1016/j.neulet.2017.05.011.

Tussyadiah, I. P., Wang, D., & Jia, C. H. (2017). Virtual reality and attitudes toward tourism destinations. In *Information and communication technologies in tourism 2017* (pp. 229–239), Springer, Cham.

Shaffer, D. W. (2017). *Quantitative ethnography*. Madison: Cathcart Press.

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45.

Turner, S., Turner, P., Carroll, F., O’Neill S., Benyon, D., McCall, R., et al. (2003). Re-creating the Botanics: Towards a sense of place in virtual environments. Paper presented at the 3rd UK Environmental Psychology Conference, Aberdeen, 23–25 June 2003.

Wang, A., Thompson, M., Roy, D., Pan, K., Perry, J., Tan, P., Eberhart, R., & Klopfer, E. (2019). Iterative user and expert feedback in the design of an educational virtual reality biology game. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2019.1678489

Weber, M. S., & Kim, H. (2015). Virtuality, technology use, and engagement within organizations. *J Appl Commun Res, 43*(4), 385–407.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (2007). A process model of the formation of spatial presence experiences. *Media Psychol, 9*(3), 493–525.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence Teleop Virt, 7*(3), 225–240.

Zaman, C. H., Yakhina, A., & Casalegno, F. (2015). nRoom: An immersive virtual environment for collaborative spatial design. In *Proceedings of the International HCI and UX Conference in Indonesia* (pp. 10-17). ACM.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.