Effect of View Sharing on Spatial Knowledge Acquisition in Remote Collaboration

Tzu-Yang Wang
Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
Nomi, Ishikawa, Japan
tywang@jaist.ac.jp

Ikkaku Kawaguchi
University of Tsukuba
Tsukuba, Ibaraki, Japan
kawaguchi@cs.tsukuba.ac.jp

Hideaki Kuzuoka
Graduate School of Information Science and Technology
The University of Tokyo
Bunkyo-ku, Tokyo, Japan
kuzuoka@cyber.t.u-tokyo.ac.jp

Mai Otsuki
Human Augmentation Research Center
National Institute of Advanced Industrial Science and Technology (AIST)
Kashiwa, Chiba, Japan
mai.otsuki@aist.go.jp

ABSTRACT

Spatial referencing is one of the important tasks in remote collaboration; for example, a person asks other collaborators to move to a specific location or pick up an object in a specific location. To achieve successful spatial referencing, one of the key points is that collaborators should obtain spatial knowledge and shared spatial knowledge about the environment. We selected the view-sharing method, which is one approach to support collaborators in understanding each other's spatial frame-of-reference, and we investigated the effect of view-sharing on spatial knowledge and shared spatial knowledge acquisition in remote collaboration. A maze exploration experiment was conducted. Participants were asked to explore the maze collaboratively with/without view-sharing. Later, to examine the participants' acquired graph knowledge, survey knowledge, and shared survey knowledge, the participants were asked to individually plan routes to move from objects to objects in the maze and draw the maze. The result showed that sharing collaborators' viewpoints improved collaborators' spatial knowledge acquisition.

CCS CONCEPTS

• **Human-centered computing** → *HCI theory, concepts and models*; *Computer supported cooperative work*; **Collaborative interaction**.

KEYWORDS

Spatial knowledge, Shared Spatial Knowledge, View Sharing, Remote Collaboration

ACM Reference Format:
Tzu-Yang Wang, Ikkaku Kawaguchi, Hideaki Kuzuoka, and Mai Otsuki. 2022. Effect of View Sharing on Spatial Knowledge Acquisition in Remote Collaboration. In *Symposium on Spatial User Interaction (SUI '22), December 1–2, 2022, Online, CA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3565970.3567699

1 INTRODUCTION

In many remote collaborations, especially for remote collaboration that involves physical tasks, people often perform **spatial referencing**, which is an action that draws other people's attention to specific locations, to navigate other remote collaborators [18, 24, 25]. For example, an instructor performs spatial referencing to ask a worker to pick up an object in a specific position during remote instruction, or a player performs spatial referencing to ask other teammates to move to a specific position in a remote collaborative game.

To perform successful spatial referencing, it is necessary that the **spatial expressions**, which a speaker uses to achieve spatial referencing, are well-designed and can be understood by others. In order to design appropriate spatial expressions, there are two important points for the speaker. First, the speaker should understand each other's spatial frame-of-reference (a spatial coordinate system used for representing positions of objects) so that he/she can perform better at perspective taking during designing spatial expressions [26], and lead to better spatial communication. Much remote collaboration research has focused on supporting collaborators in understanding each other's spatial frame-of-reference, such as sharing information about each other's viewpoints [13, 17] or providing shared visual landmarks [6, 22]. These studies mainly examined how their manipulation and developed systems improved spatial communication and efficiency in a short-term remote collaboration. However, to the best of our knowledge, none of them has paid attention to whether the participants obtained a better understanding about the environment, which is also known as spatial knowledge.

Spatial knowledge is knowledge about the configuration of an environment and locations of objects/landmarks in that environment. Obtaining spatial knowledge and **shared spatial knowledge**

SUI '22, December 1–2, 2022, Online, CA, USA

ACM ISBN 978-1-4503-9948-7/22/12…$15.00
https://doi.org/10.1145/3565970.3567699

(spatial knowledge that is obtained by all collaborators) is another important point for the speaker to design appropriate spatial expressions, especially for remote collaboration that takes place in a broad environment and for long-term remote collaboration, which same collaborators repeatedly collaborate in a same environment. In a wide environment, such as a wide warehouse, objects or items are placed separately and far away from each other, and people who obtain spatial knowledge about the environment can effectively collaborate with each other, such as efficiently planning, finding and picking up needed items in the warehouse together. For long-term remote collaboration, if collaborators can better and faster acquire spatial knowledge about an unfamiliar environment in the early stage, after that, they can collaborate effectively. Therefore, it is significant to understand how people obtain spatial knowledge in remote collaboration. Additionally, it is important to propose methods to support people in acquiring spatial knowledge of a new environment effectively so that they can adopt their spatial knowledge in later collaboration. In the psychology field, there have been some studies that focused on factors affecting individuals' spatial knowledge acquisition, but there are few studies that focused on how people acquire spatial knowledge under collaborative situations.

To fill such a research gap, we aimed to investigate how sharing knowledge of frame-of-reference supported people in acquiring spatial knowledge and shared spatial knowledge. In the current research, we leveraged the view-sharing method, one of the approaches to create shared knowledge of frame-of-reference, and examined the effect of the view-sharing method on spatial knowledge acquisition.

2 RELATED WORK

2.1 Frame-of-reference in Remote Collaboration

As people grow up, they acquire knowledge about frame-of-reference (FoR) concepts from spatial language. Frame-of-reference concepts are mental frameworks that maintain a set of spatial relations [30]. These coordinate frameworks can be based on any entity or set of entities in the space. Depending on the type of entity, frame-of-reference concepts can be categorized into two types: geocentric frame-of-reference and object-based frame-of-reference. For the geocentric frame-of-reference, the entities are stationary, such as buildings or mountains, and the axis and coordinate systems created by the entities are often stable. For the object-based frame-of-reference, the entities are things that freely move relative to the earth, such as a person or an artificial object. Furthermore, the object-based frame-of-reference includes egocentric frame-of-reference and allocentric frame-of-reference. The egocentric frame-of-reference is the case that the coordinate system is based on the speaker's own body, while the allocentric frame-of-reference is the case that the coordinate system is based on other objects or people.

In collaboration, to perform spatial referencing to other collaborators, people often select the egocentric frame-of-reference [2, 29], i.e., they describe information from their own perspective. However, since the viewpoints were not shared between people, describing information from their own perspectives might cause much misunderstanding between each other. For example, in the case of two people who face each other, "left/right" have opposite meanings, and listeners often misunderstand whether speakers are mentioning "speaker's left" or "listener's left." In a co-locate situation, people are often aware of other collaborators' positions and gaze behaviors. Thus, this helps them effectively understand other collaborators' perspectives and frame-of-reference. However, in remote collaboration situations, due to low co-presence and lack of nonverbal behaviors, it is hard for people to understand others' spatial frame-of-reference, and the spatial expressions are often misunderstood by other collaborators.

In order to improve the understanding of each other's spatial frame-of-reference and achieve a better understanding of spatial expressions, there are two common approaches. One approach is to support people to understand each other's egocentric frame-of-reference, and one common idea is to share information about collaborators' viewpoints. With such information, collaborators can comprehend others' spatial expressions based on their egocentric frame-of-reference, so the misunderstanding can be reduced. For instance, Hindmarsh et al. proposed providing remote participants with other collaborators' perspectives in a virtual environment and indicated that it improved participants' spatial communication [17].

Another approach for solving the spatial expression problem is to induce people to use allocentric frame-of-reference or geocentric frame-of-reference instead of using egocentric frame-of-reference. Compared with egocentric frame-of-reference, spatial expressions designed based on allocentric frame-of-reference and geocentric frame-of-reference are relatively understandable. To achieve this, Muller et al. designed an AR-based remote collaboration system with floating virtual cubes and virtual objects, such as a bookshelf and a potted tree. These objects were considered shared visual landmarks, and people can leverage these virtual landmarks to perform effective spatial referencing to real objects. For example, when a speaker wants a listener to pick up a book, he/she can design expressions such as "pick up the book below the potted tree" instead of "pick up the book which is on my right-hand side." The evaluation showed that the shared virtual landmarks altered participants' behavior of spatial referencing; however, the virtual landmarks did not improve collaboration efficiency [22, 23]. Instead of using virtual objects to create shared visual landmarks, Chellali et al. included a virtual character as a stable visual landmark in remote collaboration, and the collaborators were able to design understandable spatial expressions by using the spatial relationship between the virtual character and the objects [6].

These two approaches are shown to improve people's spatial communication. However, these studies mainly focused on this short-term effect, and they did not pay much attention to whether participants acquired a better understanding of spatial information with such approaches. As we mentioned in the section 1, spatial knowledge plays an key role in remote collaboration; thus, we are interested in whether supporting collaborators in understanding each other's frame-of-reference also has a positive effect on spatial knowledge acquisition.

2.2 Spatial Knowledge in Remote Collaboration

Spatial knowledge also plays a vital role in spatial referencing during remote collaboration. While people are collaborating in wide environments, such as playing cooperative shooting games, working in a large warehouse, firefighters' group training, etc., people often conduct spatial referencing to objects or locations that are not in their view. For example, a leader of a firefighter team plans routes and gives instructions to team members to reach specific rooms. This collaboration task can perform successfully only when all team members obtain spatial knowledge about the environment.

There are mainly four types of spatial knowledge: landmark knowledge, route knowledge, graph knowledge, and survey knowledge [8, 31]. Landmarks are objects that can be used to identify specific locations, and a person with landmarks knowledge can recall the landmarks that appeared in the environment. Route knowledge includes a set of landmark-action associations, such as turning left when hitting the building, and a person with route knowledge is able to move from one position to another position along a specific path. Graph knowledge is knowledge of the connectivity in an environment. A person with graph knowledge knows how locations in an environment are connected together, and he/she has the ability to plan a route from one location to another location. Survey knowledge is a spatial configuration of the environment, and a person with survey knowledge understands distances and directions between locations in the environment. For remote collaboration, since the tasks are often diversified and randomized, collaborators require high-level spatial knowledge, such as graph knowledge and survey knowledge, to flexibly design appropriate spatial expressions.

There have been many studies investigating factors that influence spatial knowledge acquisition, such as the display size [1], 2D/3D [11], visual fidelity of the environment [32], spatial ability [21], age [15, 21], information processing ability [21, 35], active learning [7, 8], etc. However, most of them focused on individual spatial knowledge acquisition, and to the best of our knowledge, there are very few studies paying attention to spatial knowledge acquisition during collaboration. One research compared the difference between individual spatial learning and collaborative spatial learning, and the result indicated that collaborative spatial learning improved participants' survey knowledge [3]. Another research indicated that participants acquired spatial knowledge with individual exploration better than collaborative exploration which was better than competitive exploration [20]. These studies only compared different ways of spatial learning (single, collaboration, and competition), but there is still no research investigating and quantitatively measuring the factors that affect the spatial knowledge acquisition in collaboration scenario.

In addition to spatial knowledge, it is also important that the spatial knowledge is shared among collaborators in remote collaboration. It is widely known that common ground, the knowledge that is mutually obtained between collaborators, plays a key role in effective communication (e.g., [5, 9, 10, 14, 33]). Although past research mainly examined common ground related to object referencing (e.g., how people name a object during collaboration), it is highly possible that having shared spatial knowledge is important for remote collaboration involving lots of spatial referencing. To the best of our knowledge, there is no research relating to shared spatial knowledge.

Overall, the contribution of the current study is to investigate the effect of understanding others' frame-of-reference on spatial knowledge acquisition and shared spatial knowledge acquisition.

3 EXPERIMENTAL DESIGN

3.1 Hypotheses

In the current experiment, we examined how participants collaboratively acquired spatial knowledge of a wide virtual environment. As for the method of supporting participants in understanding each other's frame-of-reference, we selected view-sharing function, which is one of the most common ways. Besides, for spatial knowledge, we focused on two types of high-level spatial knowledge, graph knowledge and survey knowledge, due to their importance to remote collaboration.

Since collaborators understand each other's frame-of-reference, we assumed that the collaborators could effectively acquire graph knowledge (H1), survey knowledge (H2), and shared survey knowledge (H3). Moreover, as a secondary purpose, we are also interested in how acquired spatial knowledge affected collaboration quality afterwards. We assumed that the acquired spatial knowledge affected the collaboration quality further (H4). Thus, we set the following hypotheses:

- H1: Participants plan routes to move from location to location in an environment more efficiently and correctly after they collaboratively explore that environment with a view-sharing function than without such a function.
- H2: Participants recall the absolute position of entities in an environment more correctly after they collaboratively explore that environment with a view-sharing function than without such a function.
- H3: Participants recall the absolute position of entities in an environment more similar to their partners after they collaboratively explore that environment with a view-sharing function than without such a function.
- H4: Participants perform remote collaboration tasks in an environment more efficiently after they collaboratively explore that environment with a view-sharing function than without such a function.

Note that this experiment did not examine the effect of view-sharing on the shared graph knowledge.

3.2 Material

To examine these hypotheses, a maze exploration task was conducted (Fig. 1) [3, 4, 7, 8]. The mazes were generated by Prim's algorithm [28], and the shape of the mazes was orthogonal. One 5 × 5 maze was generated for the practice session; three 7 × 7 mazes were generated for the main session. Each 7 × 7 maze included 34 pieces of wall, and there were at least 8 dead-ends. In the maze, six different animal objects were placed at different dead-ends: elephant, penguin, cat, owl, chicken, and zebra. Besides, four landmarks were placed on the walls near the intersection of roads towards different objects: closet, clock, chair, and lamp.

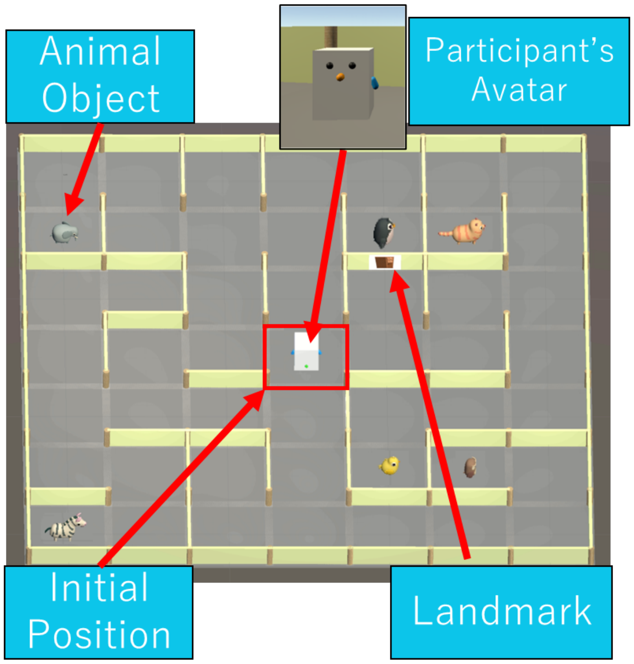

Figure 1: A sample of the maze. Six animal objects were placed at the different dead-ends, and four landmarks were placed on the walls. Note that the size of the objects and avatars are scaled up for better visibility. The participant's avatar appeared in the initial position, which is the center of the maze, at the beginning of the exploration phase and the collaboration phase. All participants' avatar designs were identical.

3.3 Method

This experiment was a within-participant design experiment with two conditions. One condition was the no-view-sharing condition, and the other was the view-sharing condition. In the view-sharing condition, the participants can see not only their own perspective but also their partner's perspective during the exploration phase (Fig. 2).

3.4 Participants

For this experiment, 16 participants (9 males and 7 females) were recruited, and eight groups were formed. The average age was 26, and SD was 7.59. Each participant received 3500 yen as compensation after the experiment. The written consent form was obtained from all participants.

3.5 Procedure

A pair of participants entered the experiment room, and each of them was asked to sit 50 cm in front of a 23.6-inch monitor (Philips 246E7Q). Two participants could not see each other physically, but they could hear each other's voice. They filled out the consent form after the experiment explanation. In the experiment explanation, the participants were informed about the maze size and the number and the type of objects and landmarks that were placed in the maze.

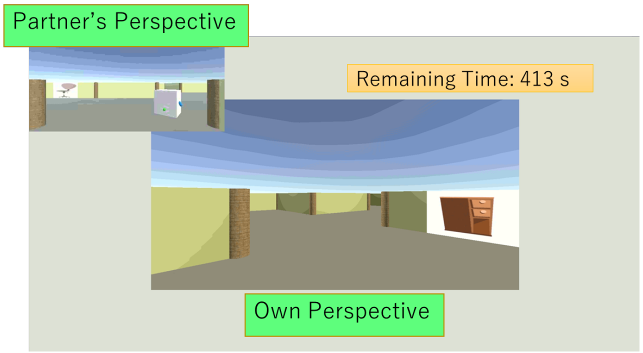

Figure 2: View-sharing function. For the view-sharing condition, participants could see both their own perspective and their partner's perspective during the exploration phase. The diagonal of the view for the own perspective was about 30cm, and it was about two times as large as the view for the partner's perspective.

Later, the practice session was conducted to ensure the participants understood the procedure. After that, the main session started. The main experiment contained four phases: exploration phase, test phase, drawing phase, and collaboration phase (Fig. 3).

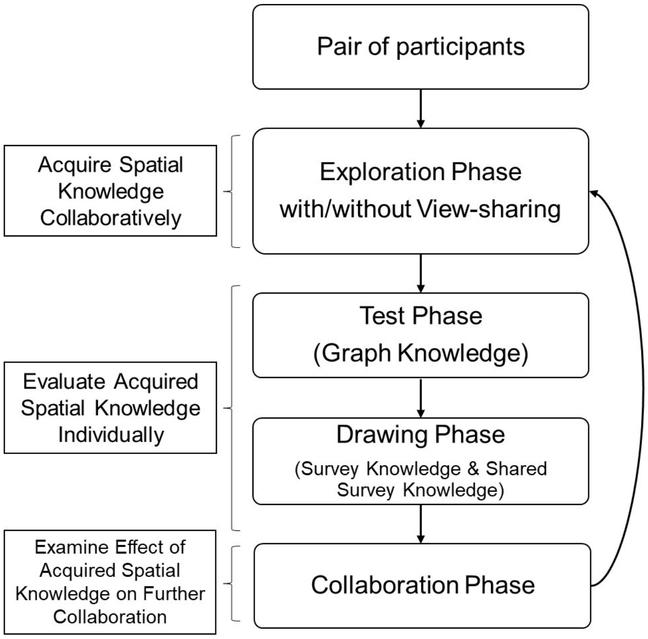

Figure 3: Diagram of the procedure in this experiment

3.5.1 *Exploration Phase.*

In the exploration phase, the monitors presented the participants' first person view. The participants saw themselves appeared in the initial position, which was the center of the maze. Later, they used a keyboard to move in the maze and collaboratively explored the maze for 7 minutes. The participants could see each other's avatars during the exploration (Fig. 1). The participants were asked to try their best to explore every place in the maze and understand its structure. Also, they were strongly requested that they should describe the maze structure they see with each other verbally. In the view-sharing condition, the participants were able to see both their own perspective and their partner's perspectives (Fig. 2).

3.5.2 Test Phase. The test phase began soon after the end of the exploration phase. The participants were asked to plan routes and move from object to object individually. First, the participants were randomly teleported to the location of one of the objects (initial object) in the maze. Later, they were requested to plan a route to move from the initial object to another object (destination object) in 15 seconds. During the planning, they were allowed to rotate their views to check their orientation, but they were not allowed to move back and forth. After 15 seconds, they followed their planned route and moved to the object in 55 seconds. They were informed that they do not necessarily follow the route they planned if they think the routes were wrong or they got lost. For those objects which were neither the initial object nor the destination object, they were replaced with red balls to prevent the participants from updating or acquiring new spatial knowledge during this phase. It should be noted that the two participants could not see each other's avatars and could not communicate or interact in this phase. After 55 seconds, the participants were teleported to another initial object no matter they successfully reached the destination object or not, and the next trial began. There were ten trials in the test phase. The initial object and the destination object were randomized, but the same object would not be assigned as the destination object in any consecutive trials.

3.5.3 Drawing Phase. In the drawing phase, each participant received a paper-based answer sheet with a maze grid (Fig. 4). Also, a list of objects and landmarks that appeared in the maze was provided. The participants were asked to draw the position of the walls, objects, and landmarks on the answer sheet. The time limit of the drawing phase was 5 minutes, and the participants were asked to stop at that time no matter he/she finished drawing or not.

3.5.4 Collaboration Phase. In the collaboration phase, the participants were asked to look at the monitor again. The monitor presented the same maze again. The maze structure and the position of the landmarks and objects in the maze were same as the maze in the exploration phase, and the only difference was that we placed coins floating at the center of each cell. The participants were asked to collaboratively collect all the coins in the maze as fast as possible. They were requested to discuss the strategy during the phase. There was no time limit in this phase, and the phase ended when all coins were collected. Note that since the purpose of this phase is to test the effect of spatial knowledge acquired in the exploration phase on further collaboration, the participants were not allowed to see their partner's perspective in either the view-sharing condition or the no-view-sharing condition to reduce the potential bias, such as the advantages of view-sharing on collaboration efficiency.

After finishing the four phases, there is a 30-minute rest to ensure that the participants recovered from the fatigue. Later the participants were asked to conduct the task again with another condition. The maze assigned to the two tasks was randomly selected from three candidate mazes, and the order of the condition was counterbalanced between groups.

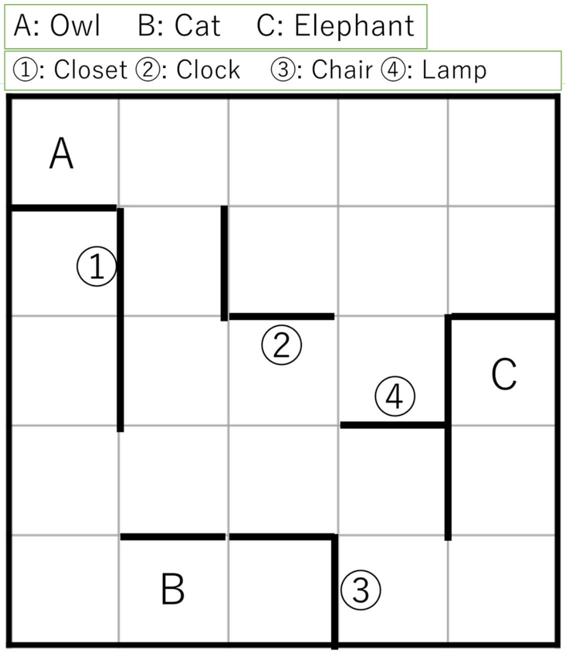

Figure 4: A sample of drawn maze

After the two tasks, each participant filled in a simulator sickness questionnaire [19] to check whether he/she experienced motion sickness during the experiment, which might affect the result. Later, the participants took another 30-minute rest. After the rest, each participant was asked to answer the Revised Purdue Spatial Visualization Test [36] and Santa Barbara Sense of Direction [16] to test his/her spatial ability.

4 RESULT

4.1 Demographic

As for the spatial ability, a group of participants did not conduct the two spatial ability tests because the experiment time was exceeded. For the other seven groups (14 participants), the average score of the Revised Purdue Spatial Visualization Test was 21 (SD=5.3), and the average score of the Santa Barbara Sense of Direction was 3.9 (SD=0.95). The result showed that the participants had a normal spatial ability, and there was no outlier, which might strongly affect the result of acquired spatial knowledge. Besides, for the motion sickness, the average score was 0.7 (SD=0.43), so there was no motion sickness occurring during the experiment.

4.2 Test phase

For the test phase, we first compared whether the conditions affect the success of each trial, i.e., whether the participant reached the destination object in time or not (Fig. 5left). A generalized linear mixed model with binomial distribution were constructed. The participant and group were considered as random intercept factors and added to the model. Analysis of deviance table with type III Wald chi-square tests showed that the success in the view-sharing condition was more than the success in the no-view-sharing condition (χ^2(1)=21.65, p<.001, d=1.64).

Later, for the succeed trials, the time each participant took to finish, i.e., moving from an initial object to a destination object, was calculated (Fig. 5right). A linear mixed model was used to model the data. The participant and group were considered as random intercept factors and added to the model. The result of the analysis of deviance table with type III Wald F tests with Kenward-Roger df showed that the time participants took in the no-view-sharing condition (M=60.63, SD=11.39) was significantly higher than the time in the view-sharing condition (M=54.55, SD=12.70) (F(1, 183.37)=11.384, p<.001, d=0.50).

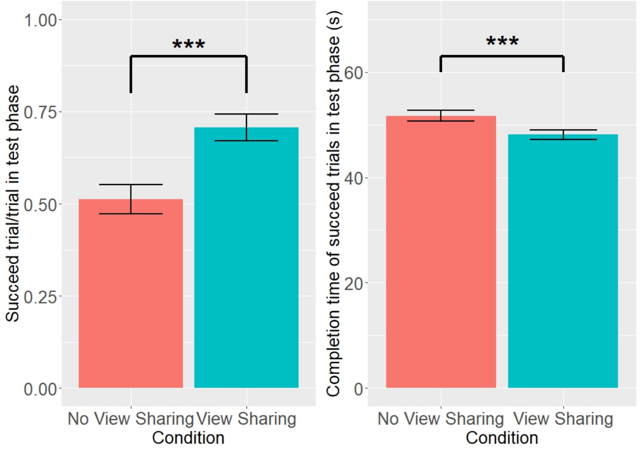

Figure 5: Left: result of succeed trial per trial; right: result of completion time for each succeed trial. All error bars represent standard errors.

4.3 Drawing phase

For the drawing phase, the correctness of drawn mazes was calculated (Fig. 6 A, B, and C). Regarding the way of calculating correctness, the positions of each wall, each landmark, and each object drawn by participants were compared with the correct answer. The wall, landmark, or object was considered correct if the absolute position was correct. Later, the number of correct walls, correct landmarks, and correct objects were summed up. Note that, since the direction of the drawn mazes and the correct maze might differ, each drawn maze was compared with the correct maze four times in four directions. The largest number of correct walls, correct landmarks, and correct objects was considered as the final result.

Due to violation of normality assumptions and the outcome was count data, generalized linear mixed models with Poisson distribution were used to model the number of the correct objects and the number of the correct landmarks. Participant was considered as a random intercept factor and added to the models. The analysis of deviance tables with type III Wald chi-square tests were calculated to examine the effect of the condition. The result of the number of correct objects showed that there was no significant difference between the number of correct objects in the no-view-sharing condition (M=2.38, SD=1.86) and the view-sharing condition (M=2.44, SD=1.59) (χ^2(1)=0.013, p=.91, d=0.03). For the number of correct landmarks, the result of the analysis of deviance table with type III Wald F tests with Kenward-Roger df showed that the number of correct landmarks in the view-sharing condition (M=0.25, SD=0.45) was significantly more than the number in the no-view-sharing condition (M=0.94, SD=0.93) (χ^2(1)=5.52, p=.019, d=1.32).

For modeling the number of correct walls, a linear mixed model was used, and the participant and the group were considered as random intercept factors and added to the model. The result of the analysis of deviance table with Type III Wald F tests with Kenward-Roger df showed that the number of correct walls in the view-sharing condition (M=17.75, SD=6.82) was significantly higher than the number of correct walls in the no-view-sharing condition (M=14.63, SD=6.55) (F(1, 15)=4.83, p=.044, d=0.78).

Besides, the similarity of the drawn mazes from participants in the same groups was calculated (Fig. 6 D, E, and F). The walls, landmarks, and objects which were drawn in the same positions were considered that the two participants had shared survey knowledge of those walls, landmarks, and objects. The number of same-position walls, same-position landmarks, and same-position objects were later summed up.

Due to violation of normality assumptions and the outcome was count data, generalized linear mixed models with Poisson distribution were used to model the number of same-position objects, the number of same-position landmarks, and the number of same-position walls. The group was considered as a random intercept factor and added to the models. The analysis of deviance tables with type III Wald chi-square tests were calculated to examine the effect of the condition. The result of the number of same-position objects showed that there was no significant difference between the number of same-position objects in the view-sharing condition (M=2.25, SD=1.28) and in the no-view-sharing condition (M=1.38, SD=1.69) (χ^2(1)=1.66, p=.20, d=0.49). For the number of same-position landmarks, the result showed that there was no significant difference between the number of same-position landmarks in the view-sharing condition (M=1.00, SD=0.93) and the number in the no-view-sharing condition (M=0.25, SD=0.71) (χ^2(1)=3.07, p=.08, d=1.39).

For modeling the number of same-position walls, a linear mixed model was constructed, and the group was considered as a random intercept factor and added to the model. The result of the analysis of deviance table with type III Wald F tests with Kenward-Roger df showed that there was no significant difference between the number of objects in the view-sharing condition (M=13.13, SD=4.76) and the no-view-sharing condition (M=10.13, SD=7.26) (F(1, 7)=2.47, p=.16, d=0.79).

4.4 Collaboration phase

For the collaboration phase, the completion time was calculated, and a linear mixed model was used to model the data (Fig 7). The group was considered as a random intercept factor and added to the model. The result showed that there was no significant difference between view-sharing condition (M=208.15, SD=47.57) and no-view-sharing condition (M=227.71, SD=67.06) (F(1, 7)= 3.01, p=.13, d=0.87).

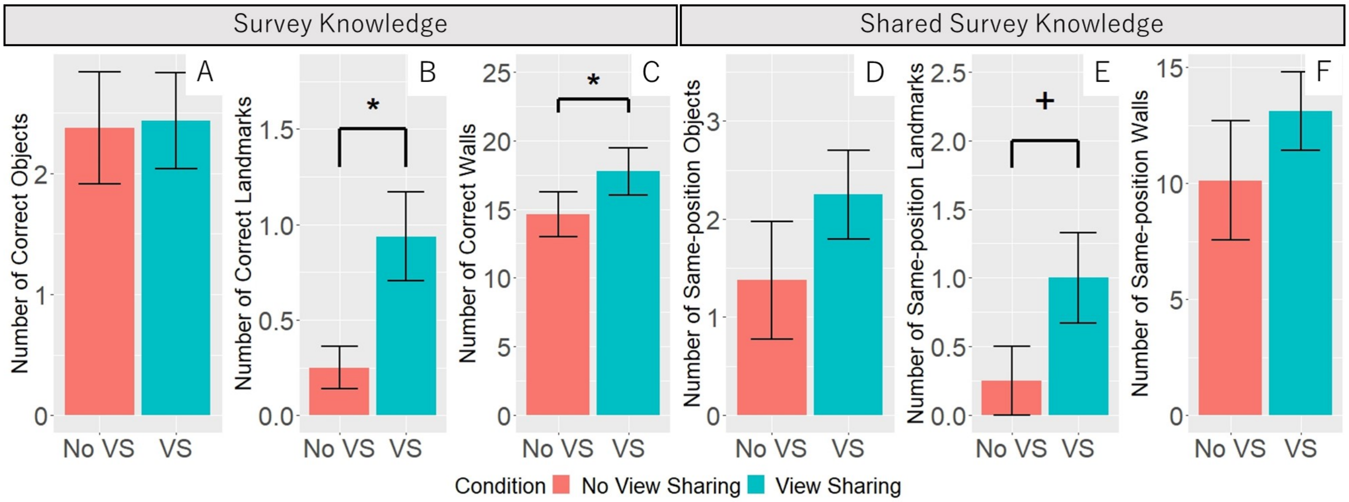

Figure 6: A: number of correctly recalled objects; B: number of correctly recalled landmarks; C: number of correctly recalled walls; D: number of objects recalled in same positions; E: number of landmarks recalled in same positions; F: number of walls recalled in same positions. All error bars represent standard errors.

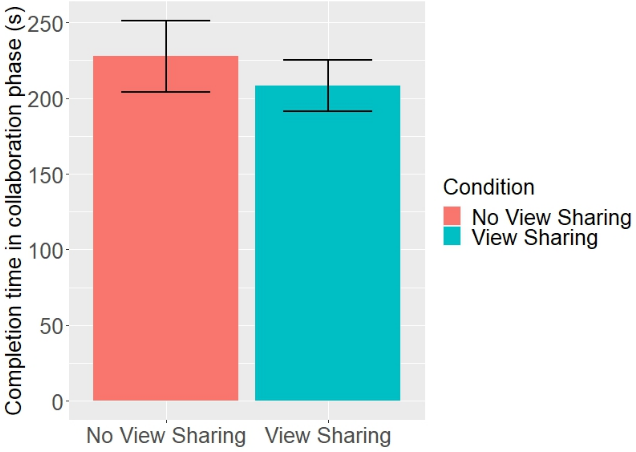

Figure 7: Result of collaboration phase. All error bars represent standard errors.

5 DISCUSSION

5.1 Acquired Spatial Knowledge

Regarding H1 (participants plan routes to move from location to location in an environment more efficiently and correctly after they collaboratively explore that environment with a view-sharing function than without such a function), the result of the test phase showed that the participants reached the destination objects more successfully and quickly in the view-sharing condition; thus, H1 was supported. These results indicated that the participants had a better understanding of the connections between objects and suggested that the participants acquired graph knowledge better in the view-sharing condition.

Regarding H2 (participants recall the absolute position of entities in an environment more correctly after they collaboratively explore that environment with a view-sharing function than without such a function), the result of the drawing phase showed that the participants correctly recalled the position of the walls and landmarks in the view-sharing condition. Although the number of correctly recalled objects had no significant difference between the two conditions, the results partially supported H2. These findings also indicated that the participants in the view-sharing condition correctly learned the spatial configuration of the maze and acquired the survey knowledge.

Taken together, our results clearly showed that the view-sharing function supported participants in obtaining spatial knowledge effectively. This also pointed out that understanding each other's frame-of-reference is positively associated with spatial knowledge acquisition. As a possible reason, past research has indicated that collaborative navigation had a positive effect on spatial knowledge [3, 20]. Buck et al. indicated that dyadic navigation improved the survey knowledge compared with individual exploration. Besides, Liang et al.'s research showed that participants gained better spatial knowledge while collaboratively exploring a virtual environment compared with competitively exploring the virtual environment. Both studies suggested that communication was a possible factor that improved spatial knowledge acquisition. Since the participants could conduct effective spatial communication and spatial referencing based on a better understanding of others' frame-of-reference, it is possible that the spatial knowledge was effectively acquired.

5.2 Shared Spatial Knowledge

Regarding H3 (participants recall the absolute position of entities in an environment more similar to their partners after they collaboratively explore that environment with a view-sharing function than without such a function), the result of the drawing phase showed that there was no significant difference between the two conditions for walls, landmarks, or objects. Thus, H3 was not supported.

It is surprising that there was a significant difference between conditions on survey knowledge but not on shared survey knowledge. One possible reason is that the sample size to test the shared survey knowledge was small. However, another possible reason is that although a better understanding of each other's frame-of-reference helps participants to encode the spatial knowledge better, simply comparing the similarity of two drawn mazes may not be the most appropriate way to assess the shared survey knowledge since the way of spatial knowledge encoding varied between participants [34, 35].

5.3 Impact of Spatial Knowledge on Quality of Remote Collaboration

Regarding H4 (participants perform remote collaboration tasks in an environment more efficiently after they collaboratively explore that environment with a view-sharing function than without such a function), although the previous results showed the participants in the view-sharing condition acquired correct spatial knowledge, the result of the collaboration phase indicated that there was no significant difference in the time participants took between conditions. Thus, the results did not support H4.

There are two possible explanations to explain why spatial knowledge did not improve efficiency. First, the efficiency of the coin collecting task was sensitive to human error. We observed several instances that the participants walked through the cells without touching the coins and did not aware that they failed to collect the coins. Therefore, they spent much time searching for those coins later, and the spent time sometimes accounted for a large proportion of the total time. As a future study, an appropriate experimental design is necessary to examine the effect of spatial knowledge on collaboration quality. Second, it is possible that the spatial knowledge was stored in the short term. The exploration phase only lasted for 7 minutes, and there were the test phase and the drawing phase before the collaboration phase began. We are aware that the spatial knowledge acquired in the exploration phase might be weakened by the time of the collaboration phase. This may be the reason why the task completion time of the two conditions was not significantly different. Although H4 is not the primary purpose of the current study, our findings remains an unsolved issue: how and what kind of spatial knowledge influences the quality of remote collaboration.

5.4 Limitation and Future work

There are some limitation in this study. First, this study chose the view-sharing method which is a common way to share frame-of-reference, but we did not examine how much the participants understood each other's frame-of-reference through the view-sharing function. The effect of view-sharing on participants' understanding of other's frame-of-reference can be affected by 1) whether the participants could effectively distribute their attention to both views and 2) whether they could effectively integrate the information from the two views. Different degree of understanding might lead to different degree of spatial knowledge acquisition, so this could be a confounding factor of this experiment. Although our experimental design has minimized the influence of other possible bias, it is still necessary to test this to strengthen our findings in the future. Second limitation is the small sample size of this research. Despite the small sample size, the current research demonstrated several interesting results. Thus, it is necessary to increase the number of participants to strengthen the findings.

Besides, our study investigated the relationship between the frame-of-reference and spatial knowledge, but there is no direct evidence of how understanding others' frame-of-reference enhanced spatial knowledge acquisition. A further study with detailed qualitative analysis to investigate how spatial knowledge was acquired differently is necessary. Also, as mentioned in section 2.2, the common ground of spatial knowledge might play an important role in remote collaboration. In this study, we only investigated the shared spatial knowledge but not the common ground of spatial knowledge. The shared spatial knowledge is the spatial knowledge that both collaborators obtain, while the common ground of spatial knowledge is the spatial knowledge that both collaborators obtain, and they know that each other has that spatial knowledge. In the current research, the shared spatial knowledge between participants in the same group was evaluated by measuring the similarity of two drawn mazes. However, it is necessary to assess how participants obtain common ground of spatial knowledge in the future.

The generalisability of the findings is another potential issue of this study. In our study, we used small desktop displays as our experiment devices. The reason for using small desktop displays was because the severe VR sickness occurred while using HMD during the pilot test. However, several past research has shown that the field-of-view and level of immersion have some effects on the spatial knowledge acquisition [12, 27]. Thus, it is necessary to repeat the experiment with different devices, such as HMDs, CAVEs, and large displays, to confirm the findings. However, in these experiments, the position and size of the two views on the user interface should be well considered since the large amount of information from wide field-of-view might make participants hard to integrate the information and weaken the effect of view-sharing on spatial knowledge acquisition.

6 CONCLUSION

In this research, we investigated the effect of view-sharing on spatial knowledge and shared spatial knowledge acquisition in remote collaboration, and the result showed that view-sharing improved participants' graph knowledge and survey knowledge acquisition. These findings highlighted the importance of sharing participants frame-of-reference, and also suggested that the process of remote collaboration might change the participants' acquired spatial knowledge differently.

ACKNOWLEDGMENTS

This work is supported by Foundation for the Fusion Of Science and Technology Research Project C.

REFERENCES

[1] Jonathan Z Bakdash, Jason S Augustyn, and Dennis R Proffitt. 2006. Large displays enhance spatial knowledge of a virtual environment. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*. 59–62.

[2] Bruce Bridgeman. 1999. Separate representations of visual space for perception and visually guided behavior. In *Advances in psychology*. Vol. 129. Elsevier, 3–13.

[3] Lauren E Buck, Timothy P McNamara, and Bobby Bodenheimer. 2020. Dyadic Acquisition of Survey Knowledge in a Shared Virtual Environment. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 579–587.

[4] Sonia Cárdenas-Delgado, Magdalena Méndez-López, M-Carmen Juan, Elena Pérez-Hernández, Javier Lluch, and Roberto Vivó. 2017. Using a virtual maze task to assess spatial short-term memory in adults. In *International Conference on Computer Graphics Theory and Applications*, Vol. 2. SCITEPRESS, 46–57.

[5] Yuan-Chia Chang, Hao-Chuan Wang, Hung-kuo Chu, Shung-Ying Lin, and Shuo-Ping Wang. 2017. AlphaRead: Support unambiguous referencing in remote collaboration with readable object annotation. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2246–2259.

[6] Amine Chellali, Isabelle Milleville-Pennel, and Cédric Dumas. 2013. Influence of contextual objects on spatial interactions and viewpoints sharing in virtual environments. *Virtual Reality* 17, 1 (2013), 1–15.

[7] Elizabeth R Chrastil and William H Warren. 2013. Active and passive spatial learning in human navigation: acquisition of survey knowledge. *Journal of experimental psychology: learning, memory, and cognition* 39, 5 (2013), 1520.

[8] Elizabeth R Chrastil and William H Warren. 2015. Active and passive spatial learning in human navigation: acquisition of graph knowledge. *Journal of experimental psychology: learning, memory, and cognition* 41, 4 (2015), 1162.

[9] Herbert H Clark. 1996. *Using language*. Cambridge university press.

[10] Herbert H Clark and Catherine R Marshall. 1981. Definite knowledge and mutual knowledge. In *"Elements of discourse understanding"*. Cambridge University Press.

[11] Andy Cockburn and Bruce McKenzie. 2002. Evaluating the effectiveness of spatial memory in 2D and 3D physical and virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 203–210.

[12] Mary Czerwinski, Desney S Tan, and George G Robertson. 2002. Women take a wider view. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 195–202.

[13] Sarah D'Angelo and Darren Gergle. 2016. Gazed and confused: Understanding and designing shared gaze for remote collaboration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2492–2496.

[14] Darren Gergle, Robert E Kraut, and Susan R Fussell. 2013. Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction* 28, 1 (2013), 1–39.

[15] Denise Head and Marlisa Isom. 2010. Age effects on wayfinding and route learning skills. *Behavioural brain research* 209, 1 (2010), 49–58.

[16] Mary Hegarty, Anthony E Richardson, Daniel R Montello, Kristin Lovelace, and Ilavanil Subbiah. 2002. Development of a self-report measure of environmental spatial ability. *Intelligence* 30, 5 (2002), 425–447.

[17] Jon Hindmarsh, Mike Fraser, Christian Heath, Steve Benford, and Chris Greenhalgh. 1998. Fragmented interaction: establishing mutual orientation in virtual environments. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. 217–226.

[18] Janet G Johnson, Danilo Gasques, Tommy Sharkey, Evan Schmitz, and Nadir Weibel. 2021. Do you really need to know where "that" is? enhancing support for referencing in collaborative mixed reality environments. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[19] Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.

[20] Hai-Ning Liang, Feiyu Lu, Yuwei Shi, Vijayakumar Nanjappan, and Konstantinos Papangelis. 2019. Evaluating the effects of collaboration and competition in navigation tasks and spatial knowledge acquisition within virtual reality environments. *Future Generation Computer Systems* 95 (2019), 855–866.

[21] Ismini E Lokka and Arzu Çöltekin. 2020. Perspective switch and spatial knowledge acquisition: Effects of age, mental rotation ability and visuospatial memory capacity on route learning in virtual environments with different levels of realism. *Cartography and Geographic Information Science* 47, 1 (2020), 14–27.

[22] Jens Müller, Roman Rädle, and Harald Reiterer. 2016. Virtual objects as spatial cues in collaborative mixed reality environments: How they shape communication behavior and user task load. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1245–1249.

[23] Jens Muller, Roman Rädle, and Harald Reiterer. 2017. Remote collaboration with mixed reality displays: How shared virtual landmarks facilitate spatial referencing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 6481–6486.

[24] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human–computer interaction* 15, 2-3 (2000), 139–178.

[25] Gary M Olson, Judith S Olson, and Gina Venolia. 2009. What still matters about distance. *Proceedings of HCIC* (2009).

[26] Sungjoon Park, Brandon K Watanabe, and Heather Burte. 2022. Perspective taking and reference frames for spatial and social cognition. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 44.

[27] Kimberly A Pollard, Ashley H Oiknine, Benjamin T Files, Anne M Sinatra, Debbie Patton, Mark Ericson, Jerald Thomas, and Peter Khooshabeh. 2020. Level of immersion affects spatial learning in virtual environments: results of a three-condition within-subjects study with long intersession intervals. *Virtual Reality* 24, 4 (2020), 783–796.

[28] Robert Clay Prim. 1957. Shortest connection networks and some generalizations. *The Bell System Technical Journal* 36, 6 (1957), 1389–1401.

[29] Ralph J Roberts Jr and Christine J Aman. 1993. Developmental differences in giving directions: Spatial frames of reference and mental rotation. *Child development* 64, 4 (1993), 1258–1270.

[30] Anna Shusterman and Peggy Li. 2016. Frames of reference in spatial language acquisition. *Cognitive psychology* 88 (2016), 115–161.

[31] Alexander W Siegel and Sheldon H White. 1975. The development of spatial representations of large-scale environments. *Advances in child development and behavior* 10 (1975), 9–55.

[32] Grégory Wallet, Hélène Sauzéon, Prashant Arvind Pala, Florian Larrue, Xia Zheng, and Bernard N'Kaoua. 2011. Virtual/real transfer of spatial knowledge: Benefit from visual fidelity provided in a virtual environment and impact of active navigation. *Cyberpsychology, Behavior, and Social Networking* 14, 7-8 (2011), 417–423.

[33] Tzu-Yang Wang, Yuji Sato, Mai Otsuki, Hideaki Kuzuoka, and Yusuke Suzuki. 2020. Developing an AR remote collaboration system with semantic virtual labels and a 3D pointer. In *International Conference on Human-Computer Interaction*. Springer, 407–417.

[34] Steven M Weisberg and Nora S Newcombe. 2016. How do (some) people make a cognitive map? Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 42, 5 (2016), 768.

[35] Wen Wen, Toru Ishikawa, and Takao Sato. 2011. Working memory in spatial knowledge acquisition: Differences in encoding processes and sense of direction. *Applied Cognitive Psychology* 25, 4 (2011), 654–662.

[36] So Yoon Yoon. 2011. *Psychometric properties of the revised purdue spatial visualization tests: visualization of rotations (The Revised PSVT: R)*. Purdue University.