Collaborative Remote Control of Unmanned Ground Vehicles in Virtual Reality

Ziming Li, Yiming Luo, Jialin Wang, Yushan Pan, Lingyun Yu, Hai-Ning Liang*
Department of Computing, *School of Advanced Technology*
Xi'an Jiaotong-Liverpool University, Suzhou, China
* Corresponding author (haining.liang@xjtlu.edu.cn)

2022 International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET) | 978-1-6654-7016-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMET54801.2022.9929783

***Abstract*—Virtual reality (VR) technology is commonly used in entertainment applications; however, it has also been deployed in practical applications in more serious aspects of our lives, such as safety. To support people working in dangerous industries, VR can ensure operators manipulate standardized tasks and work collaboratively to deal with potential risks. Surprisingly, little research has focused on how people can collaboratively work in VR environments. Few studies have paid attention to the cognitive load of operators in their collaborative tasks. Once task demands become complex, many researchers focus on optimizing the design of the interaction interfaces to reduce the cognitive load on the operator. That approach could be of merit; however, it can actually subject operators to a more significant cognitive load and potentially more errors and a failure of collaboration. In this paper, we propose a new collaborative VR system to support two teleoperators working in the VR environment to remote control an uncrewed ground vehicle. We use a compared experiment to evaluate the collaborative VR systems, focusing on the time spent on tasks and the total number of operations. Our results show that the total number of processes and the cognitive load during operations were significantly lower in the two-person group than in the single-person group. Our study sheds light on designing VR systems to support collaborative work with respect to the flow of work of teleoperators instead of simply optimizing the design outcomes.**

***Index Terms*—Virtual Reality, Collaborative Remote Control, Unmanned Ground Vehicles**

I. Introduction

In recent years, a number of studies in the field of computer-supported cooperative work (CSCW) have explored the use of emerging virtual reality (VR) technology as a means of supporting cooperative work [1]. The development of these cooperative virtual environments (CVEs) represents a radical shift away from the more commonly used single-user VR systems. CVEs are characterized by the rendering of virtual places inhabited by multiple distributed users, who are mutually represented to each other in order to support cooperative action and interaction within the shared virtual space [2]–[5]. A number of CSCW applications of this technology have previously been reported, including document editing [6], video conferencing [7], video editing [8], data analysis [9], information gathering [10] and interacting with virtual objects (building [11] and workspace [12]). In addition, a few CSCW studies have addressed the design of remote control centers for autonomous vessels [13]. All these mutual reality applications attempt to utilize CVE technology to provide arenas in which social interaction can occur and cooperative work can be supported.

Perhaps surprisingly, many practitioners have focused on VR as a virtual interface for use with CSCW systems [14]. Indeed, for many in the field of CSCW, these applications and related research represent new interface technologies for interacting with computational resources. Thus, they allow for the possibility of using a form of interface to explore the value of social science contributions to the design and evaluation of such systems [15]. For example, in relation to remote control crewless vehicles, prior studies have explored the use of VR to teleoperate robots, including unmanned ground/flying vehicles (UGVs/UFVs) [16], [17]. Such studies have shown that VR effectively supports efficient and accurate teleoperations [18], [19]. In fact, the use of a VR interface was found to lower operational costs and reduce the workload in some cases, especially when it comes to UGVs traveling within complex environments [20], [21].

However, simply designing CSCW systems for use in VR environments cannot ensure either good cooperation or a balanced workload. As teleoperation becomes more complex, it places a more significant cognitive load on the teleoperator [22], which can lead to more operational errors [23]–[25]. Although prior studies have considered familiar control layouts as a means by which interaction designers can reduce the cognitive load facing teleoperators through developing CVE functionalities [26], [27], the common failure to properly evaluate the collaboration work that takes place within virtual systems means that it remains questionable whether the cognitive load will actually be reduced. For example, while some studies have investigated the use of VR for learning and education [28]–[30] and industrial applications [31], [32], such as vehicles, cranes and trucks, they have tended to use similar control layouts in an effort to reduce the cognitive load experienced by their operations. These control layouts do not exist in a vacuum. Rather, the design of similar control layouts is in line with the principles of affordance [33], articulation work [34], [35] and grounding [36]. For instance, nuclear power plants use standardised information display layouts in order to reduce their controllers' cognitive load and help their operators to access key information quickly and accurately [36]. Moreover, online learning platforms help students to reduce the cognitive load associated with the learning process by implementing easy and familiar layouts, thereby assisting

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

students in improving their learning efficiency [37]. Without understanding the distributed cognition involved in interactive systems, it will be impossible to determine the collaborative work between people within the environment, as well as the resources and materials involved, that can serve to reduce the workload.

Thus, we consider the cognitive load stemming from cooperative work to represent a mutual means of shaping and reshaping the design of remote control in VR environments. In light of this, our main research question concerns how a collaborative multi-person approach could effectively reduce the cognitive load associated with operating a remote-control vehicle using VR means. More specifically, in the present study, we use VR for the remote control and manipulation of UGVs in order to examine the effectiveness of the task sharing and performance of participants in our pre-designed VR environment. Moreover, we explore whether the remote control of a UGV using a mechanical arm by two participants working collaboratively in a VR interface can be understood as an effective and accurate way of supporting the participants' collaboration. We conducted two experiments to evaluate our design of collaborative remote control in a VR environment. We simultaneously observed the work needed to complete the required tasks and the total number of operations performed by participants. The findings of our study contribute to the literature by suggesting designs for effective interface layouts for collaborative teleoperations that are relatively complex in nature.

The remainder of this paper is organized as follows. We present the related work in section II, wherein we position the present work in relation to prior studies. In section III, we discuss the scope of the system and provide an overview to help readers develop a picture of the experimental settings. This is important because it allows for a holistic understanding of how we implemented the technical elements to facilitate collaborative work concerning the remote control of unmanned vehicles in CVEs, as set out in section IV. Section IV details the experiments, including the applied procedures, formulated hypotheses, and results. Subsequently, section VI presents and discusses the findings. Finally, section VII concludes the paper by elucidating the limitations of the research.

II. Related work

A. Cooperation within virtual environments

VR, as a virtual interface in CSCW systems, has been the subject of significant research interest in recent years. Such interfaces are designed to support cooperative work by providing a common spatial frame of reference for collaboration of various types and in various contexts. For example, in an organizational setting, VR can take the form of video conferencing or desktop applications, although such applications are not limited to organizational settings or work practices per see and can also provide support for, for example, entertainment [38]. Cooperative work occurs in situations where multiple participants are mutually dependent on each other when it comes to completing a shared task and, therefore, are required to engage in articulation work [39]. This interdependence forms the core of the very definition of cooperative work. To have a cooperative work situation, at least two people must depend on each other to solve a shared task, and in such situations, the need for articulation work arises. In the case of immersive technologies such as VR, augmented reality (AR), and extended reality (XR), articulation work can require that multiple people 'articulate' (i.e. divide, allocate, coordinate, schedule, mesh and interrelate) their distributed individual activities through a designed digital object, thereby sharing artifacts and knowledge through a combination of voice, gesture, audio and graphical information that can be used in the real world, virtual world or both to form a flow of work [40], [41].

Furthermore, as a virtual interface for cooperative work, VR can be divided into synchronous and asynchronous collaboration with regard to the dynamic changes that occur over time. Remote locations and synchronous collaboration within a virtual environment can support users in working together by immersing them in a co-located setting. In fact, fully immersive virtual environments can support remote and real-time multi-user collaboration, interaction and information and data sharing, which is why they are used in domains such as gaming [42], entertainment education [43], therapy [44], training [45], decision making and problem solving [46]. Different tools, features, and functions are available for directly manipulating objects, navigating within the environment, encountering people, and sharing visual artifacts. In addition, collaborative tools and sharing mechanisms such as instant-, audio-, and video messaging can be featured. This enables joint contribution whereby users can simultaneously work together as data are instantly modified.

When reviewing the prior CSCW research concerning VR, we find that CSCW researchers are aware of how people who engage in work and articulate work use different strategies to reduce the effort associated with articulation work, such as awareness [34] and grounding [36]. When designing VR, we recognize that it is important to imitate real-life work tasks and mimic how cooperative work participants would be allowed to engage in awareness and grounding activities when 'testing' the remote control of unmanned vehicles. In most cases, however, the design of awareness cannot be fully understood once the experimental setting is established. Thus, the uncertainty of the VR environment and the limited knowledge of actors' cooperative work could increase the unnecessary workload among cooperating actors. Most research contexts are taken for granted, and VR practitioners tend to overlook different ways to design for peripheral awareness—allowing participants to 'see at glance' who is present and what they do. Therefore, the present study is influenced by previous VR studies but aims to elucidate the capturing and later replaying of multi-modal interactions (i.e. speech, PC, joystick, and scene manipulations) in order to enable cooperating participants to recognize and share when they have a common grounding. This common grounding will help people to expand their grounding references (i.e. pointing to certain artifacts that

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

convey information) and use other cues.

B. Measuring the distributed cognitive workload

Given the above-mentioned issues, we designed a multi-person collaboration experiment intended to measure the cognitive load so as to help us enhance the exploration of the collaborative remote control of UGVs. A reduced cognitive load will allow participants to understand things more quickly and exhibit fewer errors when performing in group setting operations that demand lower levels of cognitive load, for instance, when such operations are divided among multiple users. Yet, the cognitive load has traditionally been associated with individual work [47]–[50] and used to generate a variety of instructional efforts. Such efforts also influence the effectiveness and efficiency of collaborative work, whether computer-supported or face-to-face work, although they are often not considered when designing CVEs or researching collaborative work practices. The need to consider the cognitive load in collaborative environments involving VR applications is due to the new millennium requiring an enhanced understanding of the emerging dynamics of interaction [51]. The focal task is no longer confined to be desktop but instead reaches into the complex world of information and computer-mediated interactions. VR practitioners are expected to understand the interactions between people and technologies, including the answers to the following questions. How can we coordinate our activities in virtual environments? How do we articulate our work to co-workers? What can we do to ensure that the flow of work in VR is not subject to breakdowns? What techniques or tools can be used to successfully mediate collaborative work?

One difficulty associated with VR practitioners' measurement of cognition might be related to the fact that prior human–computer interaction (HCI) research has taken many psychological studies for granted [40]. Cause-and-effect relationships cannot explain why a high cognitive load can render completing a task difficult [52], [53] or why a cooperative workload is sometimes low [54], [55] and sometimes not [55]. This results in a near-endless need for experiments to identify the possible factors that may affect the results (i.e. the p-value) [56], [57]. Designing studies wisely, particularly in the HCI and CSCW fields, is essential [58]. In line with the approach of previous works, our study focuses on identifying how the way in which participants' cognitive load is measured during an experiment can significantly impact the outcome of the investigation. The task in question involves two participants working together to complete a remove manipulation and explore its effect on both completion and the perceived workload. Importantly, we do not aim to find a yes or no answer to the question of whether the cognitive load is reduced or not; rather, we hope to identify the cues that enable people to be aware and ground their activities during collaborations so as to trigger a reasonable cognitive workload per person.

III. System scope and overview

We developed a VR UGV remote control platform with two versions, one for two participants and another for one, in Unity3D. Therefore, in this study, there were two conditions: (1) Dual-users and (2) Single-users.

A. System architecture and workflow for Collaboration use case

Figure 1.a shows the system architecture and workflow for our collaboration interaction experiments. There are two users, User A and User B. User A is the source of the UGV's body movement signal in the system via a left joystick. User B is the source of the UGV's arm control signal in the system via a right joystick. Both users can receive the version data of the UGV through the Oculus Quest 2's HMD. The principle is that the Control PC must receive the camera singles from the RoboMaster before being simultaneously distributed to PC A and B. This step is vital because, without such singles, it is impossible to create immersive environments for both participants to control the UGV.

In addition to the intervention of the two Oculus Quest 2 sets, the collaborative experiment was completed with the participation of these devices: two laptops that support Oculus Link (a data cable used to help Oculus Quest 2 HMD rendering with the help of the computer's rendering performance), a computer that was used to assist in transmitting and counting the operational commands used and a computer that could control the UGV over the LAN.

In the collaborative experiment, there are two types of signals. One is the control signal for the movement of the vehicle's body, which is sent by User A in Figure 1 with the help of an Oculus Quest 2 left-handed joystick and is transmitted to PC A. PC A receives the signal for the movement of the vehicle and sends it directly to the Counter PC via the LAN. Similarly, the signal from User B to control the arm is sent from PC B to Counter PC by the same method with the help of a right-handed joystick. Once the Counter PC has received both signals, it will count them and send them to the Control PC controlling the UGV through LAN. Finally, the UGV will operate on the basis of the signals received.

The onboard lens from the UGV sends its field of view data in real-time to Control PC, which sends it via LAN to PC A and PC B. After that, PC A and PC B will help the Oculus HMD to display the field of view data and present it to the users. We use a Counter PC to facilitate singles before sending them to the Control PC for directly outputting all experiment data as a whole. This helps to reduce the time complexity for outputting each experiment data. In that case, data from PC A and PC B can be simultaneously transformed to the Control PC without disruption, neither from PC A nor PC B.

B. System architecture and workflow for single use case

Figure 1.b shows the system architecture and workflow for our single interaction experiments. Unlike collaborative experiments, single experiments will have one user doing all

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

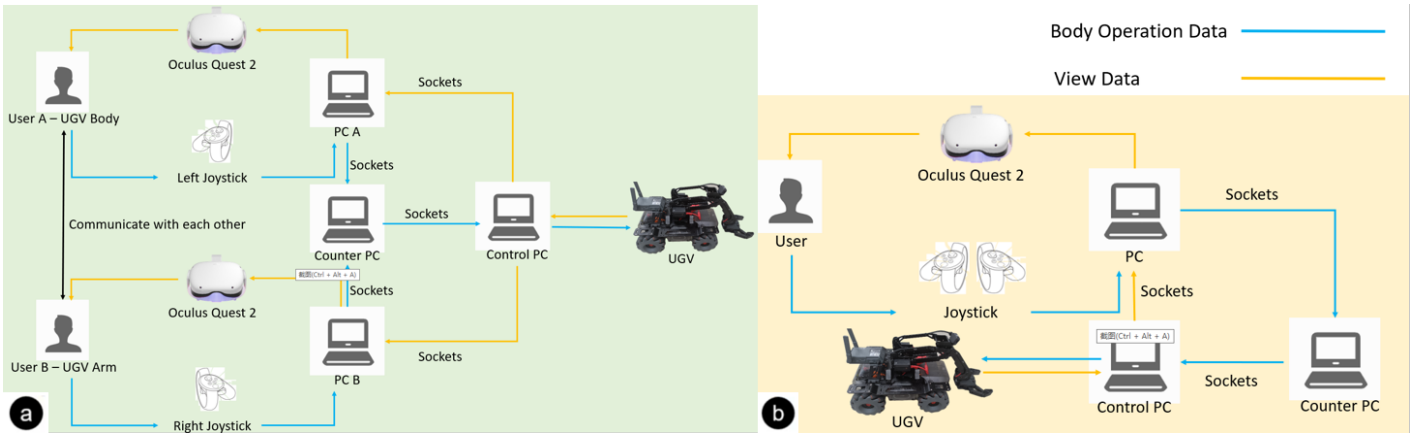

Fig. 1. (a: green background) Logical flow chart for the collaboration interaction (Users can communicate with their partner); (b: yellow background) Logical flow chart for the single interaction.

the controlling of the UGV. UGV's view data will also only be presented to this user.

The data transfer method for the single experiment is the same as for the collaborative experiment, with the only difference being that each group uses one less Oculus Quest 2 and one less laptop to assist Oculus with rendering and relaying the data.

IV. Experiment

A. User arrangement

A total of 12 participants (4 females and 8 males, aged between 20-23, M = 21.5) were invited to join this study and divided into 8 groups according to the between-subjects experiment design. 4 pairs (8 participants) randomly formed were assigned to the two-person collaborative group. The other four participants were assigned to the single-user group. Data collected from the pre-experimental questionnaire showed that all participants had no previous experience remotely controlling a UGV via a VR interface. They did not report any physical symptoms and had normal or corrected-to-normal vision. Prior to the actual experiment, they were given a pre-training practice to allow them to become familiar with the VR device, controls, and remote manipulation. All participants successfully completed the pre-training.

B. Devices Used

For this study, we have used the RoboMaster EP robot [1]. The body of this robot can be divided into a vehicle body, a robotic arm, a camera, and a WiFi module (as shown in Fig. 2a). Participants were able to observe the field of view from the robot's camera through the Oculus Quest 2 and use the Oculus Touch controller (see Fig. 2b[2].) to operate the robot's body movement and control the robotic arm.

Participants could control the UGV'S translation and rotation using the left handle (see Fig. 2b). Participants could also use the right handle to control the robotic arms and jaws (see Fig. 2b).

C. Task and Procedure

We asked participants to move the blocks (each measuring 4cm×4cm×9cm) from one place to another using the RoboMaster EP (see Fig. 2c). This task of moving the blocks was designed to allow the participants to control the robot's arm and move it simultaneously. As shown in a pilot study, the dual nature of the task was challenging enough to elicit a desired level of cognitive load.

Both single-person and two-person groups had to wear the Oculus Rift 2 and use the VR remote control platform we implemented and run in Unity to control the UGV remotely. In the single-person group, participants were required to operate the movement of the UGV and control the robotic arm. In the two-person collaborative group, one participant needed to control the movement of the UGV and the other participant had to control the robotic arm. During the experiment, the UGV and the participants were in two different rooms to realistically simulate the remote control of the UGV (see Fig. 2d). Although the participants were seated in the same room and had the opportunity to orally communicate directly, our purpose was to provide them with visual cues in the collaborative immersive environments to initiate collaboration. Thus, our experiment paid attention to the designed tasks considering direct communication as a default condition.

A simple pre-training was given to participants to complete before the formal experiment. The purpose of this pre-training was to enable participants to become familiar with remote control operation of the UGV in a VR environment, and also to enable participants in the two-person group to become familiar with collaborative remote control operations. In the pre-training participants had to try to pick up a wooden block using the UGV and move it to a designated place. Immediately after the pre-training, participants started the formal experiment.

[1]https://www.dji.com/hk/robomaster-ep

[2]https://docs.unity3d.com/560/Documentation/Manual/OculusControllers.html

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

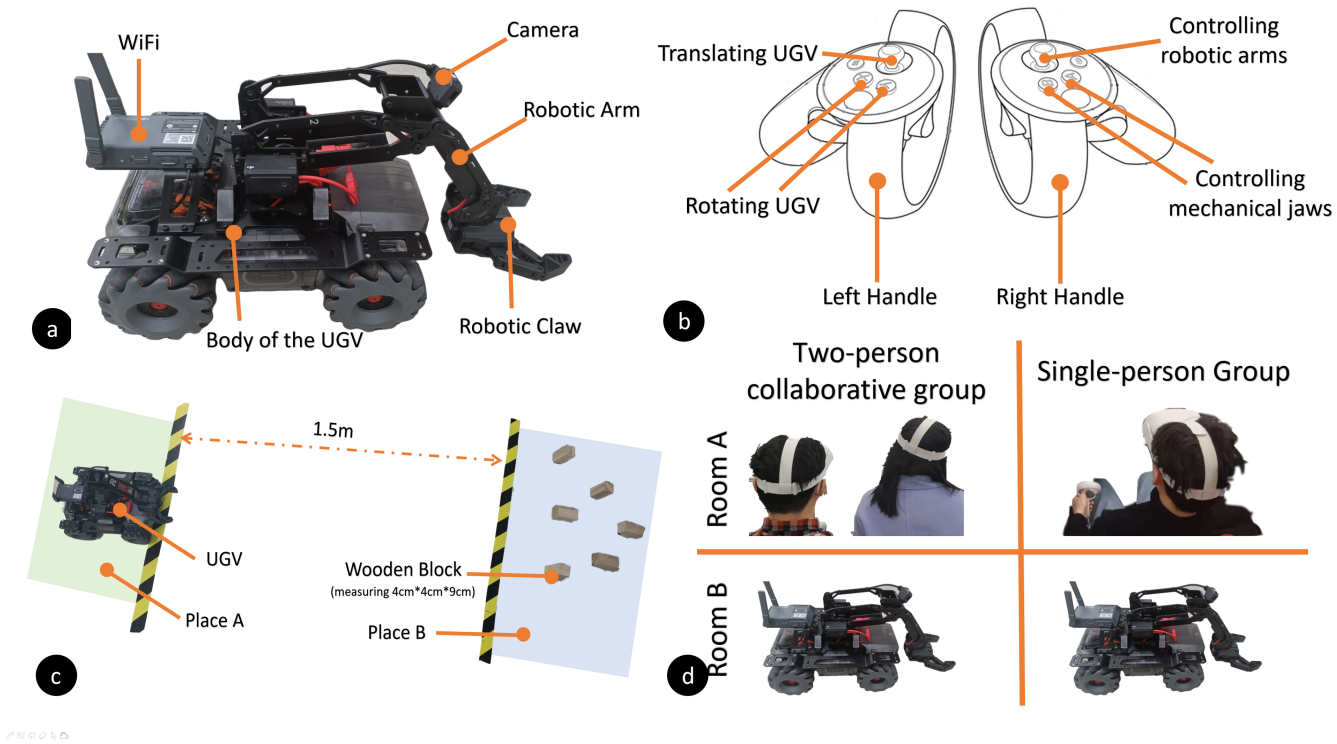

Fig. 2. (a) Structure of the UGV used in the study; (b) Description of how participants control the UGV using the Oculus Touch controllers; (c) A picture of the experimental task; and (d) The experiment environment design.

In the formal experiment, participants were given the following five tasks to complete: (1) Pick up wooden blocks (6 in total) from one side of a line (Place B); (2) Rotate the UGV and orient it towards the other side of the line (Place A); (3) Move the UGV to Place A and put the block down; (4) If there were still blocks in Place B, move the UGV to Place B and pick up these blocks; (5) Repeat tasks 1 to 4 until there are no more blocks in Place B. Upon completion of the experiment, participants were asked to complete the NASA-TLX questionnaire to measure their cognitive load during the experiment. Participants also needed to complete a Semi-Structured user experience questionnaire (UEQ) (7-point scoring scale) to allow us collect participants' ratings of the operational difficulty of the remote control approach used in the experiment.

D. Hypotheses

To answer our research question, we formulated the following four hypotheses. We focus on two particular points to investigate how articulation work is supported in VR-based collaboration: (1) the VR interface itself, and (2) social interaction in collaborative interaction in the VR environments, such as awareness and communication. Additionally, those artifacts which were used to coordinate cooperative interaction can also be a vital factor to enable the collaborative remote control of unmanned vehicles.

- H_1 : with the help of the other participant, participants in the collaborative group would take less time to complete the tasks than those in the single-user group.
- H_2 : based on Brook's Law, participants in the two-person group may use more actions to complete a task than participants in the single-person group because of the higher level of discussion and actions needed during the collaboration process (that is, participants in this group could make more missteps).
- H_3 : with the assistance of another participant, participants in the two-person group would experience a lower cognitive load in completing the experiment.
- H_4 : compared to the single-person group, the two-person group would have a lower level of difficulty to learn the operation of control UGV.
- H_5 : compared to the single-person group, the two-person group would have a lower level of difficulty to complete the given task.

V. Results & Analysis

All participants in this experiment understood and completed the given experimental tasks successfully. The data we collected were valid. For the objective data we collected during the experiment, the results of the Shapiro-Wilk test show that our data were normally distributed ($p > .05$). Results of Levene's test show that the assumption of homogeneity of variance was met in each condition ($p > .05$). We conducted independent t-tests for the comparison of objective measurements. Similar to the objective data, the subjective data also return the same results from Levene's test ($p > .05$).

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

Thus, we conducted the Mann-Whitney U test for the pairwise comparison of subjective measurements.

A. Objective Results

The time spent by participants to complete the task was obtained directly from the length of the video recorded during the experiment. The number of operations performed by the participants to complete the task was counted directly by the Unity application. As many of the commands issued by participants during the remote control process, such as UGV forward and backward, were continuous commands, the application treated and counted the commands as a state. It counted each time the participant switched between states.

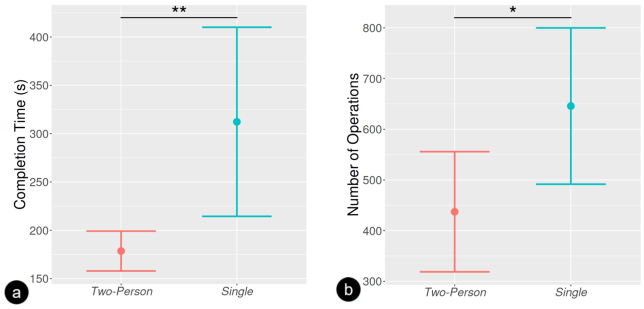

Fig. 3. (a) Mean completion time(s) of the task, and (b) mean number of operations. The error bars represent 95% confidence intervals. '**', '*' represent a '.01' and '.05' significance levels, respectively. Two-person: Two-person collaborative group; Single: Single-person group

Independent t-test shows that the mean completion time (s) of the task differed significantly between two groups ($t(6) = -4.256, p < .01$). Fig. 3a shows the mean completion time of the two-person collaborative group and single-person group. It shows that the two-person group took much less time to complete the experimental task than the single-person group, which supports H_1.

The two-person collaborative group took less time to complete tasks than the single-person group for two reasons. One is because each participant in the two-person group needed to control part of the UGV, and therefore this control was less difficult for these participants than for the single-person group. The other reason is that participants in the two-person group had less information to focus on during task completion, i.e. the cognitive load was less. Both reasons were confirmed in the subsequent analysis of the subjective data from the questionnaire.

Results of t-tests on the mean number of operations also shows a significantly difference between two groups ($t(6) = -3.412, p < .05$). Fig. 3b shows the mean number of operations of each group. It shows that the two-person group used significantly fewer operations to complete the experimental task, which does not support H_2, to our surprise. From our post-experiment interviews with participants, we found the main reason why the two-person group completed fewer total actions than the single-user group. Although there were some incorrect operations due to disagreements between some pairs in the two-person group, there were also incorrect actions made by single users due to cognitive overload. As such, we can infer that more errors were made in the single group for a different reason than in the two-person group, resulting in H_2 not being supported.

B. Subjective Results

The results of the Mann-Whitney U test for NASA-TLX questionnaire data show that the cognitive load of the two-person group was significantly lower than the single-user group ($p < .05$). This result supports H_3. As for the UEQ data's testing, the test results for UEQ data did not show significant differences in the learning difficulty between the two groups ($p > .05$), which does not support H_4. However, the test results for UEQ data showed that there were significant differences in perceived difficulty of operating the UGV between the two groups ($U = 2.5, p = .019.05$). The results of the descriptive analysis of the UEQ data show that participants in the two-person group perceived the operational difficulty of the remote control of the UGV using the VR interface to be low, whereas participants in the one-person group perceived the difficulty to be medium to high. This result supports H_5. This suggests that the use of two-person control significantly reduces the operational difficulty when controlling a UGV remotely using VR.

Based on these results, we found that two-person collaboration significantly helped participants to reduce the operational difficulty. It also significantly reduced the cognitive load of participants when performing the task.

VI. Discussion

From the above analysis, we can see that, when compared with the single group, the collaborative group took significantly less time to complete the experimental task (supports H_1); the total number of operations to complete the experimental task was lower (supports H_2); the cognitive load was significantly less (supports H_3); there was no significant difference in the difficulty perceived by the volunteers when learning to operate the UGV (supports H_4), and the difficulty perceived by the volunteers when operating the UGV during the experiment was significantly lower (supports H_5). These outcomes are attributable not only to the collaborative versus single group dichotomy but also to additional characteristics of the structure of the UGV remote control system that we designed in this project.

In the framework of our system, participants in collaborative groups can communicate their task goals directly with their partners without delay. The establishment of H_1 and the non-establishment of H_2 is strongly related to this feature of the experimental system. This is because participants in the collaborative group can confirm the target with their partner easily and quickly via direct communication, so participants in the collaborative group expended little time and manipulation in confirming the common target. We also realized that the experiment's sample size was relatively small, and the population might not be representative. These are biases of the

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

investigation. Due to Covid-19, it was hard to recruit a significant sample size with a wide age range. If more participants in the collaborative group and their partners were not in the same room when the experiment was conducted or could not communicate with each other, the results of the experiment for H_1 and H_2 might have been judged differently. However, our designed system fruitfully confirmed the articulation work of teleoperators' remote control of unmanned ground vehicles. This articulation of collaboration is essential, and it opens room for other researchers to explore systems development for collaboration in immersive environments.

To generate additional supporting evidence, the design of the experimental system for remote control operation separates the operation of the different functional areas of the UGV into different handles. This feature is manifested in the fact that the control of the movement of the vehicle is executed entirely by the left joystick and the control of the robotic arm is executed entirely by the right joystick. This design reduces the learning difficulty for participants and leads to the non-establishment of H_4. This design also makes it less likely that participants will confuse the operation when they only need to control one function, which validates H_3 and H_5.

Finally, the collaborative characteristics of the designed system are also supported by the final results of the experiment. The participants invited into this experiment had never had any experience of using a joystick to remotely control a UGV. These participants also did not have experience with remote control of UGVs. The experiment results show that our design helps teleoperators access key information quickly and accurately. In turn, their cognitive loads were reduced when using easy user interfaces.

VII. Conclusion

This study built a remote control platform for UGVs in a VR environment that supports two-person collaboration. We then used this platform to evaluate how users' perceptions, work states, and cognitive load differ when remotely controlling UGVs in this platform compared to traditional single-person remote control UGVs. We illustrate the scope of effectiveness and conditions of applicability of collaborative control in UGV remote control in VR environments. Our system shed light on the importance of awareness of multi-person collaboration and the articulation work in cooperative remote control of UGVs in VR environments. We assert that using VR as an interface for supporting remote control shall enhance the collaboration among multi-users. This contribution opens a pathway to focus on identifying and investigating those cues, tools, and coordination mechanisms to maintain the flow of work in VR environments. VR-supported cooperative systems, with such focal points as remote control, multiple teleoperators, and cognitive workload are fruitful areas for further study. Collaborative work could be greatly enhanced with further development of collaborative remote control in the VR field.

Acknowledgement

We thank the participants who generously shared their time to do the experiment. This work was supported in part by Xi'an Jiaotong-Liverpool University (XJTLU) Key Program Special Fund (#KSF-A-03) and XJTLU Research Development Fund (#RDF-17-01-54, #RDF-21-02-008, and #RDF-19-02-11).

References

[1] E. Björling, A. Kim, K. Oleson, and P. Alves-Oliveira, "I am the robot: Teen collaboration in an asymmetric," *Virtual Reality Game. Front. Virtual Real. 2: 746521. doi: 10.3389/frvir*, 2022.

[2] A. Cruz, H. Paredes, L. Morgado, and P. Martins, "Non-verbal aspects of collaboration in virtual worlds: a cscw taxonomy-development proposal integrating the presence dimension," *JUCS-Journal of Universal Computer Science*, vol. 27, no. 9, pp. 913–954, 2021.

[3] L. Chen, H.-N. Liang, F. Lu, J. Wang, W. Chen, and Y. Yue, "Effect of collaboration mode and position arrangement on immersive analytics tasks in virtual reality: A pilot study," *Applied Sciences*, vol. 11, no. 21, p. 10473, 2021.

[4] H.-N. Liang, F. Lu, Y. Shi, V. Nanjappan, and K. Papangelis, "Evaluating the effects of collaboration and competition in navigation tasks and spatial knowledge acquisition within virtual reality environments," *Future Generation Computer Systems*, vol. 95, pp. 855–866, 2019.

[5] L. Chen, H.-N. Liang, F. Lu, K. Papangelis, K. L. Man, and Y. Yue, "Collaborative behavior, performance and engagement with visual analytics tasks using mobile devices," *Human-Centric Computing and Information Sciences*, vol. 10, no. 1, pp. 1–24, 2020.

[6] M. L. Kovarik, L. S. Ott, J. K. Robinson, and T. J. Wenzel, "Getting started on active learning," in *Active Learning in the Analytical Chemistry Curriculum*. ACS Publications, 2022, pp. 13–35.

[7] J. Han and B. Smith, "Cu-seeme vr immersive desktop teleconferencing," in *Proceedings of the Fourth ACM international Conference on Multimedia*, 1997, pp. 199–207.

[8] C. Nguyen, S. DiVerdi, A. Hertzmann, and F. Liu, "Collavr: collaborative in-headset review for vr video," in *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 2017, pp. 267–277.

[9] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. H. Thomas, "Immersive collaborative analysis of network connectivity: Cave-style or head-mounted display?" *IEEE transactions on visualization and computer graphics*, vol. 23, no. 1, pp. 441–450, 2016.

[10] M. R. Morris, J. Lombardo, and D. Wigdor, "Wesearch: supporting collaborative search and sensemaking on a tabletop display," in *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 2010, pp. 401–410.

[11] P. Bjørn, M. Wulff, M. S. Petræus, and N. H. Møller, "Immersive cooperative work environments (cwe): Designing human-building interaction in virtual reality," *Computer Supported Cooperative Work (CSCW)*, vol. 30, no. 3, pp. 351–391, 2021.

[12] K. Chow, C. Coyiuto, C. Nguyen, and D. Yoon, "Challenges and design considerations for multimodal asynchronous collaboration in vr," *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–24, 2019.

[13] Y. Pan, "Reflexivity of account, professional vision, and computer-supported cooperative work: Working in the maritime domain," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–32, 2021.

[14] M. Vasarainen, S. Paavola, L. Vetoshkina *et al.*, "A systematic literature review on extended reality: Virtual, augmented and mixed reality in working life," *Internationa Journal of Virtual Reality*, 2021.

[15] T. Ludwig, O. Stickel, P. Tolmie, and M. Sellmer, "share-it: Ad hoc remote troubleshooting through augmented reality," *Computer Supported Cooperative Work (CSCW)*, vol. 30, no. 1, pp. 119–167, 2021.

[16] Y. Luo, J. Wang, R. Shi, H.-N. Liang, and S. Luo, "In-device feedback in immersive head-mounted displays for distance perception during teleoperation of unmanned ground vehicles," *IEEE Transactions on Haptics*, vol. 15, no. 1, pp. 79–84, 2022.

[17] Y. Luo, J. Wang, H.-N. Liang, S. Luo, and E. G. Lim, "Monoscopic vs. stereoscopic views and display types in the teleoperation of unmanned ground vehicles for object avoidance," in *2021 30th IEEE International Conference on Robot Human Interactive Communication (RO-MAN)*, 2021, pp. 418–425.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.

[18] M. Bout, A. P. Brenden, M. Klingegård, A. Habibovic, and M.-P. Böckle, "A head-mounted display to support teleoperations of shared automated vehicles," in *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, 2017, pp. 62–66.

[19] S. Islam, Q. Huang, F. Afghah, P. Fule, and A. Razi, "Fire frontline monitoring by enabling uav-based virtual reality with adaptive imaging rate," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2019, pp. 368–372.

[20] D. Goedicke, J. Li, V. Evers, and W. Ju, "Vr-oom: Virtual reality on-road driving simulation," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–11.

[21] V. T. Nguyen, K. Jung, and T. Dang, "Dronevr: A web virtual reality simulator for drone operator," in *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 2019, pp. 257–2575.

[22] S. Helsel, "Virtual reality and education," *Educational Technology*, vol. 32, no. 5, pp. 38–42, 1992.

[23] S. Fakhoury, Y. Ma, V. Arnaoudova, and O. Adesope, "The effect of poor source code lexicon and readability on developers' cognitive load," in *2018 IEEE/ACM 26th International Conference on Program Comprehension (ICPC)*. IEEE, 2018, pp. 286–28 610.

[24] K. Sexton, A. Johnson, A. Gotsch, A. A. Hussein, L. Cavuoto, and K. A. Guru, "Anticipation, teamwork and cognitive load: chasing efficiency during robot-assisted surgery," *BMJ quality & safety*, vol. 27, no. 2, pp. 148–154, 2018.

[25] G. Murphy and C. M. Greene, "Load theory behind the wheel; perceptual and cognitive load effects." *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, vol. 71, no. 3, p. 191, 2017.

[26] H. Nguyen and T. Bednarz, "User experience in collaborative extended reality: overview study," in *International Conference on Virtual Reality and Augmented Reality*. Springer, 2020, pp. 41–70.

[27] Y. Lee and B. Yoo, "Xr collaboration beyond virtual reality: work in the real world," *Journal of Computational Design and Engineering*, vol. 8, no. 2, pp. 756–772, 2021.

[28] T. Leung, F. Zulkernine, and H. Isah, "The use of virtual reality in enhancing interdisciplinary research and education," *arXiv preprint arXiv:1809.08585*, 2018.

[29] F. Lu, D. Yu, H.-N. Liang, W. Chen, K. Papangelis, and N. M. Ali, "Evaluating engagement level and analytical support of interactive visualizations in virtual reality environments," in *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2018, pp. 143–152.

[30] A. Pack, A. Barrett, H.-N. Liang, and D. Monteiro, "University eap students' perceptions of using a prototype virtual reality learning environment to learn writing structure," *International Journal of Computer-Assisted Language Learning and Teaching*, vol. 10, no. 1, pp. 27–46, 2020.

[31] A. Jimeno-Morenilla, J. L. Sánchez-Romero, H. Mora-Mora, and R. Coll-Miralles, "Using virtual reality for industrial design learning: a methodological proposal," *Behaviour & Information Technology*, vol. 35, no. 11, pp. 897–906, 2016.

[32] H. Lee, H. Kim, D. V. Monteiro, Y. Goh, D. Han, H.-N. Liang, H. S. Yang, and J. Jung, "Annotation vs. virtual tutor: Comparative analysis on the effectiveness of visual instructions in immersive virtual reality," in *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2019, pp. 318–327.

[33] D. Norman, *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[34] C. Heath, M. S. Svensson, J. Hindmarsh, P. Luff, and D. Vom Lehn, "Configuring awareness," *Computer Supported Cooperative Work (CSCW)*, vol. 11, no. 3, pp. 317–347, 2002.

[35] D. Randall, M. Rouncefield, and P. Tolmie, "Ethnography, cscw and ethnomethodology," *Computer Supported Cooperative Work (CSCW)*, vol. 30, no. 2, pp. 189–214, 2021.

[36] H. H. Clark and S. E. Brennan, "Grounding in communication." 1991.

[37] J. Sweller, J. J. van Merriënboer, and F. Paas, "Cognitive architecture and instructional design: 20 years later," *Educational Psychology Review*, vol. 31, no. 2, pp. 261–292, 2019.

[38] S. Benford, C. Brown, G. Reynard, and C. Greenhalgh, "Shared spaces: Transportation, artificiality, and spatiality," in *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, 1996, pp. 77–86.

[39] K. Schmidt, "Taking cscw seriously: Supporting articulation work (1992)," in *Cooperative Work and Coordinative Practices*. Springer, 2008, pp. 45–71.

[40] D. Randall, "What is common in accounts of common ground?" *Computer Supported Cooperative Work (CSCW)*, vol. 25, no. 4, pp. 409–423, 2016.

[41] L. Chen, Y. Liu, Y. Li, L. Yu, B. Gao, M. Caon, Y. Yue, and H.-N. Liang, "Effect of visual cues on pointing tasks in co-located augmented reality collaboration," ser. SUI '21. New York, NY, USA: Association for Computing Machinery, 2021.

[42] B. Brown and M. Bell, "Cscw at play: 'there'as a collaborative virtual environment," in *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, 2004, pp. 350–359.

[43] G. Pedersen and K. Koumaditis, "Virtual reality (vr) in the computer supported cooperative work (cscw) domain: A mapping and a pre-study on functionality and immersion," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 136–153.

[44] M. Begum, R. W. Serna, D. Kontak, J. Allspaw, J. Kuczynski, H. A. Yanco, and J. Suarez, "Measuring the efficacy of robots in autism therapy: How informative are standard hri metrics'," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 2015, pp. 335–342.

[45] Z. Sun and M. Li, "Design of a training system for special types of mine workers based on cscw," in *2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design ((CSCWD))*. IEEE, 2018, pp. 501–506.

[46] S. Morrison-Smith and J. Ruiz, "Challenges and barriers in virtual teams: a literature review," *SN Applied Sciences*, vol. 2, no. 6, pp. 1–33, 2020.

[47] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition: toward a new foundation for human-computer interaction research," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 7, no. 2, pp. 174–196, 2000.

[48] R. Shi, H.-N. Liang, Y. Wu, D. Yu, and W. Xu, "Virtual reality sickness mitigation methods: A comparative study in a racing game," *Proc. ACM Comput. Graph. Interact. Tech.*, vol. 4, no. 1, apr 2021.

[49] D. Yu, H.-N. Liang, X. Lu, K. Fan, and B. Ens, "Modeling endpoint distribution of pointing selection tasks in virtual reality environments," *ACM Trans. Graph.*, vol. 38, no. 6, nov 2019.

[50] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves, "Gaze-supported 3d object manipulation in virtual reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021.

[51] P. G. Roetzel, "Information overload in the information age: a review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development," *Business research*, vol. 12, no. 2, pp. 479–522, 2019.

[52] P. Ramakrishnan, B. Balasingam, and F. Biondi, "Cognitive load estimation for adaptive human–machine system automation," in *Learning Control*. Elsevier, 2021, pp. 35–58.

[53] J. Stapel, F. A. Mullakkal-Babu, and R. Happee, "Automated driving reduces perceived workload, but monitoring causes higher cognitive load than manual driving," *Transportation research part F: traffic psychology and behaviour*, vol. 60, pp. 590–605, 2019.

[54] J. Costley and C. Lange, "The moderating effects of group work on the relationship between motivation and cognitive load," *The International Review of Research in Open and Distributed Learning*, vol. 19, no. 1, 2018.

[55] H.-C. Chu, J.-M. Chen, and C.-L. Tsai, "Effects of an online formative peer-tutoring approach on students' learning behaviors, performance and cognitive load in mathematics," *Interactive Learning Environments*, vol. 25, no. 2, pp. 203–219, 2017.

[56] M. Baker, "Statisticians issue warning over misuse of p values," *Nature News*, vol. 531, no. 7593, p. 151, 2016.

[57] A. Gelman, "The problems with p-values are not just with p-values," *The American Statistician*, vol. 70, no. 10, 2016.

[58] K. Hornbæk, "Some whys and hows of experiments in human–computer interaction," *Foundations and Trends in Human-Computer Interaction*, vol. 5, no. 4, pp. 299–373, 2013.

Authorized licensed use limited to: Xi'an Jiaotong-Liverpool University. Downloaded on August 04,2026 at 08:51:54 UTC from IEEE Xplore. Restrictions apply.