Assessment of Instructor's Capacity in One-to-Many AR Remote Instruction Giving

Mai Otsuki
mai.otsuki@aist.go.jp
AIST
Japan

Tzu-Yang Wang
tywang@jaist.ac.jp
JAIST
Japan

Hideaki Kuzuoka
kuzuoka@cyber.t.u-tokyo.ac.jp
The University of Tokyo
Japan

ABSTRACT

In this study, we focused on one-to-many remote collaboration which requires more mental resources from the remote instructor than the case of one-to-one since it is "multitasking". The main contribution of our study is that we assessed instructor's capacity in one-to-many AR remote instruction giving both subjectively and objectively. We compared the remote instructor's workload while interacting with a different number of local workers, assuming tasks at an industrial site. The results showed that the instructors perceived stronger workload and the communication quality became lower when interacting with multiple local workers. Based on the results, we discussed how to support the remote instructor in a one-to-many AR remote collaboration.

CCS CONCEPTS

• **Human-centered computing** → **Computer supported cooperative work**.

ACM Reference Format:
Mai Otsuki, Tzu-Yang Wang, and Hideaki Kuzuoka. 2022. Assessment of Instructor's Capacity in One-to-Many AR Remote Instruction Giving. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3562939.3565631

1 INTRODUCTION

Remote collaboration has been an important research topic in CSCW and researchers have proposed various AR technologies to support remote skilled workers or instructors to assist local workers [1, 5, 30, 32]. However, most of the past studies have focused on one-to-one remote collaboration [18, 34]. In recent years, owing to the shortage of skilled workers, occasions that one skilled worker needs to assist multiple workers are increasing. Therefore, one-to-many remote collaboration is attracting researchers' attention. For example, Lee et al. proposed and evaluated various view-sharing AR techniques to support such collaboration [21].

One of the possible advantages of one-to-many collaboration is that, if a remote instructor can observe multiple local workers simultaneously, it will be easy for the instructor to find a worker who is in need of assistance and instantly take care of him/her. Furthermore, even when the instructor is dealing with a specific worker, the instructor can observe other workers in his peripheral view and instantly switch to another worker for help. However, such collaboration requires more mental resources from the remote instructor than the case of one-to-one, because it is "multitasking". Multitasking hurts work efficiency and task management [24, 35]. However, unlike face-to-face collaboration, remote collaboration involves devices, such as mobile/wearable devices between an instructor and a worker; thus, the instructor's workload can be reduced and also work efficiency can be improved by providing the appropriate information via devices.

We aim to develop an AR system that supports instructors and improves work efficiency in one-to-many remote collaboration (Figure 1). As the initial stage of our research, we assessed instructor's capacity in one-to-many AR remote instruction giving both subjectively and objectively. We investigated how the workload is affected depending on the number of workers when a remote instructor observes and instructs multiple local workers in parallel, assuming tabletop assembly, repair, and maintenance tasks at an industrial site. Based on the results, we discussed how to support the remote instructor in a one-to-many AR remote collaboration.

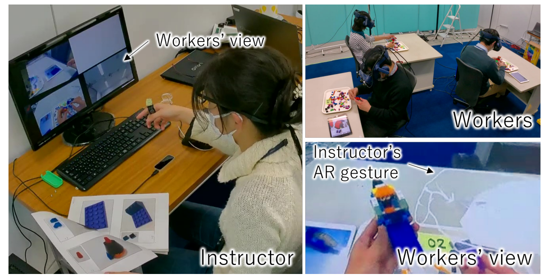

Figure 1: One-to-many AR remote collaboration system used in this study.

2 RELATED WORK

2.1 AR Remote Collaboration

2.1.1 One-to-One AR Remote Collaboration. A typical usage scenario of the AR remote collaboration system is that a remote expert instructs local unskilled workers [10, 27, 28]. By using AR, participants can share richer information such as non-verbal cues, compared to existing instructional technologies. The target of this study is collaborative physical tasks that the participants act on physical objects, such as assembly, repair, and maintenance.

There are some examples where an instructor instructs a local worker to search for and place physical objects from a remote location. Studies have evaluated the effects of sharing the eye gaze

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565631

and/or hand gestures between the instructor and worker [2, 22], and the perspectives from which remote instructors prefer to observe the worker's view [9]. Some systems can provide AR instructions onto physical objects on using small projectors [13], or through a tablet [4, 11]. These studies discussed on what to share between the instructor and the worker, in what form, and with what device. Additionally, these systems are designed to work in a large space while walking around.

Various AR remote collaboration systems for tabletop tasks have also been proposed, specifically for assembly tasks. For example, visualizing and sharing the worker's attention [8], sharing the hand gestures of a remote instructor with a worker [19, 20], and sharing gestures and facial expressions [16]. In line with these studies, in this study, we developed a system that enables an instructor to instruct a worker using voice and gestures for tasks on a tabletop, and use it as a testbed.

2.1.2 One-to-Many AR Remote Collaboration. Due to the recent shortage of skilled workers, one skilled worker needs to support multiple workers in parallel. Schott et al. developed an AR collaborative learning system that allows VR and AR users to interact with each other [29]. However, unlike our study, their system does not deal with physical tasks. Lee et al. proposed various viewpoint-sharing AR techniques to support one-to-many collaboration in physical object exploration and placement tasks and evaluated their effectiveness [21]. Norman et al. investigated how role assignment affects group task coordination and engagement during a furniture placement task in mixed reality [25] and found that when a remote participant took on the role of coordinator for two local workers, the workload was significantly higher than when all the participants took on equivalent roles and discussed the task together.

These studies evaluated the task completion time and workload in a one-to-two configuration; however, they did not discuss how the measures are affected by the number of local workers.

2.2 Multitasking Load

The one-to-many AR remote collaboration to support tasks on a tabletop, which is the target of this study, is "multitasking"; a remote skilled worker must give appropriate instructions to each worker whenever they need help (primary task), and simultaneously monitor each worker and understand the situation (secondary task). It is known that multitasking causes a significant decrease in work efficiency when switching tasks [24], and task management errors occur when the mental workload during multitasking is high [35].

Fan et al. measured the workload on participants who were tasked with monitoring multiple meters and pressing a key when they deviated from the normal range [6]. When the number of monitored objects was changed from 2, 4, and 6, the subjective workload was significantly higher, the reaction time increased, and the accuracy of finding abnormal values decreased. However, a study by compared the physiological and subjective mental workload of radiation therapists when the number of monitors changed 2 or 3, and no significant differences were identified [26].

When it comes to one person monitoring multiple workers at the same time, the case of one teacher teaching programming simultaneously to multiple students is also relevant. Systems that visualize mutual eye contact on the screen [36] or provide support through chat [12] have been proposed. However, the relationship between the number of people and workload is not mentioned. Parallel eyes [17] and Parallel Ping-Pong [31] are systems that allow multiple people to share a viewpoint using HMDs. They conducted workshops that used these systems and found that the systems yielded high cognitive workload to the participants, but there was no quantitative evaluation.

In this study, the workload will be higher than that of previous studies because the instructors not only monitor the situation but also provide instructions corresponding to the situation, for a 3D assembly task that requires spatial comprehension. To the best of our knowledge, there are no studies that support this, and thus the main contribution of our study is clarifying this quantitatively.

3 EXPERIMENT

3.1 Hypothesis

Although there are different types of one-to-many collaboration, in this study, we focused on the situation that multiple local workers work on an identical task but at different sites. In such a situation, the workers follow the pre-defined work process. However, when they encounter an unknown step, they ask for instructions from a remote instructor (skilled worker). As described in Section 2, this is a multitasking for a remote instructor. Therefore, in this situation, we hypothesized that the instructor's workload is affected depending on the number of workers; as the number of workers increases, the instructor (participant) further experiences a higher workload while interacting with them.

3.2 Experimental Design

To examine the hypothesis, we designed a one-factor within-participant experiment in which the number of workers (1-3) was the factor. In this experiment, we selected a Lego assembly task was selected as the task since it contains similar motion elements as assembly, repair, and maintenance work, and it is often used as a task to simulate these tasks in remote collaboration studies. [2, 7, 15, 18].

3.3 System Configuration

We developed a one-to-many AR remote collaboration system (Figure 1). The local worker observes the local environment using an HMD (HTC VIVE Pro) with a stereo camera (ZED mini). The remote instructor can observe the view of up to four workers simultaneously through an LCD monitor (BenQ gw2255, 21.5-inch). The screen size of each worker's view is about 10-inch regardless of the number of workers. In addition to speech, hand gestures of the remote instructor are obtained using Leap motion. The remote instructor can click the view of a target worker who needs help, then the instructor's gestures are superimposed on the selected local worker's view using AR, and shown to both the instructor and the target worker.

3.4 Experimental Procedure and Participants

The participant played the instructor's role sat in front of the LCD monitor and participated in a total of three sessions of the experiment. In each session, the participant conducted one Lego assembly task with workers. The participant received a complete version of the manual and gave instructions as requested by the workers. To simulate a situation in which the workers needed assistance from the instructor, 15% of the steps of the workers' manuals were masked, and the masked steps were randomly and evenly distributed throughout the whole procedures. Besides, the masked steps are different between workers. The workers performed assembly work independently based on the manuals step by step, and when they encountered the masked procedure, they paused the task and asked the instructor for instructions.

Because the purpose of this experiment was to assess instructor's capacity, we decided to employ three confederates as workers and trained them beforehand so that they can perform a series of tasks at a constant speed without delay and ask the instructor for assistance in the same manner. The confederates participated in all experiments and all sessions to minimize the bias caused by the differences among the individual workers.

We limited the duration of one session to 25 minutes, and each task was designed to require sufficiently longer than 25 minutes. After finishing each session, the participants completed two questionnaires (see the "Measures" section). Then, the participants rested for 10 minutes before starting the next session. After all the three sessions, semi-structured interviews were conducted with the participants to discuss the differences in the number of participants, the difficulties in giving instructions, and the usability of the system.

The number of workers and the type of Lego models were all different in the three sessions and were counterbalanced. Eleven participants took part in the experiment. Six were male and five were female, and the mean age was 26.27 years (SE=4.9). All participants were native Japanese speakers and each participant received 5,000 yen as remuneration.

3.5 Measures

3.5.1 Subjective measures. We asked the participants to respond to the NASA-TLX to measure the workload [14] and QCE to measure the quality of the communication experienced [23]. Specifically, QCE measures the three dimensions of communication (Clarity, Responsiveness, and Comfort). Each dimension includes five items on a 7-point Likert scale.

3.5.2 Objective measures. We conducted video analysis to analyze the instructor and workers behavior. In this task, the instructor (participant) gave instructions each time the worker reached the masked part of the manual. We defined "instruction time" as the time between the moment the instructor switches the target worker and the moment when s/he confirmed the completion of a step (e.g. by saying O.K.) at the end of the instruction. When there were more than two workers, there were cases where workers simultaneously asked for instructions, and the other workers had to wait for instructions while one worker received instructions. We defined this as "idle state". The time which the worker is in the idle state is referred to as the "idle time". To generalize the discussion, the instructor's "utilization rate" was calculated by dividing the total instruction time by the session time (25 min). We also counted the number of occurrences of both instruction and idle state.

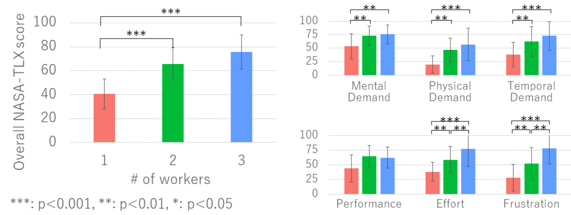

Figure 2: NASA-TLX results

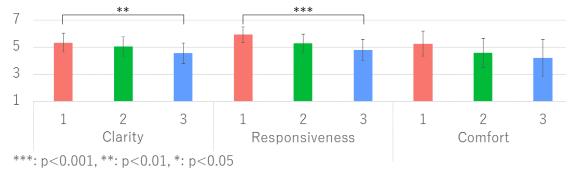

Figure 3: QCE results

4 RESULTS

4.1 Subjective Evaluation

The NASA-TLX results are shown in Figure 2. A linear mixed model was constructed for the analysis; a Type III analysis of the variance table using Satterthwaite's method revealed significant differences in NASA-TLX scores for different numbers of workers ($F(2,20) = 25.17$, $p < .001$). Post hoc tests with the Bonferroni correction showed that the scores in the 1-worker condition were significantly lower than those in the 2- and 3-worker condition ($t(20) = 4.92$, $p < .001$, $t(20) = 6.89$, $p < .001$, respectively).

In addition, a linear mixed model was constructed to analyze the result of each subscale of the NASA-TLX. The results showed significant differences in all subscale scores across conditions (mental demand: $F(2,20) = 6.30$, $p = .007$; physical demand: $F(2, 20) = 12.77$, $p < .001$; temporal demand: $F(2, 20) = 13.66$, $p < .001$; performance: $F(2, 20) = 3.89$, $p = .04$; effort: $F(2,20) = 27.08$, $p < .001$; frustration: $F(2,20) = 26.92$, $p < .001$). Post hoc tests with the Bonferroni correction indicated that subscale scores other than Performance were significantly lower in the 1-worker condition than the same subscale scores in the 2- and 3-worker conditions. Effort and frustration were also significantly lower in the 2-worker condition than in the 3-worker condition.

The results of the QCE are shown in Figure 3. The results of linear mixed models showed significant differences between conditions for clarity and responsiveness factors (clarity: $F(2,20) = 6.06$, $p = .008$; responsiveness: $F(2,20) = 9.62$, $p =.001$). Scores for the 1-worker condition were also significantly higher than the scores for the 3-worker condition (clarity: $t(20) = 3.43$ $p = .008$; responsiveness: $t(20) = 4.37$, $p < .001$).

4.2 Objective Evaluation

For significance tests, we carried out a one-way ANOVA with the number of workers as a factor. A post-hoc test for multiple comparisons using Bonferroni correction. Figure 4 shows the averages of the instruction time and idle time per instruction for each condition. ANOVA showed a significant difference in both instruction time ($F(2, 20) = 5.18$, $p < .05$, $\eta^2 = .34$) and idle time ($F(2, 20) = 69.06$, $p < .01$, $\eta^2 = .87$). Post hoc tests showed that instruction time and idle time increased significantly as the number of workers

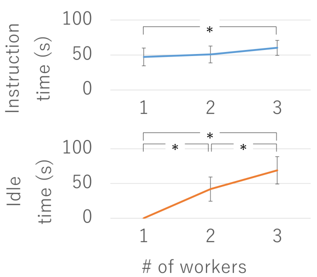

Figure 4: Average instruction time and idle time.

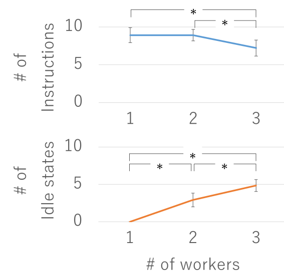

Figure 5: Average number of instructions and idle states for one worker in one session

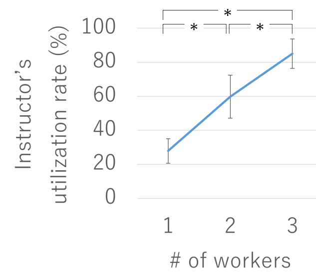

Figure 6: Instructor's utilization rate

increased (instruction time: between 1- and 3-worker condition ($p < .05$); idle time: for all cases ($p < .05$)).

To see how the work efficiency was affected when the number of workers increased, we calculated the average number of instructions and idle states per worker in one session (Figure 5). ANOVA showed a significant difference in both the number of instructions ($F(2, 20) = 11.48$, $p < .01$, $\eta^2 = .53$) and idle states ($F(2, 20) = 125.20$, $p < .01$, $\eta^2 = .93$). Post hoc tests showed that the number of instructions decreased significantly as the number of workers increased from 1- to 2-worker condition and from 1- to 3-worker condition ($p < .05$), and the number of idle states increased significantly for all cases ($p < .05$).

The utilization rate is shown in Figure 6. The utilization rate was 27.84 % in 1-worker condition, but in 3- workers condition, it was 85.07%. ANOVA showed a significant difference ($F(2, 20) = 134.72$, $p < .01$, $\eta^2 = .93$). Post hoc tests showed that instruction time increased significantly when the number of workers increased for all cases ($p < .05$).

5 DISCUSSION AND FUTURE WORK

We assessed instructor's capacity in one-to-many AR remote instruction giving. Combination of NASA-TLX results and the interview results showed that the workload of the participants (instructors) significantly increased as the number of workers increased. Specifically, "mental demand," "effort," and "physical demand" increased due to the need for attention, and "temporal demand" and "frustration" increased due to the stress of making the workers wait. The results of the QCE suggested that the stress of making workers wait decreased the instructors' understanding of communication, which might cause more conflicts during collaboration. These results support our hypothesis.

It is interesting to find that, although the result of utilization rate showed that the 3-worker condition made the instructor almost three times busier than 1-worker condition, the number of instructions per participant decreased and the instruction time per instruction increased. This indicated that the efficiency of instruction giving decreased in the 3-worker condition. The video analysis showed that while receiving a request from a worker, the instructor often also checked the current progress of the workers. The low efficiency might be because the instructors had insufficient capacity to handle such kind of multitasking well in the 3-worker condition, in such a high utilization rate. Our interview results also supported this assumption that seven of eleven participants commented that they could track the progress of one or two workers, but not three. The low efficiency further increased the worker's idle time and number of idle states.

To overcome this issue, following two points need to be solved when developing a system to support the instructor who is giving instructions to multiple workers; (1) keep track of each worker's task progress especially when the number of workers increases, and (2) reduce the time required for each instruction.

For (1), a possible approach is to support instructors in understanding the status of workers, such as presenting each worker's progress on the screen, clearly indicating who should be instructed next, , and visualizing what the worker is looking at. For (2), based on previous studies [2, 8], the instructions can be easier by sharing the instructor's and worker's gaze each other. It is also possible that providing linguistic support; speech recognition can be incorporated (e.g., if the instructor says "red", the red blocks will be highlighted), and AR labeling can be assigned to prompt lexical entrainment [3] and establish a common ground among participants at an early stage [33]. Furthermore, supporting workers through system may be another way to indirectly support instructors. One possible method is to record the instructions given to a certain worker and replay the recorded instructions if another worker has trouble with the same procedure.

For future work, we plan to add the above functions and investigate how each function contributes to the reduction of the utilization rate and workload and the improvement in work efficiency. We also need to consider introducing different technologies. As for the ways to share hand gestures, for example, a technology like MirrorTablet [20] allows an instructor to express richer expressions by sharing the real hand gestures. As for a visual display for the instructor, we may consider using 3D display, which might give better sense of working environments for the instructor. We need further experiments to understand how these technologies affect one-to-many instruction giving.

ACKNOWLEDGMENTS

This work was supported by Telecommunications Advancement Foundation.

REFERENCES

[1] Ryan Anthony J. de Belen, Huyen Nguyen, Daniel Filonik, Dennis Del Favero, and Tomasz Bednarz. 2019. A systematic review of the current state of collaborative mixed reality technologies: 2013–2018. *AIMS Electronics and Electrical Engineering* 3, 2 (2019), 181–223. https://doi.org/10.3934/ElectrEng.2019.2.181

[2] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376550

[3] Susan E Brennan. 1996. Lexical Entrainment in Spontaneous Dialog. *Proceedings, 1996 International Symposium on Spoken Dialogue, ISSD-96* (1996), 41–44.

[4] Sicheng Chen, Miao Chen, Andreas Kunz, Asim Evren Yantaç, Mathias Bergmark, Anders Sundin, and Morten Fjeld. 2013. SEMarbeta: Mobile sketch-gesture-video remote support for car drivers. In *ACM International Conference Proceeding Series*. 69–76. https://doi.org/10.1145/2459236.2459249

[5] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human Computer Studies* 131, May (2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011

[6] Xiaoli Fan, Chaoyi Zhao, Xin Zhang, Hong Luo, and Wei Zhang. 2020. Assessment of mental workload based on multi-physiological signals. *Technology and Health Care* 28, S1 (2020), S67–S80. https://doi.org/10.3233/THC-209008

[7] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive worker assistance. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16* (2016), 934–939. https://doi.org/10.1145/2971648.2971706

[8] Susan R. Fussell, Leslie D. Setlock, and Robert E. Kraut. 2003. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proceedings of the conference on Human factors in computing systems - CHI '03*. ACM Press, New York, New York, USA, 513. https://doi.org/10.1145/642611.642701

[9] Lei Gao, Huidong Bai, Mark Billinghurst, and Robert W. Lindeman. 2020. User Behaviour Analysis of Mixed Reality Remote Collaboration with a Hybrid View Interface. In *32nd Australian Conference on Human-Computer Interaction*. ACM, New York, NY, USA, 629–638. https://doi.org/10.1145/3441000.3441038

[10] Danilo Gasques, Janet G. Johnson, Tommy Sharkey, Yuanyuan Feng, Ru Wang, Zhuoqun Robin Xu, Enrique Zavala, Yifei Zhang, Wanze Xie, Xinming Zhang, Konrad Davis, Michael Yip, and Nadir Weibel. 2021. ARTEMIS: A Collaborative Mixed-Reality System for Immersive Surgical Telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–14. https://doi.org/10.1145/3411764.3445576

[11] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. World-stabilized annotations and virtual scene navigation for remote collaboration. In *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*. ACM Press, New York, New York, USA, 449–459. https://doi.org/10.1145/2642918.2647372

[12] Philip J. Guo. 2015. Codeopticon: Real-time, one-to-many human tutoring for computer programming. *UIST 2015 - Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology* (2015), 599–608. https://doi.org/10.1145/2807442.2807469

[13] Pavel Gurevich, Joel Lanir, Benjamin Cohen, and Ran Stone. 2012. TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vol. 31. ACM, New York, NY, USA, 619–622. https://doi.org/10.1145/2207676.2207763

[14] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Vol. 52. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9 arXiv:arXiv:1011.1669v3

[15] Weidong Huang, Leila Alem, Franco Tecchia, and Henry Been-Lirn Duh. 2018. Augmented 3D hands: a gesture-based mixed reality system for distributed collaboration. *Journal on Multimodal User Interfaces* 12, 2 (jun 2018), 77–89. https://doi.org/10.1007/s12193-017-0250-2

[16] Steven Johnson, Madeleine Gibson, and Bilge Mutlu. 2015. Handheld or Handsfree? Remote Collaboration via Lightweight Head-Mounted Displays and Handheld Devices. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*. ACM Press, New York, New York, USA, 1825–1836. https://doi.org/10.1145/2675133.2675176

[17] Shunichi Kasahara, Mitsuhito Ando, Kiyoshi Suganuma, and Jun Rekimoto. 2016. Parallel Eyes: Exploring Human Capability and Behaviors with Paralleled First Person View Sharing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1561–1572. https://doi.org/10.1145/2858036.2858495

[18] Seungwon Kim, Gun Lee, Mark Billinghurst, and Weidong Huang. 2020. The combination of visual communication cues in mixed reality remote collaboration. *Journal on Multimodal User Interfaces* 14, 4 (2020), 321–335. https://doi.org/10.1007/s12193-020-00335-x

[19] David Kirk, Tom Rodden, and Danaë Stanton Fraser. 2007. Turn it this way: Grounding Collaborative Action with Remote Gestures. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*. ACM Press, New York, New York, USA, 1039. https://doi.org/10.1145/1240624.1240782

[20] Khanh Duy Le, Kening Zhu, and Morten Fjeld. 2017. MirrorTablet: Exploring a low-cost mobile system for capturing unmediated hand gestures in remote collaboration. *ACM International Conference Proceeding Series* (2017), 79–89. https://doi.org/10.1145/3152832.3152838

[21] Geonsun Lee, Hyeong Yeop Kang, Jong Min Lee, and Jung Hyun Han. 2020. A User Study on View-sharing Techniques for One-to-Many Mixed Reality Collaborations. *Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020* (2020), 343–352. https://doi.org/10.1109/VR46266.2020.1581166222244

[22] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2018. A User Study on MR Remote Collaboration Using Live 360 Video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Vol. 1. IEEE, 153–164. https://doi.org/10.1109/ISMAR.2018.00051

[23] Leigh Anne Liu, Chei Hwee Chua, and Günter K. Stahl. 2010. Quality of Communication Experience: Definition, Measurement, and Implications for Intercultural Negotiations. *Journal of Applied Psychology* 95, 3 (2010), 469–487. https://doi.org/10.1037/a0019094

[24] Stephen Monsell. 2003. Task switching. *Trends in Cognitive Sciences* 7, 3 (2003), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

[25] Mitchell Norman, Gun A. Lee, Ross T. Smith, and Mark Billinghurst. 2019. The impact of remote user's role in a mixed reality mixed presence system. *Proceedings - VRCAI 2019: 17th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (2019). https://doi.org/10.1145/3359997.3365691

[26] Joseph K. Nuamah, Prithima R. Mosaly, Robert Adams, Kathik Adapa, Bhisham S. Chera, Lawrence B. Marks, and Lukasz M. Mazur. 2021. Assessment of Radiation Therapy Technologists' Workload and Situation Awareness: Monitoring 2 Versus 3 Collocated Display Monitors. *Advances in Radiation Oncology* 6, 1 (2021), 100572. https://doi.org/10.1016/j.adro.2020.09.012

[27] Eimei Oyama, Naoji Shiroma, Norifumi Watanabe, Arvin Agah, Takashi Omori, and Natsuo Suzuki. 2016. Behavior navigation system for harsh environments. *Advanced Robotics* 30, 3 (feb 2016), 151–164. https://doi.org/10.1080/01691864.2015.1113888

[28] Eimei Oyama, Kohei Tokoi, Ryo Suzuki, Sousuke Nakamura, Naoji Shiroma, Norifumi Watanabe, Arvin Agah, Hiroyuki Okada, and Takashi Omori. 2021. Augmented reality and mixed reality behavior navigation system for telexistence remote assistance. *Advanced Robotics* 35, 20 (2021), 1223–1241. https://doi.org/10.1080/01691864.2021.1976670

[29] Danny Schott, Patrick Saalfeld, Gerd Schmidt, Fabian Joeres, Christian Boedecker, Florentine Huettl, Hauke Lang, Tobias Huber, Bernhard Preim, and Christian Hansen. 2021. A VR/AR environment for multi-user liver anatomy education. *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2021* (2021), 296–305. https://doi.org/10.1109/VR50410.2021.00052

[30] Mickael Sereno, Xiyao Wang, Lonni Besancon, Michael J Mcguffin, and Tobias Isenberg. 2020. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 2626, c (2020), 1–1. https://doi.org/10.1109/TVCG.2020.3032761

[31] Kazuma Takada, Midori Kawaguchi, Akira Uehara, Yukiya Nakanishi, Mark Armstrong, Adrien Verhulst, Kouta Minamizawa, and Shunichi Kasahara. 2022. Parallel Ping-Pong: Exploring Parallel Embodiment through Multiple Bodies by a Single User. In *Augmented Humans 2022*. ACM, New York, NY, USA, 121–130. https://doi.org/10.1145/3519391.3519408

[32] Peng Wang, Xiaoliang Bai, Mark Billinghurst, Shusheng Zhang, Xiangyu Zhang, Shuxia Wang, Weiping He, Yuxiang Yan, and Hongyu Ji. 2021. AR/MR Remote Collaboration on Physical Tasks: A Review. *Robotics and Computer-Integrated Manufacturing* 72, May (2021), 102071. https://doi.org/10.1016/j.rcim.2020.102071

[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 *HCII 2020. Lecture Notes in Computer Science*. 407–417. https://doi.org/10.1007/978-3-030-50017-7_30

[34] Tzu-Yang Wang, Yuji Sato, Mai Otsuki, Hideaki Kuzuoka, and Yusuke Suzuki. 2020. Effect of body representation level of an avatar on quality of AR-based remote instruction. *Multimodal Technologies and Interaction* 4, 1 (2020), 1–19. https://doi.org/10.3390/mti4010003

[35] Christopher D. Wickens, Justin G. Hollands, Simon Banbury, and Raja Parasuraman. 2015. Engineering Psychology and Human Performance. *Engineering Psychology and Human Performance* (2015). https://doi.org/10.4324/9781315665177

[36] Nancy Yao, Jeff Brewer, Sarah D'Angelo, Mike Horn, and Darren Gergle. 2018. Visualizing Gaze Information from Multiple Students to Support Remote Instruction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, Vol. 2018-April. ACM, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188453