Evaluating Workload in One-to-Many Remote Collaboration

Tzu-Yang Wang*
School of Knowledge Science
Japan Advanced Institute of Science and Technology
Nomi, Ishikawa, Japan
wang.tzuyang@gmail.com

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

ABSTRACT

In this study, we focused on one-to-many remote collaboration, which is becoming one of the essential topics in CSCW. We conducted an experiment comparing the remote instructors' workload while interacting with different number of local workers. The results showed that the remote instructors perceived strong workload when interacting with multiple local workers.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; **Mixed / augmented reality**; **Collaborative interaction**.

KEYWORDS

Workload, Remote Collaboration, One-to-many

ACM Reference Format:
Tzu-Yang Wang, Mai Otsuki, and Hideaki Kuzuoka. 2022. Evaluating Workload in One-to-Many Remote Collaboration. In *The15th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '22), June 29-July 1, 2022, Corfu, Greece.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3529190.3534717

1 INTRODUCTION

For decades, many researcher have paid much attention on supporting remote experts/instructors to support local workers. However, most studies focused on one-to-one remote collaboration (e.g., [3, 6]). Due to the shortage of experts, there are recently more situations that one expert needs to support multiple workers. Thus, researchers started aiming at one-to-many remote collaboration. For example, Lee et al. investigated different view-sharing AR techniques to support such kind of collaboration [4].

Compared with one-to-one remote collaboration, one-to-many cases may demand more mental resources for remote experts, such as comprehending each worker's situation and giving appropriate support individually. Thus, remote experts may experience high workload. Our goal is to develop a suitable one-to-many remote collaboration system. As an early stage of our research, we aimed at investigating how remote experts experienced workload while interacting with different numbers of local workers.

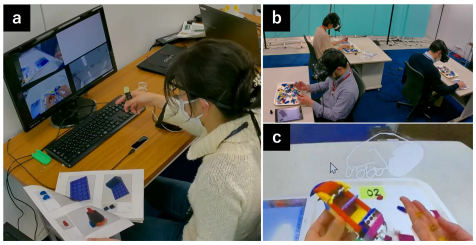

Figure 1: One-to-many remote collaboration system used in this research. (a) Remote: Participant can give instructions by voice and hand gestures via display. Each display shows the worker's view. (b) Local: Workers assembles LEGO blocks. (c) A worker's view. Workers can see the participant's AR hand gesture.

2 EXPERIMENTAL DESIGN

2.1 Hypothesis, Task and Apparatus

We hypothesized that instructors experienced higher workload while interacting with more workers at the same time. To examine the hypothesis, we developed a one-to-many remote collaboration system (Fig. 1). Local workers used HMD with a stereo camera (htc Vive pro, and ZED Mini) to observe the local environment, and an remote instructor could observe at most four local workers' views through an LCD monitor simultaneously. Additionally, the remote instructor's hand gestures were captured by Leap motion, and the remote instructor could freely choose to overlay his/her hand gestures on one of the local users' views.

As for the task, we adopted the LEGO assembly task, which is a conventional method for remote collaboration. There were one instructor and one to three workers in the task. Both the instructor and workers received a manual containing the assembly procedures. To simulate the situation that workers need assistance by the instructor, the workers received uncompleted manuals that 15% of the procedures were missing. The workers asked the instructor for help while encountering the missing steps. Since our main target of this

*The author is a former research assistant in National Institute of Advanced Industrial Science and Technology (AIST)

PETRA '22, June 29-July 1, 2022, Corfu, Greece

ACM ISBN 978-1-4503-9631-8/22/06…$15.00
https://doi.org/10.1145/3529190.3534717

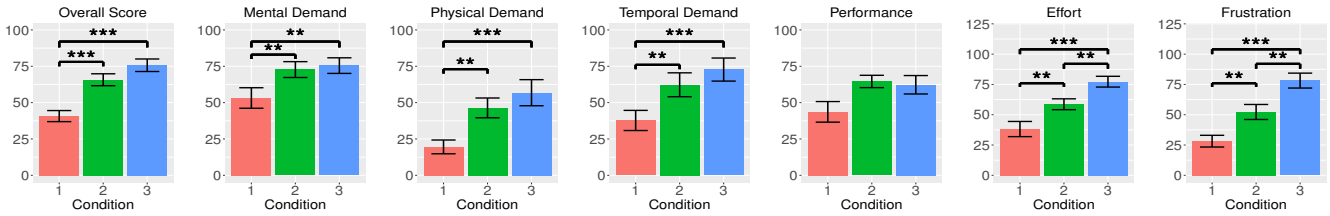

Figure 2: NASA-TLX overall score and subscale score

investigation was the instructors, a group of three workers in this experiment were well-trained confederates and they participated in all sessions, to minimize any bias that might be caused by the difference between each worker group.

2.2 Participants

Eleven Japanese participants were recruited. There were six males and five females, and the average age was 26.27 (SE = 4.9). Each participant received 5000 yen as compensation.

2.3 Procedures

The participant sat in front of a monitor and started a three-session experiment. In each session, he/she played the role of instructors, receiving a complete version manual and supported the workers when necessary. The number of workers was 1, 2, or 3 depending on the session. The task was stopped in 25 minutes although the required time for the assembly task was longer than 25 minutes. After the task, the participant filled in three questionnaires (NASA-TLX for measuring the experienced workload [2], SUS for measuring the usability of the remote collaboration system [1], and QCE for measuring the communication experience quality [5]). After filling out the questionnaire, the participant rested for 10 minutes and started another session. The number of workers and the LEGO models in the three sessions was different and randomized.

3 RESULTS

For NASA-TLX, we constructed a linear mixed model to analyze the result of NASA-TLX (Fig. 2). The Type III Analysis of Variance Table showed that there was a significant difference in NASA-TLX score between different numbers of workers ($F(2,20) = 25.166$, $p < .001$). Post hoc test with Bonferroni correction showed that the NASA-TLX score in the 1-worker condition was significantly lower than the score in the 2-worker condition ($t(20) = 4.92$, $p < .001$) and the score in the 3-worker condition ($t(20) = 6.89$, $p < .001$).

Additionally, linear mixed models were constructed to analyze the result of each sub-scale of NASA-TLX. The models showed that there was significant differences in all sub-scale scores between conditions. Post hoc test with Bonferroni correction showed that except for "Performance", other sub-scale score in the 1-worker condition was significantly lower than the same sub-scale score in the 2-worker condition and in the 3-worker condition. Also, for "Effort" and "Frustration", the scores in the 2-worker condition were significantly lower than those scores in the 3-worker condition.

As for SUS, the result of linear mixed model showed a significant difference between conditions ($F(2,20) = 12.13$, $p < .001$). Post hoc test showed that the SUS score in the 1-worker condition was higher than the score in the 2-worker condition ($t(20) = 2.74$, $p = .004$) and the score in the 3-worker condition ($t(20) = 4.91$, $p < .001$).

For QCE, following the original factor structure, three factors (clarity, responsiveness, comfort) were analyzed by linear mixed model. The result of clarity and responsiveness showed a significant difference between conditions ($F(2,20) = 6.06$, $p = .008$, and $F(2,20) = 9.62$, $p = .001$, respectively). In addition, the score in the 1-worker condition was higher than the score in the 3-worker condition ($t(20) = 3.43$, $p = .008$, and $t(20) = 4.37$, $p < .001$, respectively).

4 DISCUSSION AND CONCLUSION

The result of this research strongly indicated that the participants experienced stronger workload while interacting with more than one worker. Especially in terms of both "Effort" and "Frustration", as increasing the number of workers, workload became significantly higher. This finding supports our hypothesis. Interview results indicate that the instructor could not pay enough attention to a particular worker when other workers seemed to be waiting for the instructor's support. According to the result of QCE, the gained stress also reduced the comprehension of the communication and might generate more conflicts during collaboration. The result of SUS suggests that the the current remote collaboration system should be improved for one-to-many situations. As a future work, it is necessary to conduct video analysis to figure out the key components causing high workload and redesign an appropriate one-to-many remote collaboration system based on the findings.

ACKNOWLEDGMENTS

This work was supported by The Telecommunications Advancement Foundation. We appreciate Mr. Li He supporting data analysis.

REFERENCES

[1] John Brooke. 1996. Sus: a "quick and dirty'usability. *Usability evaluation in industry* 189, 3 (1996).
[2] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909
[3] 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.
[4] Geonsun Lee, HyeongYeop Kang, JongMin Lee, and JungHyun Han. 2020. A user study on view-sharing techniques for one-to-many mixed reality collaborations. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 343–352.
[5] 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.
[6] 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), 3.