Study of communication modalities for teaching distance information

Francesco Fastelli, Cassandre Simon, Aylen Ricca, Amine Chellali

Study of communication modalities for teaching distance information

Francesco Fastelli[#]
IBISC Lab, Univ Evry, Université Paris Saclay

Cassandre Simon[†]
IBISC Lab, Univ Evry, Université Paris Saclay

Aylen Ricca[§]
Arts et Metiers Institute of Technology, LISPEN, UBFC

Amine Chellali*
IBISC Lab, Univ Evry, Université Paris Saclay

ABSTRACT

We present an exploratory study to compare the haptic, visual, and verbal modalities for communicating distance information in a shared virtual environment. The results show that the visual modality decreased the distance estimation error while the haptic modality decreased the completion time. The verbal modality increased the sense of copresence but was the least preferred modality. These results suggest that a combination of modalities could improve communication of distance information to a partner. These findings can contribute to improving the design of collaborative VR systems and open new research perspectives on studying the effectiveness of multimodal interaction.

Index Terms: Human-centered computing—Interaction Techniques; Human-centered computing—Virtual Reality.

1 INTRODUCTION

Mastering technical skills, such as tool manipulation, requires time and practice. During the early learning stages, guidance from a teacher is necessary to improve skill acquisition. This is called "augmented feedback"; information augmented by a teacher using verbal statements or visual aids [1]. In some cases, the teacher takes the learner's hand to show the correct movement to perform. Augmented feedback improves satisfaction [2], reduces errors [3], and enhances skill acquisition [3]. However, VR technologies supporting this learning model are still limited.

This exploratory work studies how an instructor can guide a learner to move a tool with the correct amplitude. The verbal, visual, and haptic instructions are compared for their impact on user performance and user experience. The main contribution of this work is to inform on the role of each modality during teacher-learner communication in shared virtual environments (SVE).

2 USER STUDY

2.1 Experimental task

A new task was designed to mimic a tool manipulation under the guidance of an instructor in a SVE. It consisted of picking up a small 3D sphere from a starting position and placing it at a target position using a virtual tool controlled by a haptic arm. The target position was unknown for the participant who received indications about it from an instructor (experimenter) using the verbal, visual, or haptic modality. Thus, the task was divided into two steps:

Step 1: the movement to perform was described to participants using one modality according to the experimental condition.

Step 2: the participants had to replicate the described movement on their own in the VE as quickly and accurately as possible. They used a haptic arm to pick a 3D sphere, move it to the target position following the described movement, and place it.

2.2 Participants, experimental design and, conditions

Twenty-one participants were recruited for the experiment. The study followed a within-subjects design with one factor (modality) and three conditions: verbal (VB), visual (VS), and haptics (HP).

VS: Participants received visual instructions by watching a pre-recorded 3D animation in the VE. This animation consisted of a static realistic virtual hand manipulating the tool to pick up a sphere and place it at the target position simulating the instructor's hand movement when performing the same task.

VB: Verbal instructions were given by the instructor through reading a script inviting the participant to move the ball according to one direction (left/right, up/down, forward/backward) and with an amplitude in centimeters (e.g., "five centimeters to the left").

HP: Haptic instructions were received by participants by holding a haptic arm's stylus. The arm was then moved following a pre-recorded path from the starting position to the end position simulating the instructor performing the same movement.

The presentation order of the conditions was counterbalanced to avoid any learning effect. Each participant performed twelve trials for each condition. Movement amplitudes ranged from 3 to 13 centimeters and were randomly picked up for each trial. The amplitudes were counterbalanced between conditions.

2.3 Experimental setup, procedure, and measures

A Vive Pro HMD was used for visualization, and a Geomagic Touch device was used for interactions and haptic instructions (Figure 1). The VE was developed on Unity3D. To improve distance estimation, it replicated as faithfully as possible the real one (exact object sizes and colors). During step 1 of VB and HP conditions, the VE was composed only of two perpendicular planes. In the VS condition and during step 2, it also included a sphere and a static hand avatar holding a tool (Figure 1).

The participants filled in a demographics questionnaire and were familiarized with the VE components and interactions. After that, the actual experiment started. After completing the 12 trials for one condition, participants answered a copresence questionnaire. Then, they performed 12 other trials for the following condition. After finishing all trials and questionnaires of all conditions, they filled in a modality comparison questionnaire.

Objective measurements included distance estimation error (Euclidean distance between the centers of the moving sphere and the target sphere) and the pick and place completion time. Subjective measures included answers to the copresence (Qa) and the modality comparison (Qb) questionnaires.

Figure 1: The VE (top left) and the experimental setup with the instructor (left) guiding the participant (right)

[#]francesco.fastelli@universite-paris-saclay.fr
[†]cassandre.simon@univ-evry.fr
[§]aylen.ricca@ensam.eu
*amine.chellali@univ-evry.fr

3 Results and Discussion

The results show a significant main effect of modality on distance estimation error ($\chi^2 = 12.61$, $p = 0.002$). The Wilcoxon tests show that the mean error was significantly lower in the VS condition than VB and HP conditions. This is in line with previous research suggesting that visual augmented feedback is a more effective learning strategy [4]. While the information provided by each modality is different, the visual modality can be considered as the best means for teaching distance estimation.

The one-way repeated measure ANOVA shows a main effect of modality on completion time ($F_{(2,40)} = 11.26$, $p < 0.001$). The post-hoc tests show that the meantime was significantly lower in the HP condition than in the VB and the VS conditions. Furthermore, participants evaluated this modality as the most difficult for memorizing instructions (Qb6). This suggests that they tried to follow the instructions quickly after receiving them while still "fresh" in their memory. However, this hypothesis needs to be confirmed.

According to the copresence questionnaire, the verbal modality was generally the most preferred (Table 1). The participants found the virtual experience much closer to a real-world meeting and felt a stronger connection with the instructor when using this modality (Qa4, Qa6). In addition, they thought that the instructor was warmer (Qa7) and was more helpful (Qa8) in this condition than in the visual one. Besides, they found the verbal instructions more understandable than the haptic ones (Qa10).

There is a contrast between the copresence and the objective measurements results. The participants' preference of the verbal modality might be explained by the experimenter being physically close to them and the verbal instructions coming directly from the "real world". In contrast, the visual and haptic instructions were provided in the VE through a visual animation or the haptic arm. Hence, the sense of copresence could have been impacted by the communication means: direct (for verbal) and mediated (for haptic and visual). In the future, it will be interesting to compare the three conditions with the same level of mediation (for instance, with an instructor located in another room or with pre-recorded voice messages displayed on a headset).

Table 1: Comparison of the copresence questionnaire answers (only significant values are included)

Question	Friedman test: χ^2 (p)	Wilcoxon test (p)		
		VS-VB	VS-HP	VB-HP
Qa4	7.89 (0.01)	NS	NS	0.03
Qa6	12.23 (0.002)	0.006	0.045	NS
Qa7	7.96 (0.019)	0.012	NS	NS
Qa8	10.14 (0.006)	NS	NS	0.021
Qa10	9.48 (0.009)	NS	NS	0.024

The haptic modality generated a stronger connection with the instructor than the visual modality (Qa6). The low scores in the visual modality could be associated with using a non-animated hand avatar. Indeed, the low kinematic fidelity of the partner's avatar has already been reported to impact this dimension [5] negatively. In contrast, the haptic modality has been reported to increase the feeling of closeness and intimacy with others [6].

The subjective comparison between modalities contrasts with the copresence results but is in line with task performance (Figure 2). Participants thought they were learning more with the haptic and visual modalities (Qb8), which were more effective in receiving spatial information than the verbal modality (Qb10). The latter is also rated as the least engaging (Qb9). In addition, the differences between the visual and haptic modalities were not significant. This suggests that both were generally well accepted. Finally, the haptic modality was the most disturbing (Qb5). The novelty of this communication may explain this means the participants have never experienced that. Nevertheless, this did not impact either their performance or their user experience. Thus, it will be explored in the future to improve teacher-learner communication in SVE.

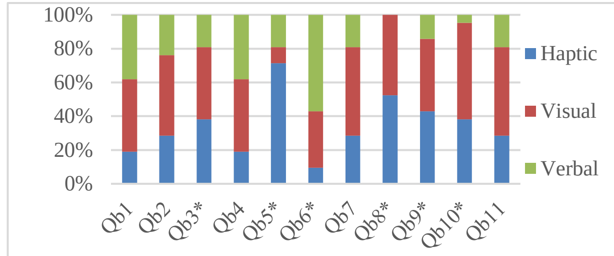

Figure 2: Participants' most preferred modalities (* indicates significant differences with p-values <0.05)

4 Conclusion

This work is part of a research project aiming to design SVE dedicated to technical skills learning. The present study results indicate that the visual modality is the most accurate for teaching spatial tool manipulation. In contrast, the haptic modality permitted a faster execution of the instructions. On the other hand, the verbal modality increased the sense of copresence but was generally the least preferred modality.

These findings give various insights on designing collaborative interactions for spatial skills learning in VE. Indeed, they suggest that each modality can bring additional features to improve the learning experience and performance and that multimodal interactions could be the most appropriate approach. Hence, we plan to study the impact of combining modalities on the learning experience and performance in the future. We also plan to study the effects of modalities on the teacher's experience. This will help us design more appropriate user interfaces supporting technical skills transfer between a teacher and a learner in immersive SVE.

5 Acknowledgements

This work was supported by a grant from the Agence Nationale de la Recherche (ANR-20-CE33-0010 Show-me).

References

[1] T. D. Lee, S. P. Swinnen, and D. J. Serrien, Cognitive Effort and Motor Learning, *Quest*, volume 46, pages 328–344, 1994.

[2] C. MacKenzie, T. M. Chan, and S. Mondoux, Clinical Improvement Interventions for Residents and Practicing Physicians: A Scoping Review of Coaching and Mentoring for Practice Improvement, *AEM Education and Training*, volume 3, pages 353–364, 2019.

[3] B. Lovell, What do we know about coaching in medical education? A literature review, *Medical education*, volume 52, pages 376–390, 2018.

[4] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review, *Psychonomic bulletin & review*, volume 20, pages 21–53, 2013.

[5] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, and S. Otmane, Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments, *Personal and Ubiquitous Computing*, volume 25, pages 467–484, July 2020.

[6] E.-L. Sallnäs, K. Rassmus-Gröhn, and C. Sjöström, Supporting presence in collaborative environments by haptic force feedback, *ACM Transactions on Computer-Human Interaction*, volume 7, no. 4, pages 461–476, December 2000.