The Augmented Reality affective training increases the care communication skill and empathy

Atsushi Nakazawa (nakazawa.atsushi@i.kyoto-u.ac.jp)
Kyoto University
Miyuki Iwamoto
Kyoto University
Ryo Kurazume
Kyushu University
Masato Nunoi
Sugiyama Jogakuen University
Masaki Kobayashi
Tokyo Medical Center
Miwako Honda
Tokyo Medical Center

Article

Keywords:

Posted Date: June 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1667324/v1

Additional Declarations: No competing interests reported.

The Augmented Reality affective training increases the care communication skill and empathy

Atsushi Nakazawa[1,*,+], **Miyuki Iwamoto**[1,+], **Ryo Kurazume**[2,+], **Masato Nunoi**[4], **Masaki Kobayashi**[3], **and Miwako Honda**[3,+]

[1]Kyoto University, Dept. of Informatics, Kyoto, 606-8501, Japan.
[2]Kyushu University, Faculty of Information Science and Electrical Engineering, Fukuoka, 819-0395, Japan.
[3]Tokyo Medical Center, Tokyo, 152-8902, Japan.
[4]Sugiyama Jogakuen University, School of Human Sciences, Aichi, 470-0136, Japan.
[*]nakazawa.atsushi@i.kyoto-u.ac.jp
[+]these authors contributed equally to this work

ABSTRACT

It is important for caregivers of people with dementia (PwD) to have good patient communication skills as it has been known to reduce the behavioral and psychological symptoms of dementia (BPSD) of PwD as well as caregiver burnout. However, acquiring such skills often requires one-on-one affective training which can be costly. In this study, we propose affective training using augmented reality (AR) for supporting the acquisition of such skills. The system uses a see-through AR glasses and a nursing training doll to train the user in both practical nursing skills and affective skills such as eye contact and patient communication. The experiment was conducted with 39 nursing students. The participants were assigned to either the Doll group, which only used a doll for training, or the AR group, which used both a doll and the AR system. The results showed that eye contact significantly increased and the face-to-face distance and angle decreased in the AR group, while the Doll group had no significant difference. In addition, the empathy score of the AR group significantly increased after the training. Upon analyzing the correlation between personality and changes of physical skills, we found a significant positive correlation between the improvement rate of eye contact and extraversion in the AR group. These results demonstrated that affective training using AR is effective for improving caregivers' physical skills and their empathy for their patients. We believe that this system will be beneficial not only for dementia caregivers but for anyone looking to improve their general communication skills.

Introduction

Caregiving for people with dementia (PwD) is becoming increasingly important as the number of people with dementia has been growing worldwide[1–3]. In Japan, the number of PwD is expected to exceed seven million in 2025, and the number of caregivers that will be needed is estimated to be 2.53 million. However, the number of actual caregivers has been decreasing and is estimated to be only 2.15 million in 2025[4]. One of the causes of the decrease in caregivers is burnout. Dementia can cause symptoms similar to those of mental illness, known as behavioral and psychological symptoms of dementia (BPSD), which can make it challenging for caregivers to handle their patients. As a result, more and more caregiving staff members and family caregivers are experiencing burnout and leaving the profession[5].

Improving nursing communication skills have been reported to reduce the aforementioned problems[6–9]; however, such skill training requires additional costs. Moreover, there are a limited number of capable instructors, so it is impractical for all caregivers to receive such training. Virtual reality (VR) technology has drawn attention as an alternative way of instructing the affective aspects of nursing. The VR techniques can be categorized into two types. In the first type, the participants experience being a PwD in a virtual world[10–12]; for example, their vision and hearing are limited and their perceptions of scenes and/or people are distorted. By using these systems, participants can experience the difficulties that PwD face on a daily basis. Many studies reported that the participants' empathy for PwD increased after using these systems[13–21]. However, the systems did not encourage participants to improve their communication skills for PwD.

The second type has participants behave as a caregiver and learn skills for communicating with PwD using VR, though little work has done in this area. Elliman developed a game where the participant can converse with PwD in VR[22]. UBISIM is a VR simulation system in which participants can learn medical procedures and interact with virtual patients[23]. However, these systems were designed mainly to teach medical procedures rather than patient communication skills. In addition, the educational effects of these systems with respect to affective training are still unclear.

In this study, we developed an augmented reality (AR)-based nursing communication training system and evaluated its effectiveness. As a reference for ideal care communication, we use the caregiving methodology *Humanitude* developed by

Geneste and Marescotti[24], which has been introduced in more than 600 hospitals and nursing homes in Europe. Several studies have reported the effectiveness of Humanitude in reducing BPSD and the burden on caregivers[25]. In addition to verbal communication, the Humanitude system also focuses on non-verbal communication techniques such as eye contact, face-to-face distance, and physical touch[26], which are often difficult to learn from classroom lectures only.

In the development of the affective communication training system, AR technology has the following advantages over the VR-based systems:

1. Because participants can see the real view using the see-through AR headset, they can grasp the physical reality of the communication with patients, such as eye contact and face-to-face distance to the patient's face. In contrast, the scenes and patients are represented with fully virtual computer graphics in VR.
2. Participants can perceive the levels of their caregiving skills through the AR display or voice annotation to improve their communications.
3. Participants can learn skills requiring physical contact, such as touching the patient's body and changing their clothes, which cannot be done in VR.
4. Participants can train their communication skills independently using the system, which increases their opportunities to practice. This also reduces the need to gather with instructors in person, reducing training costs and the risk of spreading illnesses.
5. The system is not limited to training communication skills for PwD as the avatars' behavioral models can be replaced for other purposes.

The main objectives of this study are:

1. observe the effects of training on the physical skills of the participant groups using AR and the existing method, where the physical skills include the occurrence of eye contact, face-to-face distance and angle, and length of utterances;
2. observe psychological changes in the participants, i.e., the difference in empathy scores before/after the nursing communication training using different training methods;
3. clarify the relationship between the participants' personality traits and the effects of training, which are important in determining the potential target candidates for using the AR system.

The remainder of the paper is as follows. In the section 2, the study design, study procedure, the AR system and the data analysis are described, followed by results and discussion.

Material and Methods

Study design and participants

A two-group randomized controlled study was conducted. Participants were recruited from students in the clinical nursing department at Bukkyo University, Japan. The demographics of the participants (N = 38) are listed in Table 1. Participants ranged from 18 to 28 years old, with an average age of 20.9 years old. All of them were born in Japan and are native Japanese speakers. The students ranged from first to fourth years, and 82% of them were female. They were paid 2,000 JPY (about 16 USD) for their two-hour participation. Written informed consent was obtained from all patients for the participation of the study and publication of accompanying images. They were randomly assigned to either the control (Doll) or the experimental (AR) group.

Table 1. Participants

	Control (Doll)		Experimental (AR)	
	M	F	M	F
All	3	14	4	17
1st grade	0	2	0	5
2nd grade	0	3	0	3
3rd grade	0	3	2	4
4th grade	3	6	2	5
Av. age	21.06		20.81	

Measure

All participants filled out and submitted a shortened version of the Japanese Big-Five Scale[27] and the Jefferson Scale of Empathy (JSE) Health Profession Students' version[28]. A 7-point Likert scale was used for both surveys. The Big-Five was administered and collected once before the experiment, whereas the JSE was collected before and after the experiment to observe the effects of training on the participants' empathy.

To measure the physical behavioral features, first-person videos were recorded from head-mounted cameras worn by the participants. From the videos, we then retrieved face-to-face distance and pose, the occurrence of eye contact between the participants and simulated patients, and the length of the caregiver's speech. These skills are vital aspects of the Humanitude care technique.

Apparatus

Simulated patients

Twelve female faculty members in the nursing department at Bukkyo University participated as simulated patients. Before the experiment, they attended a three-hour lecture on portraying a simulated patient with dementia. The participants' training groups were not disclosed to the simulated patients.

Life-size nursing doll

Life-size nursing dolls (Sakura, Kyoto Kagaku) were used for training. The doll was positioned sitting in a wheelchair, and the participants practiced communication and assigned tasks in independent practice sessions. For the AR group, the face of the doll was covered by a white cloth because the face was presented by a AR headset, whereas it was not covered for the Doll group.

First- and third- person video

The Tobii Pro Glasses 3 were used to collect the first-person video and gaze data of the participants while they were simulating patient care tasks in the pre- and post-evaluations. In addition, a standard video camera was used to record the sessions. Although the Tobii systems collected the participants' gaze data, we did not use the data because the number of frames where the data were successfully taken varied between individuals due to technical limitations. Thus, instead of using the gaze data, we used the first-person videos (frontal view) for further analysis.

AR system

The AR-based care communication training system HEARTS (Humanitude AR Training System) developed by Kurazume et. al[29] was used to train the participants in the AR group. The system configuration of HEARTS is shown in Fig.1. The system consists of an AR device (Microsoft HoloLens 2) and the nursing doll. To imitate the interaction between a patient and a caregiver, we superimposed a three-dimensional CG model of a patient's face onto the head of the doll using AR technology. The gaze direction, blinking, and facial expressions of the simulated patient were dynamically controlled on the basis of the caregiver's movement. Six facial expressions corresponding to happiness, anger, surprise, sadness, fear, and neutral were implemented as shown in Fig.2. The eye contact and verbal communication between the patient and the caregiver were detected and evaluated online by the AR device. The evaluation results are indicated by "Good," "Bad," or the number of asterisks on the screen of the HoloLens. The criteria is as follows:

- **Eye contact**
  "Good" if the frontal axis of the HoloLens passes through the AR's face and the relative angle between the frontal axis of the HoloLens and the avatar's frontal axis is less than 25 degrees, otherwise "Bad." The facial expression also changes to "happiness" if the eye contact is "Good."
- **Verbal communication**
  "Good" if the sound pressure is stronger than the threshold, otherwise "Bad.
- **Face-to-face distance**
  One, two, or three asterisks corresponding to distances over 1400 mm, between 700 and 1400 mm, and under 700 mm.

If there is no eye contact, conversation, a warning "No communication" is issued and the facial expression changes to "sadness" or "fear." In addition, the numerical score for eye contact is evaluated by the ratio of the total time the caregiver makes eye contact to the total nursing care time. The numerical score of the verbal communication skill is determined by the ratio of the total speaking time to the total nursing care time. These scores are displayed on the screen of the HoloLens as shown in Fig.3.

Procedure

The evaluation was conducted in accordance with the procedure approved by the ethics committee of the Tokyo Medical Center numbers R21-055 and R21-079. The study protocol adheres to the Declaration of Helsinki and follows the latest ethical guidelines for clinical research in Japan, and is registered in the Japanese clinical research registry system that meets the criteria of the International Committee of Medical Journal Editors (ICMJE) (University Hospital Medical Information Network Clinical Trials Registry Number: UMIN000046670).

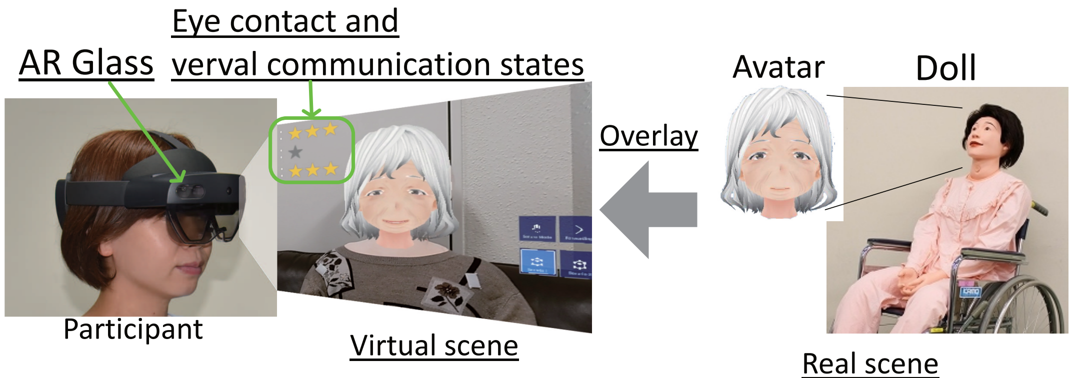

Figure 1. System configuration of HEARTS

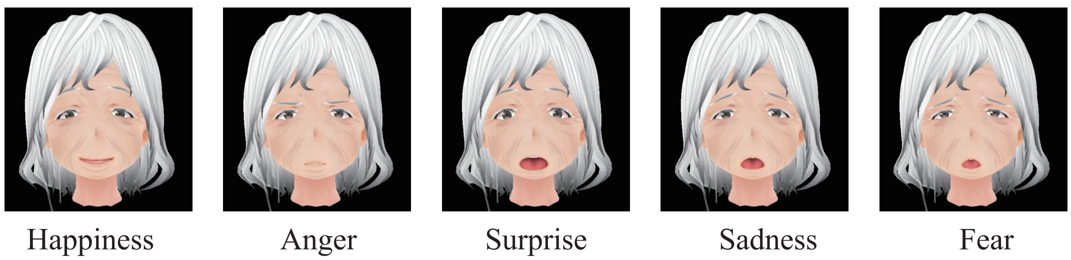

Figure 2. Facial expressions of the avatar.

The overview of the procedure is shown in Fig. 4. Two days before the experiment, the participants were asked to read a Humanitude textbook, which provides a brief overview of the fundamental aspects and skills of Humanitude, and then they answered the JSE and Big-Five through online questionnaires. Afterwards, the participants answered a ten-question mini test about Humanitude to encourage them to learn about Humanitude from the textbook.

On the day of the experiment, the participants were first asked to perform care tasks for a simulated patient with dementia, e.g., the patient did not give clear responses to the participants. The care tasks included helping the patient change clothes and put on socks. During this experiment, the participants wore an eye tracking device, and the simulated patient wore a first-person camera. The entire trial was recorded by a video camera. Then the participants were randomly assigned to either the control or experimental group. The participants were asked to practice the same task by themselves for 30 minutes. In the control group, the participants practiced the task using a patient-care doll, while the participants assigned to the experimental group used the AR headset and the doll with the facial image overlaid on its head. After the independent training session, the participants performed the same tasks with the same simulated patient. Finally, they filled out the JSE and a questionnaire on the feasibility and usability of the training.

Analysis of the first person video

We used the first person videos worn by participants to retrieve eye contact and face-to-face posture information, which is important for evaluating the Humanitude skills[30]. Although we obtained the participants' gaze data, we did not use the information because the success rates of the data collection varied greatly between individuals. We used the video starting from the beginning (knocking the door) to the end of the care sessions (leaving from the room) for the analysis.

Face to face distance and pose

The face-to-face distance and pose (angle) between a participant and a simulated patient were obtained using the image recognition (facial parts detection) algorithm illustrated in Fig. 5. First, the facial parts of a simulated patient were detected using Amazon Rekognition, and then the distance and angle to the face were obtained using Perspective-n-Point (PnP) pose computation[31]. Here, camera parameters of the first person cameras and predefined facial size were used. From the analysis, we obtained (1) the facial detection rate that represents how likely the participant looked at the simulated patient's face, (2) face-to-face distance, and (3) roll, pitch, and yaw angles.

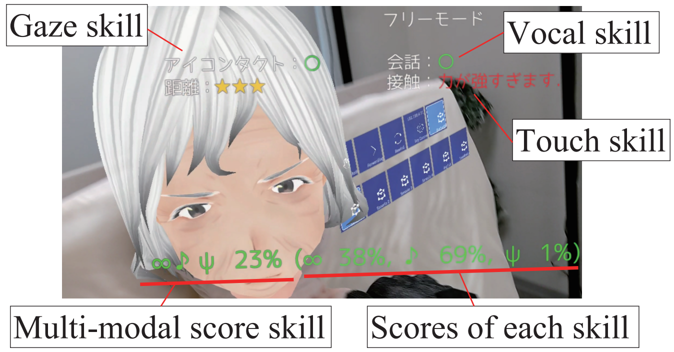

Figure 3. Gaze, speech scores displayed on HEARTS

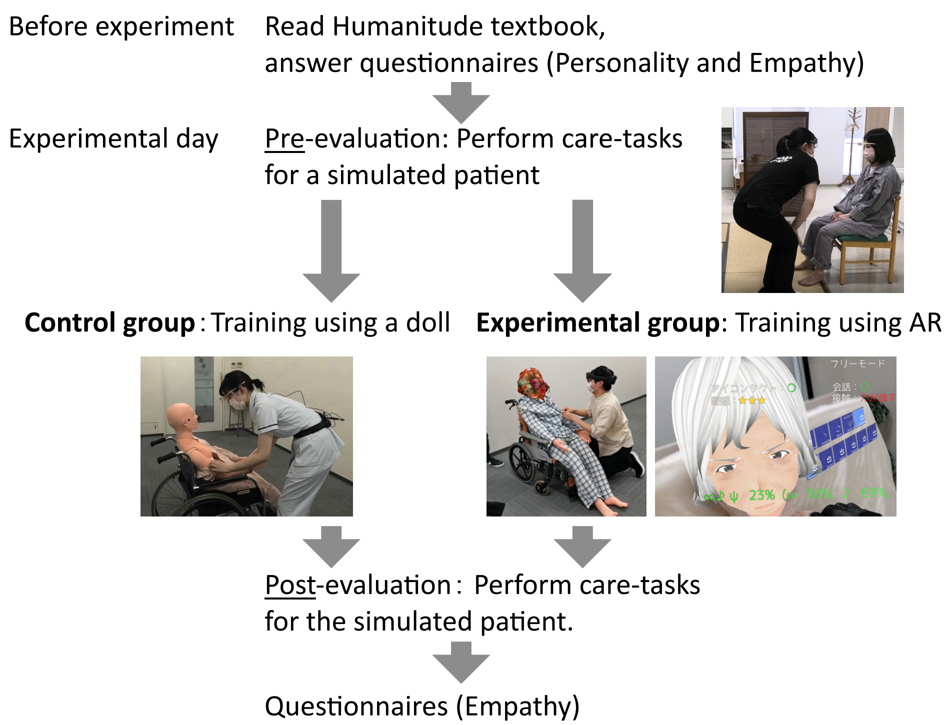

Figure 4. Overview of the experiment.

Eye contact

The eye contact was measured by using a deep neural network-based eye contact detection algorithm[32,33], which has been shown to perform as accurately as human experts. We inputted the first-person video and facial regions to the algorithm and obtained eye contact scores. We assumed the frame where the output eye contact score was greater than 0.8 as the engagement of eye contact.

Utterance

The length of the participants' utterances were detected from the video data. We manually annotated the video and obtained the total utterance duration of each session.

Results

The JSE scores, face-to-face distance and pose, the rate of eye contact, and length of caregivers' speech between the pre- and post-evaluations are compared. The paired t-test (single sided) was used for statistical evaluations, and the error bars in figures denote standard deviation.

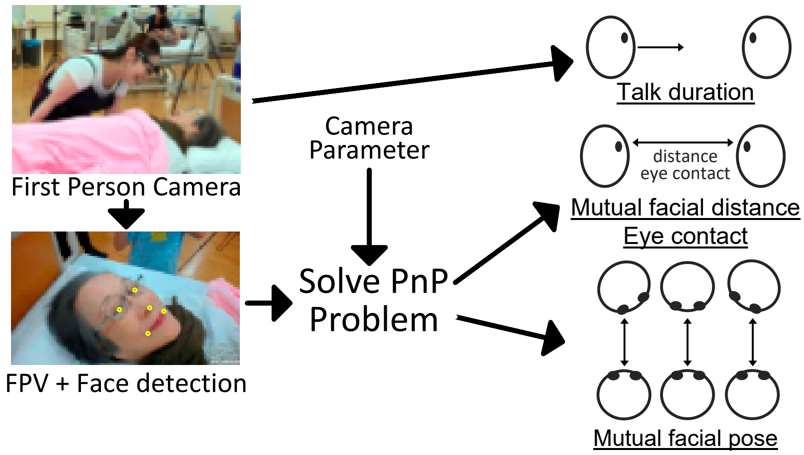

Figure 5. Estimation of face-to-face distance and pose (angles).

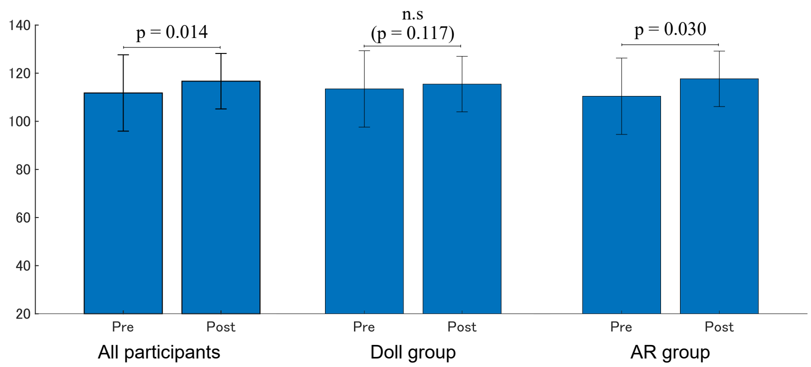

Figure 6. Jefferson Scale of Empathy scores before and after training sessions.

Empathy

Fig. 6 shows the JSE scores before and after the training sessions. Note that the scores ranged from 20 (lowest) to 140 (highest). The average score of all participants significantly increased (Pre: 111.79 (15.87), Post: 116.68 (11.53), p=0.014), that of the doll group had no significant change (Pre: 113.47 (13.13), Post: 115.47 (12.84), p=0.117), whereas that of the AR group significantly increased (Pre: 110.43 (17.99), Post: 117.65 (10.58), p=0.030)).

Face to face distance and pose

Fig. 7 shows the face detection rates which indicate how frequently the participant looked at the simulated patient. The face detection rate significantly increased (Pre: 0.157 (0.105), Post: 0.184 (0.110), p = 0.012) in all participants, and that of the AR group also significantly increased (Pre: 0.168 (0.122), Post: 0.212 (0.122), p = 0.011). However, no significance was found in the Doll group (Pre: 0.145 (0.080), Post: 0.150 (0.085), p = 0.323).

As shown in Fig. 8, the face-to-face distance significantly decreased in all participants (Pre: 980.77 (445.28)[mm], Post: 899.35 (369.48)[mm], p=0.045), but no significance was found in the Doll group (Pre: 1048.96 (582.35)[mm], Post: 933.32(359.49) [mm], p = 0.095) and the AR group (Pre: 925.58 (297.09)[mm], Post: 871.84(383.92)[mm], p=0.153).

Fig. 9 shows the occurrenece rates of when the face-to-face distance is less than 700 [mm], which is the same threshold used in the AR system. They significantly increased in all participants (Pre: 0.096 (0.080), Post: 0.122 (0.087), p = 0.001) and in the AR group (Pre: 0.109 (0.093), Post: 0.145 (0.096), p = 0.002), but no significance was found in the Doll group (Pre: 0.081 (0.058), Post: 0.094 (0.068), p = 0.108).

Fig. 10 shows the results of face-to-face posture (angles). The mean yaw-angles significantly decreased in all participants (Pre: 17.19 (8.41)[deg], Post: 15.58(6.39) [deg] p=0.040) and in the AR group (Pre: 18.16 (9.35)[deg], Post: 15.63 (6.73) [deg], p = 0.010), while no significance was found in the Doll group (Pre: 15.98(7.18)[deg] to 15.51(6.14)[deg], p = 0.382). The smaller yaw angle indicated that participants looked at the simulated patients' face in the frontal direction, leading to improved face-to-face communication skills for the participants in the AR group.

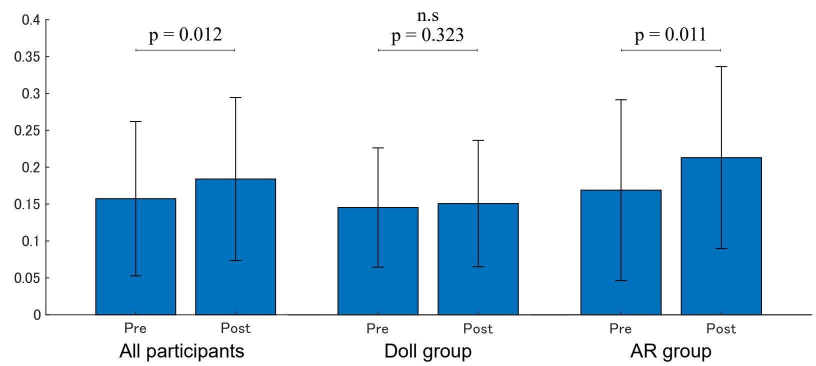

Figure 7. Face detection rate.

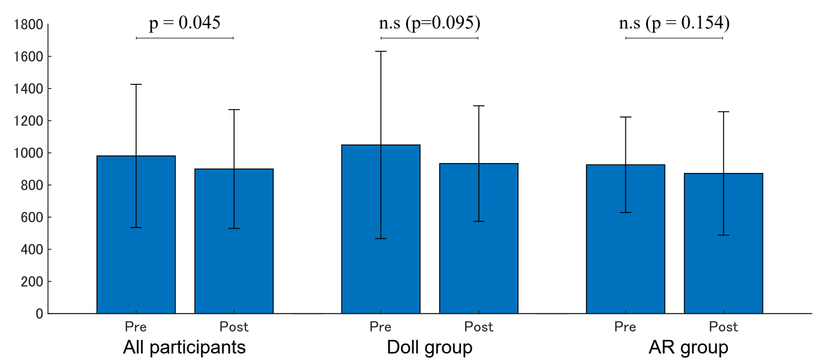

Figure 8. Face-to-face distance ([mm]).

Eye contact

Fig. 11 shows the occurrence rate of eye contact which is computed as the duration of eye contact divided by the session length. A significant increase was found in the AR group (Pre: 0.023 (0.025), Post: 0.032 (0.029), p=0.039), while no significance was found in the Doll group (Pre: 0.016 (0.016), Post: 0.015 (0.018), p=0.501).

Utterance

Fig. 12 shows the utterance duration normalized by the session length. A significant increase was found in both the Doll and AR groups (Doll - Pre: 0.141 (0.073), Post: 0.221 (0.098), p=0.0005; AR - Pre: 0.147 (0.066), Post: 0.267 (0.102), p=0.000).

Personality and behavior

The correlation between personality and physical behavioral parameters was examined. A significant correlation was found between extraversion (ranging from 5 to 35) and the occurrence of eye contact (ranging from 0.0 to 1.0) as shown in Fig.13. Specifically, a significant correlation was found in the post-evaluation in all participants (Pre: $r = 0.198$ (p=0.230), Post: $r =$

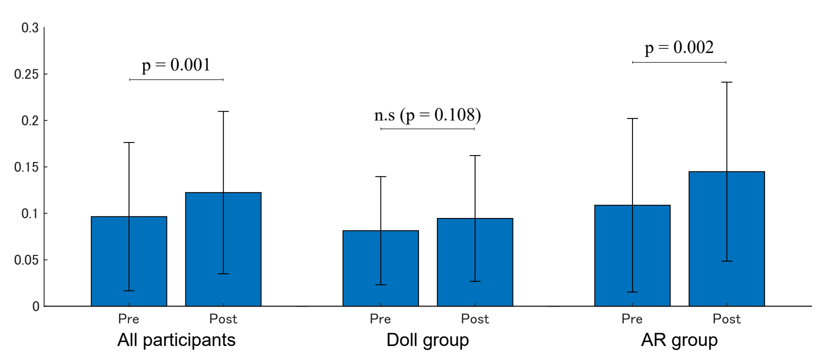

Figure 9. Occurrence rate of where face-to-face distance is less than 700 [mm].

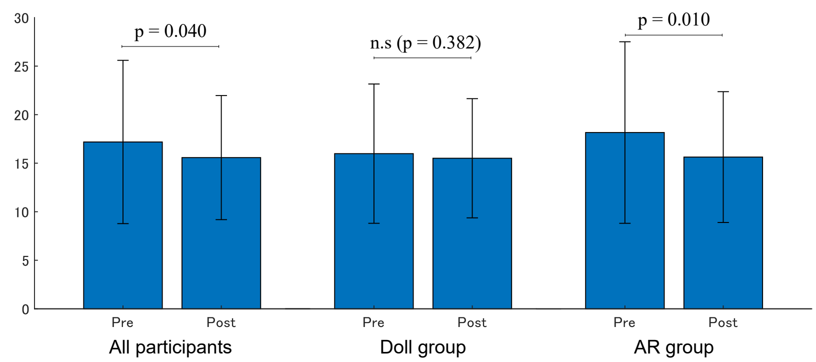

Figure 10. Face-to-face angle (yaw rotation [deg]).

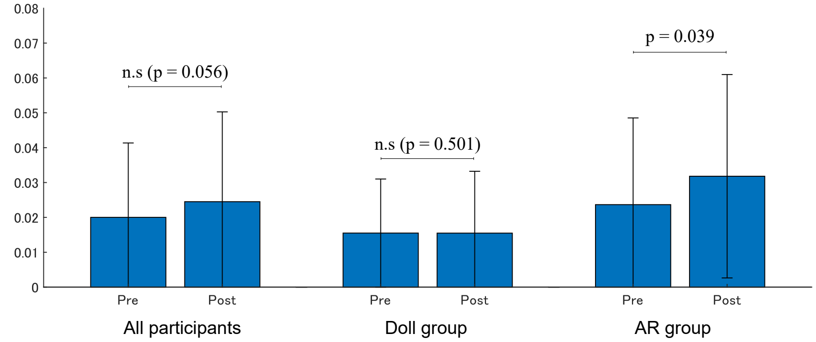

Figure 11. Occurrence rate of eye contact.

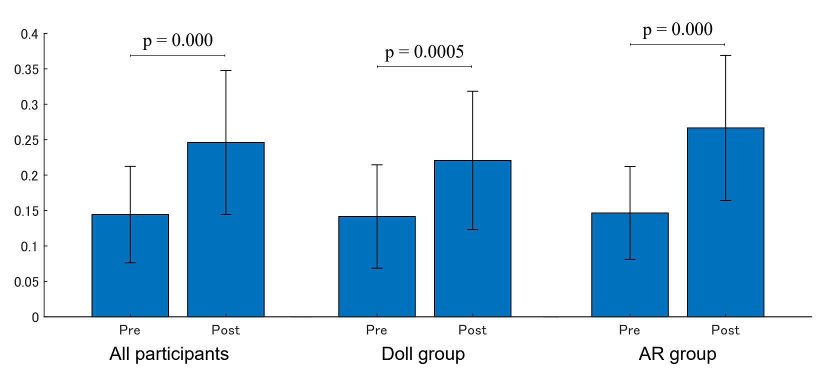

Figure 12. Utterance.

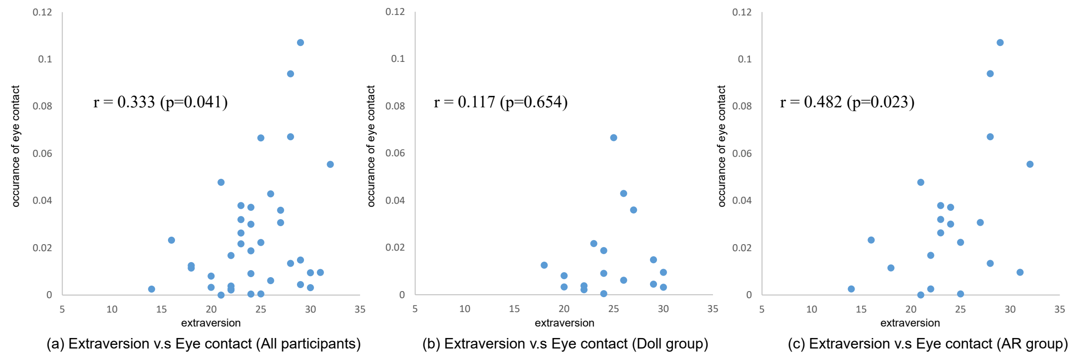

Figure 13. Correlation between the extraversion and the occurance rate of eye contact in the post-evaluation.

0.333 (p=0.041)) and in the AR group (Pre: $r = 0.278$ (p=0.210), Post: $r = 0.482$ (p=0.023)), but no significance was found in the Doll group (Pre: 0.069 (p=0.792), Post: $r = 0.117$ (p=0.654)).

Feasibility and acceptability

In response to the prompt “The training time was adequate.”, 95% of the participants in the AR group answered "strongly agree" or "agree", in contrast to 59% of the participants in the Doll group. In response to “I felt as if I was seeing the actual patient in front of me.”, 90% of the participants in the AR group and 70% of the participants in the Doll group answered "strongly agree" or "agree".

In the open-ended response after the training, participants in the AR group expressed that they became more aware of the importance of continuous communication, e.g. making eye contact and speaking with the patient. They stated, "I thought it was good that the system let me know when I was not communicating well during care", "The system told me when my eye contact was even a little off, so I realized that not actually making eye contact even though I thought I was", "I became more conscious of making eye contact with patients and communicating more during care", and "I realized that communication during care is important." The responses to the general impact of the AR system were largely positive. For example, one participant wrote, "It was good that the system informed me when I was not communicating with the patient." They also shared what they would like to see improved for individual practice; for example, "The patient’s only utterance was laughter, and it would have been nice to have a few more statements or changes in emotion," "It was good to be able to check the evaluation in real-time but I wish there had been some model answers or hints on how to get a higher evaluation in learning."

Discussion

In this study, the impact of the AR-based affective training for nursing students was evaluated by means of a two-group randomized controlled study. The overall results indicate that the system had a more of a positive impact on the participants in the AR group than those trained by the existing (Doll) method. In particular, the frequency of the patient’s faces appeared in the first-person video, the occurrence of smaller face-to-face distances, and the occurrence of eye contact with the patient significantly increased after the training, while they did not increase significantly in the Doll group. These parameters indicated that the participants adjusted themselves to look at the patients’ faces and eyes more frequently at a closer distance, which is one of the important skills in Humanitude skill system and dementia care. At the same time, the face-to-face yaw angle was also significantly decreased in the AR group, which means the participants tended to looked at the patient’s face in the frontal direction. This is another important skill factor in dementia care as PwDs become difficult to identify from non-frontal faces[34]. These findings demonstrate that the AR-based affective training is effective for learning physical skills needed in dementia care, such as making eye contact, decreasing face-to-face distance, and facing the direction of the patient.

Interestingly, although the system was designed so that participants were in the role of a caregiver rather that PwD, the empathy score significantly increased in the AR group while it did not in the Doll group. This may be because the AR group received affective feedback from the avatar through positive and negative facial expressions and laughter, whereas the Doll group did not receive any feedback. The findings indicate that compared with the conventional training method, the AR-based skill training system had a greater impact on the participants’ acquisition of both physical and affective skills for dementia care.

The existing VR systems which enabled participants to experience life as PwD increased their empathy through a first-person experience. Our proposed system also increased the participants empathy even though it was a third-person experience because they could empathize with the avatars’ responses, e.g., smiling, anger. We believe that the participants’ empathy for PwD can be further increased by using both types of systems at the same time.

Another interesting finding is the correlation between extraversion and eye contact. In the first meeting with the simulated patient (Pre-evaluation), no significance was found, whereas a significant correlation was found in the AR group in the second meeting (Post-evaluation). This indicates a positive relationship exists between extraversion and eye contact. Furthermore, the affective communication training increased eye contact, and participants with higher extraversion scores could effectively improve their eye contact skill by using the AR system.

Limitations and Future work

Caregivers' facial expressions are important when they communicate with PwDs; however, our system does not evaluate the participants' facial expressions. Similarly, the system does not evaluate the caregiver's physical touch with the patient or the content of the caregiver's speech. The caregiver's facial expressions, touch, and content of their speech will need to be integrated in the future to evaluate the quality of their caregiving skills for people with PwD. In this study, the participants were nursing students, but it is not clear whether the same results can be obtained with more experienced caregivers.

The avatar used in the study had a cartoon-like face because its facial expressions could be noticed easily by novices (nursing students), so we thought it was more appropriate. The relationship between the facial rendering style (cartoon-like or realistic) and the outcome may also need to be investigated in the future.

References

1. World Health Organization *et al.* Dementia: Fact sheet N 362 (2012).
2. Larson, E. B., Yaffe, K. & Langa, K. M. New insights into the dementia epidemic. *New Engl. J. Medicine* **369**, 2275–2277 (2013).
3. Boseley, S. Dementia research funding to more than double to £66m by 2015. *The Guard.* (2012).
4. Cabinet Office, Government of Japan. Aged society – white paper (2017). [Online; accessed 16-Dec-2017].
5. Biquand, S. & Zittel, B. Care giving and nursing, work conditions and humanitude®. *Work* **41**, 1828–1831 (2012).
6. Kobayashi, M. & Honda, M. The effect of a multimodal comprehensive care methodology for family caregivers of people with dementia. *BMC Geriatr* **21**, 434 (2021).
7. Machiels, M., Metzelthin, S. F., Hamers, J. P. & Zwakhalen, S. M. Interventions to improve communication between people with dementia and nursing staff during daily nursing care: A systematic review. *Int J Nurs Stud* **66**, 37–46 (2017).
8. Morris, L., Horne, M., McEvoy, P. & Williamson, T. Communication training interventions for family and professional carers of people living with dementia: a systematic review of effectiveness, acceptability and conceptual basis. *Aging Ment Heal.* **22**, 863–880 (2018).
9. Eggenberger, E., Heimerl, K. & Bennett, M. I. Communication skills training in dementia care: a systematic review of effectiveness, training content, and didactic methods in different care settings. *Int Psychogeriatr* **25**, 345–358 (2013).
10. Hattink, B. J. *et al.* [Experiencing dementia: evaluation of Into D'mentia]. *Tijdschr Gerontol Geriatr* **46**, 262–281 (2015).
11. Wijma, E. M., Veerbeek, M. A., Prins, M., Pot, A. M. & Willemse, B. M. A virtual reality intervention to improve the understanding and empathy for people with dementia in informal caregivers: results of a pilot study. *Aging Ment Heal.* **22**, 1115–1123 (2018).
12. Slater, P., Hasson, F., Gillen, P., Gallen, A. & Parlour, R. Virtual simulation training: Imaged experience of dementia. *Int J Older People Nurs* **14**, e12243 (2019).
13. Jütten, L. H., Mark, R. E. & Sitskoorn, M. M. Can the Mixed Virtual Reality Simulator Into D'mentia Enhance Empathy and Understanding and Decrease Burden in Informal Dementia Caregivers? *Dement Geriatr Cogn Dis Extra* **8**, 453–466 (2018).
14. Jütten, L. H. *et al.* Testing the effectivity of the mixed virtual reality training into d'mentia for informal caregivers of people with dementia: protocol for a longitudinal, quasi-experimental study. *BMJ open* **7**, e015702 (2017).
15. Farina, L., Asnicar, J., Chan, O. H. & Pachana, N. A. Getting to know you but also to know me: changes in self-perceptions of aging among aged care workers after a Virtual Dementia Tour. *Int Psychogeriatr* **33**, 97–98 (2021).
16. Veerbeek, M., Willemse, B., Prins, M. & Pot, A. Development of a virtual reality-experience to improve empathy in caregivers of people with dementia. *Gerontologist* **56**, 538–538 (2016).
17. DEVELOPMENT OF A VIRTUAL REALITY-EXPERIENCE TO IMPROVE EMPATHY IN CAREGIVERS OF PEOPLE WITH DEMENTIA. *The Gerontol.* **56**, 538–538, DOI: 10.1093/geront/gnw162.2172 (2016). https://academic.oup.com/gerontologist/article-pdf/56/Suppl_3/538/7933268/gnw162.2172.pdf.

18. Chen, F. Q. *et al.* Effectiveness of Virtual Reality in Nursing Education: Meta-Analysis. *J Med Internet Res* **22**, e18290 (2020).

19. Jones, C., Jones, D. & Moro, C. Use of virtual and augmented reality-based interventions in health education to improve dementia knowledge and attitudes: an integrative review. *BMJ Open* **11**, e053616 (2021).

20. Dean, S., Halpern, J., McAllister, M. & Lazenby, M. Nursing education, virtual reality and empathy? *Nurs Open* **7**, 2056–2059 (2020).

21. Plotzky, C. *et al.* Virtual reality simulations in nurse education: A systematic mapping review. *Nurse Educ Today* **101**, 104868 (2021).

22. Elliman, J., Loizou, M. & Loizides, F. Virtual reality simulation training for student nurse education. In *2016 8th international conference on games and virtual worlds for serious applications (VS-games)*, 1–2 (IEEE, 2016).

23. UBISIM. https://www.ubisimvr.com/. [Online; accessed 26-Apr-2022].

24. Gineste, Y. & Pellissier, J. *Humanitude: comprendre la vieillesse, prendre soin des hommes vieux* (A. Colin, 2007).

25. Honda, M., Ito, M., Ishikawa, S., Takebayashi, Y. & Tierney, L. Reduction of behavioral psychological symptoms of dementia by multimodal comprehensive care for vulnerable geriatric patients in an acute care hospital: A case series. *Case reports medicine* **2016** (2016).

26. Sumioka, H., Shiomi, M., Honda, M. & Nakazawa, A. Technical challenges for smooth interaction with seniors with dementia: Lessons from humanitude™. *Front. Robotics AI* **8**, DOI: 10.3389/frobt.2021.650906 (2021).

27. Namikawa, T. *et al.* Development of a short form of the japanese big- five scale, and a test of its reliability and validity. *The Jpn. J. Psychol.* **83**, 91–99, DOI: 10.4992/jjpsy.83.91 (2012).

28. Fields, S. K. *et al.* Measuring empathy in healthcare profession students using the Jefferson Scale of Physician Empathy: health provider–student version. *J Interprof Care* **25**, 287–293 (2011).

29. Kurazume, R. *et al.* Development of ar training systems for humanitude dementia care. *Adv. Robotics* **36**, 344–358, DOI: 10.1080/01691864.2021.2017342 (2022).

30. Nakazawa, A. *et al.* First-person video analysis for evaluating skill level in the humanitude tender-care technique. *J. Intell. & Robotic Syst.* **98**, 103–118 (2020).

31. Fischler, M. A. & Bolles, R. C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**, 381–395, DOI: 10.1145/358669.358692 (1981).

32. Chong, E. *et al.* Detection of eye contact with deep neural networks is as accurate as human experts. *Nat. communications* **11**, 1–10 (2020).

33. eye-contact-cnn. https://github.com/rehg-lab/eye-contact-cnn (2020). [Online; accessed 26-Apr-2022].

34. Adduri, C. A. & Marotta, J. J. Mental rotation of faces in healthy aging and alzheimer's disease. *PLoS One* **4**, e6120 (2009).

Acknowledgements

This work was supported by JST CREST Grant Number JPMJCR17A5.

Author contributions statement

A.N. conceived the experiments, A.N., R.K, M.H. and M.N. made the experimental design, A.N., M.I. and M.H. conducted the experiments. R.K. developed the AR system, A.N., M.I. and M.K. analyzed the results. All authors reviewed the manuscript.

Additional information

The corresponding author is responsible for submitting a competing interests statement on behalf of all authors of the paper.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due to the privacy/ethical restrictions, but are available from the corresponding author on reasonable request.