2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 979-8-3503-2891-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR-Adjunct60411.2023.00087

Stop Bad Real-time Feedback!: Estimation of the Timing of Feedback that Negatively Impacts Presenters for Presentation Training in Virtual Reality

Yuichiro Fujimoto*
Nara Institute of Science and Technology

Zhou Hangyu†
Nara Institute of Science and Technology

Taishi Sawabe‡
Nara Institute of Science and Technology

Masayuki Kanbara§
Nara Institute of Science and Technology

Hirokazu Kato¶
Nara Institute of Science and Technology

Abstract

Training in virtual reality (VR) using a head-mounted display (HMD) has been applied to various fields, and recent attention has been focused on real-time feedback to the trainee. However, trainees' acceptance of the presented feedback is dependent on the state of the ongoing task as well as their internal state. If the trainees are not ready to accept the feedback from the system, it may potentially disturb the ongoing training. This study is premised on it being desirable for a VR training system to recognize the trainee's state and present feedback at an appropriate time. As a first step, this study focused on using a case scenario about public speaking training to confirm the appropriate timing of feedback in terms of trainees' subjective impressions. Specifically, a presentation training system that uses VR-HMD to collect observable multimodal information (including body posture and movement, gaze, heartbeat, perspiration, voice, and slide operation) was developed, which provided real-time feedback using icons. We collected the trainee's subjective scores of disturbance for each feedback in various situations through a Wizard of Oz experiment. Data analysis revealed two things; (1) Real-time feedback negatively affects the trainee's presentation in as many as 30 % of cases, (2) More than half of such negating-impact feedback can be estimated from observable information. This finding is expected to contribute to the development of future training systems that provide real-time feedback avoiding inappropriate timing for the trainee.

Index Terms: Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality

1 Introduction

With the advent of head-mounted displays (HMDs) training systems in virtual spaces have attracted extensive attention in diverse fields recently. The virtual reality (VR) training system can capture head movements, controller operations, voice, and gaze information. With the captured information, the system is potentially able to estimate user performance and provide interactive feedback to the trainee. To present effective feedback, "what" and "how" to provide feedback are important. This has been the subject of numerous studies (e.g., [4, 28]). Instead, this study focuses on the timing (the "when") of feedback presentation. In terms of timing, feedback can be broadly divided into real-time feedback (concurrent feedback), which is presented during training, and summary feedback (terminal feedback), which is presented after training or during short breaks in the training [20]. Summary feedback can provide broad suggestions for overall performance. Real-time feedback is expected to be more effective in improving micro-operations because it can point out problems as soon as a system discovers them.

This study focuses on a public speaking training as a case study scenario using real-time feedback. We assumed a case in which a user (trainee) wearing an HMD gives a presentation to a virtual audience in a virtual space, and the system points out problems in the user's current performance that need to be improved with the corresponding icon. The positive aspects of real-time feedback in VR training systems have been attracting attention [28]. However, if most of user's cognitive resources are devoted to an ongoing task, displaying additional real-time feedback and prompting for its confirmation may inhibit the ongoing task [16] (i.e., the presentation in this study). To the best of our knowledge, there are no studies that focus on the negative aspects of real-time feedback and the solutions for that in the VR presentation training scenario.

Therefore, the first research question in this study was defined as "To what extent does real-time feedback negatively affect the presentation in VR training?" The secondary research question was also defined: "Is it possible to estimate the feedback at the timing that could negatively affect the presentation from the observable information up to that timing?". A presentation training system using a VR-HMD to collect multimodal information (including body posture/movement, gaze, heart activity, perspiration, voice and operating the presentation slides) was developed for this study to assist with this determination. We collected the participants' subjective scores on behavioral change and disturbance feeling for each feedback in various situations through a Wizard of Oz (WoZ) experiment with this system. As a result of the data analysis, we confirmed that real-time feedback occasionally had negative effects on presenters. Furthermore, we confirmed the possibility of estimating more than half of the feedback that causes such negative effects from the multimodal data. These findings lead the development of future training systems that provide real-time feedback avoiding inappropriate timing for the trainee.

2 Related Work

2.1 Interruptibility of Information Presentation

The acceptability of a source's information presentation is largely determined by the properties of the information (e.g., importance and urgency) and the properties of the receiver (interruptibility) [32]. As for the receiver property, an example may be that a person is irritated to receive a phone call at a time they are absorbed in watching a World Cup soccer game on TV in their living room because their attention is already focused elsewhere. On the other hand, a person may appreciate (and not at all feel annoyed) a phone call when they are bored with nothing to do in their spare time. McFarlane proposed a taxonomy of eight perspectives for human interruption [25].

Office workers have a variety of tasks and often multitask while switching between them; as a result, there are many studies on

*e-mail: yfujimoto@is.naist.jp
†e-mail: zhou.hangyu.zh1,t@is.naist.jp
‡e-mail: t.sawabe@is.naist.jp
§e-mail: kanbara@is.naist.jp
¶e-mail: kato@is.naist.jp

DOI 10.1109/ISMAR-Adjunct60411.2023.00087

the effects of such task-switching (e.g., [18, 19]). Conscious task switching by oneself in multitasking is also called self-interruption, and excessive switching has been shown to reduce intellectual productivity. The target situation in this study of receiving feedback during a presentation can also be interpreted as multitasking, where the main task is the presentation and the secondary task is feedback confirmation.

2.2 Estimation of Appropriate Timing of Information Presentation

The effects of the timing of presenting information to people have been widely studied in the field of HCI. Mark et al. pointed out that inappropriate interruptions from human-to-human communication via computers reduce the productivity of workers engaged in intellectual activities [23]. For example, when a worker is concentrating on his/her office work, a ringing phone, a colleague talking to him/her, or excessive pop-ups of email software can interrupt work and increase a person's mental workload [1, 22]. To address this problem, Hudson et al. estimated interruptibility based on the idea that if the system can recognize the state of the worker, it can control the timing of information presentation [6, 14]. One another mainstream idea is that moments of low cognitive load (referring to the total amount of mental effort used in working memory [30]) are more likely to interrupt an ongoing task [26]. Therefore, cognitive load at each timing of the system has been studied using physiological signals that reflect mental workload, such as pupil size [7, 15] and electroencephalogram [24]. Iqbal and Bailey suggested that the mental cost of interruption depends on the stage of the ongoing task and that task models are able to predict the moment of possible interruption [16].

In recent years, a new trend emerges to study the appropriate timing and method of information presentation in situations where others in the real world talk to the VR-user [9, 12]. However, to the best of our knowledge, no study to date has investigated the effects of timing in the presentation of realtime feedback information in training scenarios.

2.3 System Feedback in Public Speaking

A number of systems (e.g., [10, 21, 31]) have been proposed for VR public speaking, both research and commercial products, using summary feedback. Palmas et al. compared the acceptability of real-time feedback using icons with indirect feedback through the behavior of a virtual audience in training with VR-HMDs to see differences in their training effects [28].

Chollet et al. compared real-time feedback and summary feedback conditions in a VR system and found that real-time feedback was more effective in maintaining subjects' motivation [3]. In their survey paper, Hatala et al. discuss the advantages and disadvantages of real-time feedback and summary feedback in different tasks, drawing on the findings of various papers [13]. Yadav et al. analyzed the relationship between anxiety at each time point and various physiological index data, which was collected by having participants speak in public [36]. According to their future work, they plan to predict the degree of anxiety from observable information and use it for real-time feedback. The final purpose of our study is similar to Yadav et al.'s in terms of controlling the real-time feedback, although we do not focus on anxiety reduction.

2.4 Position of This Work

In summary, there already exist studies that compare real-time feedback with summary feedback in training (e.g., [3,11,29,33]). However, if our target scenario is regarded as multitasking, the trainee's internal state can change within a short time (e.g., several seconds), as in the office work [15]. Based on this idea, we focus solely on determining if "current time" is the right time for feedback or not in the presentation training in the real-time feedback.

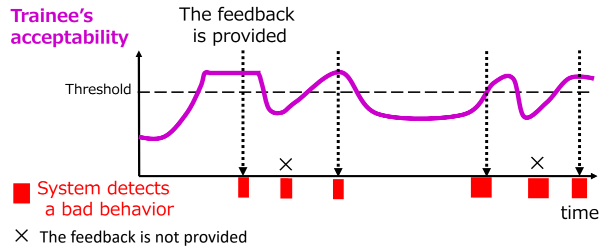

Figure 1: Ideal system behavior to the user's acceptability.

3 Methods

3.1 Expected Behavior of Final System

Figure 1 shows a simplified representation of the envisioned system behavior. The system is assumed to have a function that can sequentially observe the trainee's behavior during the presentation and detect a point that needs improvement (e.g., "You are looking only at the slides instead of at the audience!"). Once the system detects points in need of improvement, it then estimates the user's acceptance of the information presentation. If the acceptance is low, the system will not provide feedback. Then, when the trainee is determined to have a high level of acceptance of feedback, and if the points of improvement are still evident, this feedback is presented to the trainee. We believe that such a system will increase the effectiveness of training. On the other hand, if a bad point is found but feedback is not presented repeatedly, that would be also undesirable in terms of training effects. In this study, however, the first step is to focus on avoiding feedback at times that could negatively affect the presentation (i.e., feedback at times of low acceptability).

3.2 User's Internal Model in Training with VR System

The ultimate goal of the system is the trainee's acquisition of ideal presentation skills. Many skills are not acquired in a single training session and continued training is important for skill improvement. Recently, Fussell et al. proposed a new model for continuous learning with VR systems based on the technology acceptance model (TAM) [5], which is tied to the final skill improvement [8]. In this model, "Attitude toward use" is directly related to the trainee's continuity of training (the skill improvement), and it is related to "perceived usefulness" and "perceived ease of use" for the training system. "Perceived usefulness" is defined as "the degree to which a user believes that using VR for the training would enhance his or her performance". "Perceived ease of use" is defined as "the degree to which a user believes that using VR for the training would be free of effort". This means not only the ease of understanding and operating the system but also the low mental load involved in using the system.

Based on the model, we newly set the possible factors of real-time feedback which are related to both. "Perceived usefulness" is expected to be related to "perceived behavior changing", which is "the degree to which the user believes that the feedback has improved his/her subsequent behavior". This can be also seen as a subjective, short-term training effect. To keep the "perceived ease of use", feedback should not negatively affect the presentation. This was defined as the "disturbance level for feedback", which was "the degree to which the ongoing presentation activity is disturbed by feedback". These two subjective values would be affected by the timing of the feedback. A tentative model summarizing these relationships is shown in Figure 2.

3.3 Observable Information

The experiment below confirms the possibility of estimation of subjective values described in Section 3.2 from observable information.

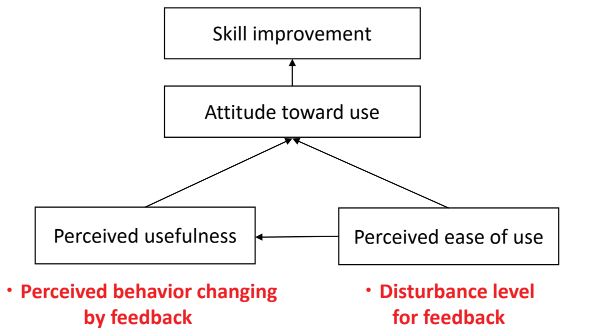

Figure 2: Expected training effect model in real-time feedback system for public speaking. Red factors are employed in this study. Black parts come from the model in [8].

Although many studies have examined the predictability of performance and anxiety level during presentations (e.g., [34–36]), there is no other study that has evaluated the impression for each of real-time feedback separately. Therefore, taking into account the information used in the above studies, this study collected multi-modal information including audio, body movement, gaze, and physiological data.

4 Experimental System

An experimental system was designed and implemented to analyze if the realtime feedback negatively affects participants' presentation and to collect multimodal data at the feedback moment.

4.1 Overview

The system simulated an environment where a presentation is made in front of audience with three virtual people, using a VR-HMD. Figure 3 shows a user's view of the virtual environment. Presentation slides are displayed behind the user, below eyes. The user can switch between slides by clicking the VR controller in his or her hand. In addition, the light beams emitted from the tip of the VR controller can be used as a pointer during the presentation.

Although many methods have been studied to automatically determine points for presentation improvement from observed information, the current system's accuracy is often inferior to human judgment. Considering this, we elected to adopt a WoZ method in which a human evaluator points out areas that can be improved by the presenter. Specifically, the system was designed so that the trainee receives feedback from the evaluator by keyboard input when he or she finds something that can be improved.

When a key is pressed, a corresponding icon is presented to the trainee after a delay selected from 0 to 5 seconds, based on a uniform distribution to collect various timing data. After having given their presentation, the trainees gave a score to each piece of feedback received. In the experiment, body movements, eye gaze, voice, heart rate, perspiration, and slide-switching information were continually recorded.

4.2 Hardware and Software

The system consisted of a personal computer (PC), a HMD (HTC VIVE Pro Eye (Dual OLED, 1440 × 1600 pixel for each eye)) with eye tracker, Microsoft Azure Kinect that measures body information, and E4 wristband to record heart rate (HR) and electrodermal activity (EDA) data. All system components ran in Unity 2019.2.10f1 on a single PC (OS: Windows10 Home, CPU: AMD Ryzen 5 3600 6-Core Processor, RAM: 64GB, GPU: NVIDIA GeForce RTX 3060). The E4 wristband data at a rate of 60 times

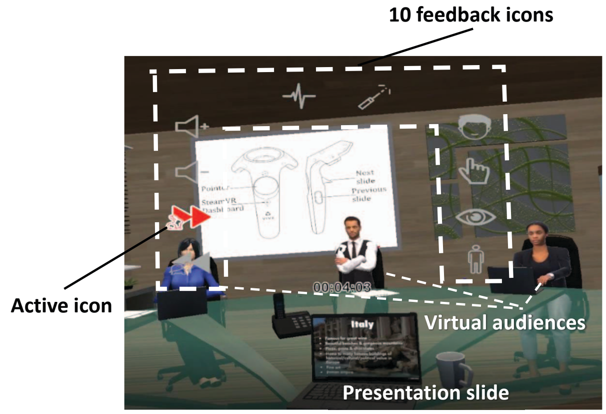

Figure 3: User's view in the training environment

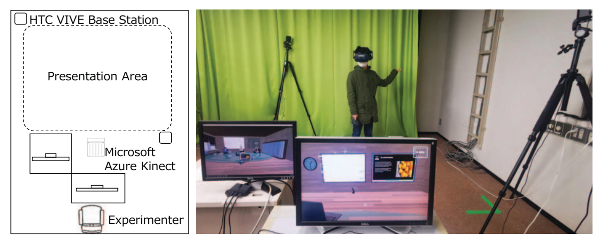

Figure 4: System configuration and the experimenter view

per second and sent the data to a dedicated online storage after it was saved to the device. The system layout is shown in Figure 4.

4.3 Feedback Method

The method of feedback also has a significant impact on a trainee's acceptance of the feedback. In this study, to focus only on the timing of feedback, the method was limited to one type – a simple icon, similar to Palmas et al. [28]. According to the Presentation Evaluation Metrics of Human Global Communications Co., Ltd, [1], ten types of icons were defined (Table 1). These can be roughly divided into three categories: (1) Voice, including increasing and decreasing loudness, depending on the loudness of the voice and the speed of speaking, respectively; (2) Physical behavior, including the presence of smiling, eye contact, hand movements, and body posture; and (3) Presentation skills, including appropriateness of inflection in speech, and appropriateness of pointer use.

Real-time feedback would be counterproductive if it is constantly and excessively interfered with by ongoing training. On the other hand, however, feedback should always be strong enough for the user to notice it. The intensity parameters of icon presentation include size, presentation position, presentation time, and the method of stimulus change during the presentation. Considering this, the following final parameters are determined through repetitive preliminary experiments and adjustments by three graduates. As shown in Figure 3, these 10 icons are always displayed in gray in the HMD coordinate system, in a position that avoids the central field of view (such as [17]). The icon flashed red twice and then remained red for five seconds to indicate to the user that a problem related to the icon had been found. After five seconds, the icon color returned to gray and the next keystroke was accepted. No

[1] https://human-gc.jp/english/

more than two icons were presented simultaneously to avoid confusing the trainee.

4.4 Virtual Audience

To increase the realism of the presentation environment, three virtual people were placed as audience [27]. The virtual audience was designed to randomly take general listening behaviors (nodding, taking notes, etc.). Two software platforms were used to create the virtual audience: Greta [2] and RENDERPEOPLE [2].

4.5 Recorded Data

Six types of information were recorded during the experiment in the following format:

(1) Audio was recorded as a common wav file.

(2) Body posture and movements were recorded by Microsoft's Azure Kinect at 30 fps as the 3D position and rotation in the Kinect depth camera coordinate system for 25 joints, excluding joints on the face.

(3) Eye gaze was calculated as the vector formed by the approximate intersection of the right eye and left eye gaze directions and the midpoint of the right eye and left eye, by the eye camera SDK in Vive Pro Eye, and was recorded at 60 fps.

(4) Blood Volume Pulse (BVP) was recorded at 64 fps by the E4 wristband. In addition, Electrodermal Activity (EDA) was recorded at 4 fps by the same device.

(5) Slide switching is performed by the key operation of the VR controller in the user's hand. All keystroke timings and page numbers were recorded.

(6) As for the user's property, the user was asked to answer a questionnaire before the experiment to indicate his/her anxiety and skill of public speaking as a subjective value on three levels, respectively.

In addition, the time when the evaluator pushed the key to present a feedback icon, the time when the icon actually started to be presented after the delay, and the type of icon were also recorded, respectively.

5 Experiment

This section describes an experiment using the system in Section 4 to observe how the real-time feedback negatively affects the ongoing presentation and to collect data to confirm the relationship between trainees' impression scores for each feedback and the observable information. The former was evaluated qualitatively by means of recorded video and audio.

Table 1: Icon list

Category	Name	Icon	Description
Voice	VoiceLoud		Speak louder.
	VoiceQuiet		Speak in a low voice.
	VoiceFast		Speak faster.
	VoiceSlow		Speak slower.
Physical behavior	Smile		Act naturally, smile.
	EyeContact		Look at the audience frequently.
	Gesture		Use meaningful gestures. Do not make distracting movements.
	Posture		Stand naturally and comfortably.
Presentation Skill	Rhythm		Stress key words and words that show relationships.
	Pointer		Use the pointer appropriately.

[2]https://renderpeople.com/free-3d-people/

5.1 Overview

Participants wore the HMD and gave presentations in a virtual space. The evaluator (experimenter) always observed the presentation, and if she found behavior to be improved, she made keystrokes corresponding to each feedback icon (Figure 4: Experimenter's viewpoint during the experiment). Icons were presented in front of the participants' view, and based on these, they were asked to improve their behavior without stopping their presentations. Immediately following their presentations, participants were shown a video recording of their views in the virtual environment and asked to respond to their scores for each piece of feedback received. After the experiment, the impact of the feedback was qualitatively confirmed by observing the presenter's behavior and voice immediately after the real-time feedback in the recorded video. In addition, the relationship between each subjective score and the observable information was analyzed to ascertain if it is possible to estimate undesirable timing for participants.

5.2 Participants

Eighteen participants (all graduate students in information science and materials science, all non-native English speakers, 8 males and 10 females, aged 27 to 37 years ($M = 28.61, SD = 4.30$) at a local university were recruited by e-mail to participate in this study. None spoke English natively, although all participants' level of English were above the pre-advanced: I speak and understand well but still make mistakes and sometimes people do not understand me clearly. Four participants reported low anxiety levels when giving presentations, six said they experienced medium anxiety, and eight experienced high anxiety, according to the results of the questionnaire completed before the presentation. Each participant was paid an honorarium of 2,000 JPY after two hours of work in the experiment.

5.3 Procedures

First, the experimenter provided an overview of the experiment to each participant and obtained informed consent from each participant. The 5 minutes baseline BVP and EDA of participants are recorded first. Participants were asked to sit on a chair and wear an HMD and an E4 wristband. A high-fidelity meditation platform called Meditation VR [3] in Steam was used to help participants relax. After this, each participant was taught how to use the system and how to give their presentation using this system. Then, the eye tracker was calibrated and the experimenter ensured each participant had time to get familiarize themselves with the system. Participants were especially asked to remember the meaning of each feedback icon.

After setting up the system, participants are asked to choose one slide (from an option of 10 slides) across five topics (foods, lifestyle, spiritual, sports, travel). They were each given 10 minutes to organize their presentations without wearing HMD, and five minutes free to practice presenting while wearing the HMD. After this initial training, the participants began their presentations and the system began recording the data when the experimenter signaled the start of the training session. The evaluator (experimenter) observed the situation and, following the presentation metric, made keystrokes corresponding to each piece of feedback once she found improvement points. Each presentation was five to seven minutes long.

Immediately after each presentation, participants responded to each feedback with subjective scores in Section 5.4 while watching

[3]https://store.steampowered.com/app/1301850/Meditation_VR/

their own viewpoint videos during the training. After a 10-minute break, the above procedure was repeated with another presentation topic. Finally, a brief interview was conducted to obtain opinions on the feedback method and the training effectiveness of the system.

5.4 Subjective Scores

The scores to which the participants responded for each feedback after each presentation are listed below. These correspond to the elements discussed in section 3.2. Each score was answered on a five-point scale.

(1) Perceived behavior changing (PBC)： How much did you change your behavior according to the feedback you received? 1: Nothing changed. 3: I changed my behavior once, but forgot about it soon. 5: I changed my behavior immediately, and kept it in mind for the whole presentation.

(2) Disturbance level for feedback (DL): How much did this feedback disturb or distract you? 1: Not disturbed/distracted at all. 3: It disturbed/distracted me little, but I can deal with it. 5: It disturbed/distracted me a lot, and even interrupted the presentation for a while.

Note that PBC is the effect of feedback until the end of the presentation, and DL is the impression of feedback immediately after receiving it.

6 Results

The total number of feedbacks presented to all participants was 228. Except for nine unnoticed feedbacks and four feedbacks with eye tracker errors, the remaining 215 were analyzed.

6.1 Negative Impact of Feedback on Presentation

We qualitatively analyzed the observable disturbing effect of the feedback on the ongoing presentation. The experimenter checked observable changes in the presentation for 15 seconds after each of the feedback with the recorded audio and video. The following three main negative effects were observed. The first is the case that the presenter suddenly stops speaking for more than three seconds immediately after presenting feedback (even though it is not during the slide changeover, etc.). The sudden speech stop gives a rather negative impression to the audience when it occurs in a real presentation. Such cases were observed 35 times (/ 215). The second is the case of a clear increase in filler words (e.g., Uh, well). Due to individual differences in the frequency of filler word use, we considered a negative case to be one in which the use of filler words increased more than twice in the 15 seconds immediately after the feedback, compared to the 15 seconds before the feedback. Such cases were observed 33 times (/ 215). The third was the repetition of the same word. We considered this case when the same two or more words were repeated within 15 seconds after the feedback. Such cases were observed 16 times. Some of these three cases occurred simultaneously, and the total number of times with any of the negative effects was 65 (/ 215). In other words, in as many as 30% of the cases, a negative impact (observable for the audience) was observed on the presenter immediately after the feedback.

The Pearson's correlation coefficient between its occurrence and the DL score was 0.36 ($p < 0.01$). There was no significant correlation between its occurrence and PBC.

6.2 Estimation of Perceived Behavior Change and Disturbance Level for Feedback

Simple models to estimate each score (i.e., PBC and DL) were created from all the recorded multi-modal data with the Random Forest. A 10-fold cross-validation was performed using the 215 data. For PBC, the correlation coefficient was 0.45. The correlation coefficient for DL was 0.56. On the other hand, adding the participants' property data (i.e., anxiety level as described in Section 4.5) and the icon type improved the correlation coefficient to 0.53 for PBC, and 0.62 for DL.

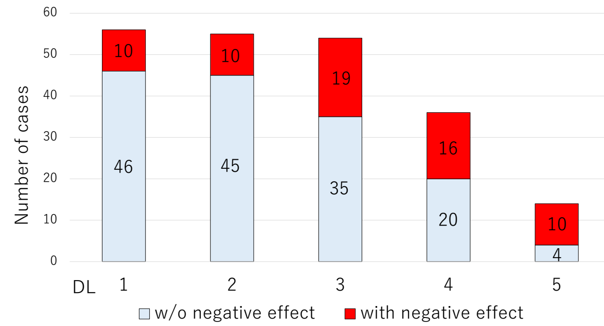

Figure 5: Number of cases with negative effects included in each disturbance level (DL)

6.3 Detection of Feedback Timing that Negatively Impacts Presenters

Finally, we describe the possibility of predicting feedback at times that could have a negative impact by using the DL values. Only DL was used here for estimation because PBC was not correlated with the occurrence of negatively affected feedback. Figure 5 shows the number of cases with negative effects included in each DL value. The higher the DL value, the higher the percentage of cases that had a negative impact at a given DL value. From this result, for example, if feedback is not displayed simply when the DL is greater than or equal to 3, 70 % ($(19+16+10)/65$) negative cases can be avoided. Similarly, by using the estimated DLs in Section 6.4, 62% negative cases ($40/65$) can be avoided. However, note that this is the result of fitting the model to the test data. The actual estimability is discussed further in Section 7.

7 Limitations and Future Work

At present, this study has a number of unexplored issues that leave gaps for future research.

First, all participants for data collection at this stage are graduate students, and the results obtained depend on them. It is essential to enrich the data by continuously conducting the same experiment on participants of various properties. The construction of a model for each participant is also a potential future issue.

Features calculated using the measured data are rudimentary and need to be replaced with more sophisticated ones (e.g., combined multiple data). For instance, the results showed that the timing of looking at the slides was undesirable in terms of DL; however, if the speaker is speaking without stuttering, the timing of looking at the slides may not result in lower DL. In this case, such timing could probably be detected by using a combination of head direction and voice information.

In addition, it must be verified soon whether withdrawing feedback at times that negatively affect presentation in training really leads to training effects (the reverse could also be true as another training strategy, e.g., overload training). We plan to conduct an experiment with the system that can provide real-time feedback at a time decided by DL to confirm its training effect as the next step.

8 Conclusion

This study confirmed the influence of timing of real-time feedback in the VR system for public speaking training through an experiment using the WoZ method. First, we confirmed that as much as 30% of real-time feedback had a negative effect on presentation. Next, referring to Fussell et al.'s model [8], we defined the disturbance level, etc. for each feedback and showed that it could be estimated with reasonable accuracy from the observable multimodal information up to the feedback timing. Furthermore, it was found that more than half of the feedbacks with negative effects can be determined from the estimated disturbance level. As next, we plan to conduct an experiment with the real-time feedback system with timing control by the estimated DL to confirm its training effect.

ACKNOWLEDGMENTS

This work was funded by JST CREST Grant Number JPMJCR19A, Japan.

REFERENCES

[1] S. R. Barley, D. E. Meyerson, and S. Grodal. E-mail as a source and symbol of stress. Organization Science, 22:887–906, 2011.

[2] C. Pelachaud. Greta: Realtime platform to control socio-emotional virtual characters. Accessed 16 Feb 2022.

[3] M. Chollet, S. Marsella, and S. Scherer. Training public speaking with virtual social interactions: effectiveness of real-time feedback and delayed feedback. Journal on Multimodal User Interfaces, 2021.

[4] I. Damian, C. S. S. Tanz, T. Baur, J. Schoningz, K. Luytenz, and E. Andr. Augmenting social interactions: Realtime behavioural feedback using social signal processing techniques. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 15), pp. 565–574, 2015.

[5] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw. User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35(8):982–1003, 1989.

[6] J. Fogarty, S. Hudson, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang. Predicting human interruptibility with sensors. ACM Transactions on Computer Human Interaction, 1:119–146, 2005.

[7] L. Fridman, B. Reimer, B. Mehler, and W. T. Freeman. Cognitive load estimation in the wild. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 18), pp. 1–9, 2018.

[8] S. G. Fussell and D. Truong. Using virtual reality for dynamic learning: an extended technology acceptance model. Virtual Reality, 26(1):249–267, 2021.

[9] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. Notifivr: Exploring interruptions and notifications in virtual reality. IEEE Transactions on Visualization and Computer Graphics, 24:1447–1456, 2018.

[10] S. V. Ginkel, J. Gulikers, H. Biemans, O. Noroozi, M. Roozen, T. Bos, R. van Tilborg, M. van Halteren, and M. Mulder. Fostering oral presentation competence through a virtual reality-based task for delivering feedback. Computers Education, 2019.

[11] S. V. Ginkel, D. Ruiz, A. Mononen, C. Karaman, A. de Keijzer, and J. Sitthiworachart. The impact of computer-mediated immediate feedback on developing oral presentation skills: An exploratory study in virtual reality. IEEE Transactions on Affective Computing, 2020.

[12] M. Gottsacker, N. Norouzi, K. Kim, G. Bruder, and G. Welch. Diegetic representations for seamless cross-reality interruptions. In Proceedings of the IEEE Virtual Reality and 3D User Interfaces, pp. 310–319, 2021.

[13] R. Hatala, D. A. Cook, B. Zendejas, S. J. Hamstra, and R. Brydges. Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. Advances in Health Sciences Education, 2014.

[14] S. Hudson, J. Fogarty, C. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. Lee, and J. Yang. Predicting human interruptibility with sensors: a wizard of oz feasibility. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 03), pp. 257–264, 2003.

[15] S. T. Iqbal, P. D. Adamczyk, X. S. Zheng, , and B. P. Bailey. Towards an index of opportunity: Understanding changes in mental workload during task execution. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 05), pp. 311–320, 2005.

[16] S. T. Iqbal and B. P. Bailey. Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 05), pp. 1489–1492, 2005.

[17] N. Janaka, C. Haigh, H. Kim, S. Zhang, and S. Zhao. Paracentral and near-peripheral visualizations: Towards attention- maintaining secondary information presentation on ohmds during in-person social interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 22), pp. 1–14, 2022.

[18] N. Janaka, S. Zhao, and S. Sapkota. Can icons outperform text? understanding the role of pictograms in ohmd notifications. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 23), pp. 1–23, 2023.

[19] I. Katidioti, J. P. Borst, and N. A. Taatgen. What happens when we switch tasks: Pupil dilation in multitasking. Journal of Experimental Psychology, 20:380—396, 2014.

[20] P. E. King, M. J. Young, and R. R. Behnke. Public speaking performance improvement as a function of information processing in immediate and delayed feedback interventions. Communication Education, 49:365–374, 2000.

[21] V. Ltd. Virtual speech, September 2017.

[22] G. Mark, V. Gonzalez, and J. Harris. No task left behind? examining the nature of fragmented. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 05), pp. 321–330, 2005.

[23] G. Mark, S. Voida, and A. Cardello. A pace not dictated by electrons: An empirical study of work without e-mail. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 12), pp. 555–564, 2012.

[24] S. Mathan, S. Whitlow, M. Dorneich, P. Ververs, and G. Davis. Neurophysiological estimation of interruptibility: Demonstrating feasibility in a field context. In Proceedings of the 4th International Conference of the Augmented Cognition Society, 2007.

[25] D. C. McFarlane. Interruption of people in human-computer interaction. PhD Thesis, George Washington University, 1998.

[26] Y. Miyata and D. A. Norman. The Control of Multiple Activities - in User Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Associates, 1986.

[27] F. Mostajeran, M. B. Balci, F. Steinicke, S. Kuhn, and J. Gallinat. The effects of virtual audience size on social anxiety during public speaking. In Proceedings of the IEEE Virtual Reality and 3D User Interfaces (IEEE VR 2020), pp. 303–312, 2020.

[28] F. Palmas, R. Reinelt, J. E. Cichor, D. A. Plecher, and G. Klinker. Virtual reality public speaking training: Experimental evaluation of direct feedback technology acceptance. In Proceedings of the IEEE Virtual Reality and 3D User Interfaces, pp. 463–472, 2021.

[29] R. Sigrist, G. Rauter, R. Riener, and P. Wolf. Terminal feedback outperforms concurrent visual, auditory, and haptic feedback in learning a complex rowing-type task. Journal of Motor Behavior, 2013.

[30] J. Sweller. Cognitive load during problem solving: Effects on learning. Cognitive Science, 12:257–285, 1988.

[31] V. H. Technologies. Virtual orator, 2015.

[32] M. Vastenburg, D. Keyson, and H. de Ridder. Considerate home notification systems: A field study of acceptability of notifications in the home. Personal and Ubiquitous Computing, 12:555—566, 2008.

[33] C. M. Walsh, S. C. Ling, C. S. Wang, and H. Carnahan. Concurrent versus terminal feedback: it may be better to wait. Academic Medicine, 2009.

[34] T. Wörtwein, M. Chollet, B. Schauerte, L.-P. Morency, R. Stiefelhagen, and S. Scherer. Multimodal public speaking performance assessment. In Proceedings of ACM on International Conference on Multimodal Interaction (ICMI '15), pp. 43–50, 2015.

[35] T. Wörtwein, L.-P. Morency, and S. Scherer. Automatic assessment and analysis of public speaking anxiety: A virtual audience case study. In Proceedings of International Conference on Affective Computing and Intelligent Interaction (ACII2015), pp. 187–193, 2015.

[36] M. Yadav, M. N. Sakib, E. H. Nirjhar, K. Feng, A. Behzadan, and T. Chaspari. Exploring individual differences of public speaking anxiety in real-life and virtual presentations. IEEE Transactions on Affective Computing, 2020.