Impression of a Job Interview training agent that gives rationalized feedback

Should Virtual Agent Give Advice with Rationale?

Nao Takeuchi
Graduate School of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
m1m21a29@st.oit.ac.jp

Tomoko Koda
Faculty of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
tomoko.koda@oit.ac.jp

ABSTRACT

The COVID-19 pandemic has had a significant socio-economic impact on the world. Specifically, social distancing has impacted many activities that were previously conducted face-to-face. One of these was the training that students receive for job interviews. Thus, we developed a job interview training system that will give students the ability to continue receiving this type of training. Our system recognized the nonverbal behaviors of an interviewee, namely gaze, facial expression, and posture and compares the recognition results with those of models of exemplary nonverbal behaviors of an interviewee. A virtual agent acted as an advisor gives feedback on the interviewee's behaviors that need improvement. In order to verify the effectiveness of the two kinds of feedback, namely, rationalized feedback (with quantitative recognition results) vs. non-rationalized one, we compared interviewees' impression. The results of the evaluation experiment indicated that the virtual agent with rationalized feedback was rated as more reliable but less friendly than the non-rationalized feedback.

CCS CONCEPTS

• Human-centered computing ~ Interaction design ~ Empirical studies in interaction design;

•Human-centered computing ~ Human computer interaction (HCI)~Empirical studies in HCI;

•Human-centered computing ~ Human computer interaction (HCI)~HCI design and evaluation methods ~ User models;

•Human-centered computing ~ Interaction design ~ Interaction design process and methods ~ User interface design;

MMAsia '21, December 1–3, 2021, Gold Coast, Australia

ACM ISBN 978-1-4503-8607-4/21/12…$15.00
https://doi.org/10.1145/3469877.3493598

KEYWORDS

job interview training, rationalized feedback, virtual agent nonverbal behaviors, multi-modal interaction, gaze recognition, posture recognition, facial expression recognition

ACM Reference format:

Nao Takeuchi and Tomoko Koda. 2021. Impression of a Job Interview training agent that gives rationalized feedback ~Should Virtual Agent Give Advice with Rationale?~: In Proceedings of ACM Multimedia Asia (ACMMMAsia'21). ACM, New York, NY, USA, 4 pages.
https://doi.org/10.1145/3469877.3493598.

1 Introduction

Form its onset in late 2019, COVID-19 has had a significant social and economic impact, worldwide, and people have been asked to avoid human contact as much as possible. As a result, the face-to-face training of interviewees, which could be useful in the employment search process, is now avoided. Interview training could help students acquire skills by experiencing the content and flow of job interviews and can increase their confidence in their search for employment. However, interview training has been limited due to the number of interviewers and time available for interviewing [1]. Moreover, the impact of COVID-19 has made interview training more difficult to conduct. This suggest that there is an increasing need for a system that allows students to train for job interviews independently.

There is a growing body of research demonstrating the power of the social signals that people consciously or unconsciously exhibit in a variety of situations, such as job interviews and group discussions. Visual nonverbal behavior during a dialog account for 55% of all the information conveyed [2]. Washburn et al. pointed out that the outcome of an interview is affected more by the nonverbal behaviors of an interviewee than their verbal behaviors [3]. Moreover, Arvey et al. noted that nonverbal behaviors such as gaze, body movements, and tone of voice greatly influence the interviewee's evaluation [4]. These studies show that the use of nonverbal behavior and its impact on job interview success has been a major focus in research.

In recent years, social signal processing (SSP) techniques using multimodal information have been used for dialog analysis [5] and have been applied to AI-based interview recruitment systems [6, 7, 8, 9] and interview training systems [10, 11, 12, 13]. Specifically, there are those that visualize the information of the nonverbal behavior and provide feedback during or after the interview [14, 15, 16, 17], and those that change the behavior of the interviewer, i.e., the virtual agent [18, 19, 20]. However, most of the research conducted in the field of social signal processing focused on the recognition of emotions based on speech and facial expressions and paid less attention to posture recognition. In addition, some studies [6, 7, 9] proposed high-end computer systems that are not affordable for general users.

It has also been reported that practicing interviews with a virtual agent as an interviewer was more effective in improving interviewees' skills and increasing self-disclosure compared to conventional methods [15,21,22]. In addition, virtual agents acting as interviewers and advisors to practice interviews have been reported to improve interview performance, interviewee employability, and reduce interview anxiety [17]. However, these studies focus on the effectiveness of using virtual agents, but not on the effectiveness of the dialogue strategy of the agent.

We focus on dialogue strategy to improve the effectiveness of the feedback. It has been suggested that dialogue strategies that clearly communicate what the problem is and how to act are more effective than indirect expressions in motivating people to improve their behavior [23]. In spoken interaction systems in driving simulators, it has been found that using dialogue strategies that show the rationale that people use to give advice is more effective for acceptance rates [24]. Since we use SSP to evaluate the interviewee's non-verbal behaviors, the system can provide quantitative evidence of the recognition results to the agent, and the agent can give rational feedback with quantitative evidences, i.e., "Your knees are 50cm apart." instead of giving non-rationale feedback, i.e., "Your knees are apart too much."

The aim of the study is to investigate the effectiveness of rationalized feedback from the agent in terms of user's impression on the agent, i.e., friendliness, performance, trust. This paper reports a preliminary result of our evaluation experiment that compared users' impression on rationalized feedback (with quantitative recognition results) vs. non-rationalized one.

2 Job Interview Training System

2.1 System Overview

The system was developed using Unity, FaceAPI [25], OpenPose [26], TobiiEyeTracker4C [27], and a webcam (Figure 1). This system consists of three phases: a demographic input phase, a mock interview phase, and a feedback phase. The demographic input phase was used to input the number of users and gender. During the mock interview phase the interview was video-recorded from a front-left angle in order to obtain the nonverbal behaviors of the interviewee, including gaze, facial expression, and posture. The captured video was analyzed by the following procedures (see 2.3) and played back in the feedback phase. The system paused the video where feedbacks were needed, and the virtual agent provided feedbacks on any points for improvement (Figure 2).

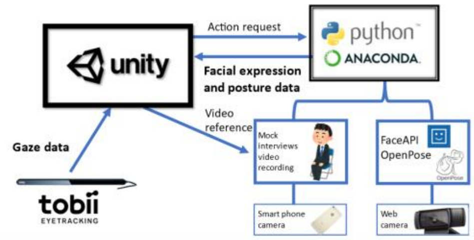

Figure 1: System Overview

Figure 2: Feedback phase
Left: Experimental set up; Right: View of the display on the right screen, during standby state (top) and feedback state (bottom)

2.2 Detectable Non-verbal Behaviors

The nonverbal behaviors were acquired at 1-second intervals for gaze and 3-second intervals for facial expression and posture. From the acquired behaviors, the following information was obtained: gaze rate, user's gaze moving off the interviewer's face for more than five seconds, number of times the gaze point moved to the upper right or upper left, level of smile or straight face,
six facial expressions (anger, contempt, disgust, fear, sadness, surprise), posture (forward and backward leaning), legs open, legs opening gradually, shake of the neck, and protrusion of the elbow.

2.3 Method of Detecting

2.3.1 Gaze Detection Method

We used a collision-detection method in order to detect inappropriate gazes. In order to determine when the interviewee's gaze moved off of the interviewer's face, we preset an area-of-interest (AOI) on the interviewer's entire face. The system determined inappropriate gaze when the interviewees gazing point moved out of the AOI for five seconds. The gaze rate was calculated by dividing the number of frames in which the interviewee was looking at the interviewer's face by the total number of frames and displaying it as a percentage. The number of times the gazing point moved to the upper right or upper left was determined by setting up another AOI in the upper right and upper left areas of the screen (next to the interviewer). The system determined inappropriate gaze if the gazing point entered these areas more than 10 times. Eventually, these metrics were used to determine the feedback given by the virtual agent during the feedback phase.

2.3.2 Facial Expression Detection Method

For inappropriate facial expression detections, we used "smile" and "emotion" from FaceAPI. The smile and straight face scores were set at 0 and 1, respectively. A smile was determined as a result when a score of 0.5 or more was detected, and a straight face was determined as the result when 0 was detected three times in a row.

2.3.3 Posture Detection Method

To detect posture, we observed whether the interviewee leaned forward or backward, opened their legs, opened their legs gradually, shook their neck, or protruded their elbows.

Inappropriate postures were detected by comparing a correct posture model with the posture of the interviewee. A correct posture model was created for each gender using OpenPose under the guidance of the Employment Department of our university. For example, whether an interviewee was leaning forward and backward was judged when there was a difference of more than 20 degrees between the model and the interviewee, legs in an open position was judged when the legs of the interviewee was wider than the width of model's legs, and protrusion of the elbow was judged when there was more than 15 degrees difference between the model and the interviewee.

2.4 Feedback Algorithm

While the video taken during the mock interview is played back, the system can pause the video at any time according to the feedback algorithm and a virtual agent provides feedback. The weight of this algorithm was set in the order of gaze, facial expression, and posture, based on the order of importance during the interview. For example, “At this time, your gaze was off the interviewer for a period of time. Let’s pay attention.”, “Your expression was stiff at this time.”, “At this time, your legs were gradually opening. Let's be careful.” etc.

3 Experiment

3.1 Overview of the Experiment

The purpose of the experiment was to compare the effectiveness of two types of feedback, rationalized and non-rationalized feedback, used by the virtual agent in a mock interview practice situation.

Participants used the system and received feedback from the virtual agent in the rationalized (WR hereafter) and non-rationalized conditions (NR hereafter) in a within-participant design.

Examples of NR feedback were “Aren't you opening your legs too much at this time?”, “At this time, your elbow is sticking out”. Examples of the WR feedback were, “Your legs are more than 30 cm apart, aren't your legs too open at this time?”, “At this time, your elbows are sticking out more than 25 degrees.” The participants answered a questionnaire on the impressions of the agent in both conditions. The experiment using human participants was approved by the Life Science Committee of our university.

3.2 Experimental Setup

The participants were 10 university and graduate students (8 males and 2 females, aged between 20 and 22). As evaluation criteria, Adjective Check List (ACL) for Interpersonal Cognition for Japanese [28] and a virtual agent impression evaluation questionnaire that we created ourselves with reference to [29].

ACL for Interpersonal Cognition for Japanese was rated by the 7-point SD method, and the impression evaluation questionnaire (shown in Table 1) were rated by the 7-point Likert scale (low 1−7 high). For the impression evaluation questionnaire, the 14 items in Table 1 were used.

Table 1: Impression evaluation questionnaire

Trust	Q1	I felt I could trust this virtual agent.
	Q2	I felt more comfortable with the presence of this virtual agent.
	Q3	I felt I could accept the advice and suggestions of this virtual agent.
	Q4	I felt that this virtual agent was speaking with intention.
Friendliness	Q5	This virtual agent frustrated me.
	Q6	I had a good feeling about this virtual agent.
	Q7	I fell I can get along with this virtual agent.
	Q8	I think I'll get tired of this virtual agents soon.
	Q9	I want this virtual agent to be like a family member or a best friend.
Performance	Q10	This virtual agent felt like it was working exactly the way someone had designed it to work.
	Q11	I think the presence and advice of this virtual agent will be a good practice.
	Q12	With this virtual agent, I think it will be more fun than if I were to practice alone.
	Q13	With this virtual agent, I think I'll feel less anxious than if I were to practice alone.
	Q14	I think I'd like to use this virtual agent myself.

4 Results

4.1. Results of Factorial Analysis

Factor analysis (FA hereafter) was conducted on the virtual agent’s impression ratings in order to extract the factors that composes our interpersonal impressions toward the virtual agent. The results of FA using the principal factor method extracted two factors (shown in Table2). The First factor was named as “Reliability factor” (composed of adjectives such as Unshy, Grand, Reasonable, Pertinent, and Positive), the second as “Annoyance factor” (composed of adjectives such as Hateful, Gloomy, and social). Cronbach's coefficients alpha for the factors were 0.71 for “Reliability factor”, and 0.64 for “Annoyance factor”, which showed high enough internal consistency of the extracted factors.

Table 2: Two factors and adjectives for interpersonal impressions

	Adjective-pair	Factor	
		1	2
Reliability factor	Unshy – Shy	.782	.137
	Grand – Servile	.711	-.237
	Reasonable – Unreasonable	.659	.399
	Pertinent – Impertinet	.544	.250
	Positive – Passive	.506	-.185
Annoyance factor	Hateful – Lovabie	-.337	.728
	Gloomy – Cheerful	.213	.649
	Hard-hearted – Soft-hearted	.128	.468
	Mature – Immature	.115	.175
	Unpleasant – Pleasant	-.218	.036
	Irresponsible – Responsible	.175	-.280
	Friendly – Unfriendly	.096	-.552
	Broad-minded – Narrow-minded	.008	-.871
	Incautious – Cautious	-.056	-.045
	Unsocial – Social	-.624	.071

4.2. ACL for Interpersonal Cognition for Japanese

Figure 3 shows the results of ACL. In Figure 3, for example, for the item "Negative - Proactive," the closer the score was to 1, the more "Negative," and the closer the score was to 7, the more "Proactive.

The results of Wilcoxon signed rank test showed that the WR condition showed a significantly higher "responsible" rating than the NR condition. In terms of “friendly”, “broad-minded” ratings (p=<0.05), the NR condition was rated significantly higher than the WR condition.

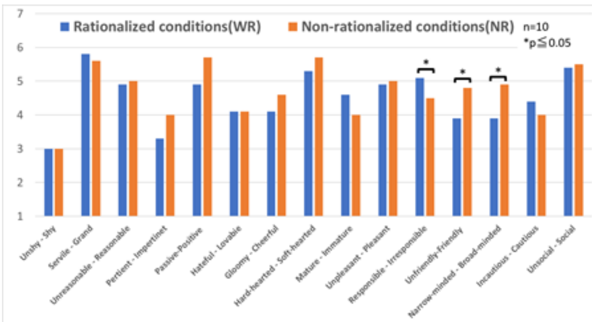

Figure 3: Results of ACL for Interpersonal Cognition for Japanese (left adjective is 1, right adjective is 7)

4.3. Impression Evaluation Questionnaire

Figure 4 shows the results of the impression evaluation questionnaire. The results of Wilcoxon signed rank test showed that the WR condition had a significantly higher rating of “I think the presence and advice of this virtual agent will be a good practice” than the NR condition, and a significant trend in the rating of “I felt I could trust this virtual agent.”.

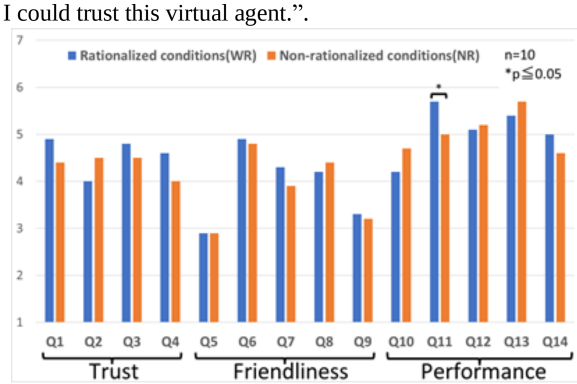

Figure 4: Results of the impression evaluation questionnaire

5 Discussion

Firstly, we discuss the result of the factor analysis. Two factors, “Reliability factor” and “Annoyance factor” were extracted. The “Reliability factor” is a factor that include adjectives such as reasonable, pertinent, and positive. The "Annoyance factor" indicates annoyance of the agent which is the opposite of friendliness. These results suggest that the main factor in forming interpersonal impressions of the agent in this experiment was reliability. This suggests the virtual agent that gives interview feedback gave an impression related to reliability followed by unfriendliness.

In addition, the adjectives that form the reliability factor suggest that the virtual agent was perceived as proactive, unafraid, and articulate, while the adjectives of the annoyance factor suggest the virtual agent caused frustration for the user. This frustration might be caused by accurate advice which was hard to accept by the participants. This suggests that the virtual agent was perceived as a proactive advisor that gives articulate feedback, sometimes too articulate to accept by the participants.

The results of the ACL for Interpersonal Cognition for Japanese suggest that WR was more unapproachable, narrow-minded, but responsible than NR. The results of the impression evaluation questionnaire suggest that WR was more trustworthy and more likely to lead to good practice than those in NR.

The results indicated that the virtual agent who provides rationalized feedback were rated higher in terms of its responsibility, trustworthiness and providing good service but lower in terms of its friendliness. These results were similar to the results of the factor analysis, suggesting that the virtual agent that provides rationalized feedback were appropriate in interview practice situations.

Future study should conduct a long-term evaluation experiment with more participants since interview practice was not a one-time event. It is important to verify whether users keep their feeling of reliability toward the virtual agent that gives rationalized feedback, while maintaining preferable impression on the virtual agent enough to keep using the job interview training system during the course of multiple uses.

6 Conclusion

The purpose of the study is to investigate the effectiveness of rationalized feedback from the agent in terms of user’s impression on the agent, i.e., friendliness, performance, trust.

The proposed system uses a Tobii eye tracker for gaze recognition and camera images for facial expression and posture recognition. The system compared the recognition results of the interviewee's nonverbal behaviors with exemplary models. A virtual agent acted as an advisor gives feedback on the interviewee’s behaviors that need improvement. To test the effectiveness of the two kinds of feedback, rationalized and non-rationalized, we conducted an experiment to compare the impressions of the interviewees.

The results of the evaluation experiments showed that the main factor in forming interpersonal impressions of the agent in this experiment was reliability. This suggests the virtual agent that gives interview feedback gave an impression related to reliability followed by unfriendliness. In addition, virtual agent was perceived as a proactive advisor that gives articulate feedback, sometimes too articulate to accept by the participants.

Finally, the agent with rationalized feedback was rated as more trustworthy, reliable, and likely to lead to good practice but less friendly than the non-rationalized virtual agent.

ACKNOWLEDGMENTS

A part of this work was supported by Grant-in-Aid for Scientific Research "KAKENHI (C) 20K11926".

REFERENCES

[1] Yuko Matsuda, Minoru Nagasaku, Kunijiro Arai : Infuence of job-hunting anxiety on job-hunting: From the viewpoint of coping, The Japanese Journal of Psychology, Vol. 80, No. 6, pp. 512-519, 2010
[2] Mehrabian A,Silent messages: Implicit communication of emotions and attitudes,Wadworth Publishing.Co.,California, 1981
[3] Washburn P.V., Hakel M.D. ,Visual cues and verbal content as influences on impressions after simulated employment interviews.Journal of AppliedPsychology, pp.58,137-140, 1973
[4] R. D. Arvey ,J. E. Campion, "The employment interview: A summary and review of recent research," Personnel Psychology, vol. 35, no. 2, pp. 281–322, 1982
[5] Shogo Okada, Yoshihiro Matsugi, Yukiko Nakano, Yuki Hayashi, Hung-Hsuan Huang, Yutaka Takase, Katsumi Nitta, Estimating Communication Skills based on Multimodal Information in Group Discussions, Journal of the Japanese Society for Artificial Intelligence,Vol.31,No.6, A130-E, 2016
[6] MIDAS Information Technology Co., Ltd., https://www.inair.co.jp/
[7] ZENKIGEN Co., Ltd., https://harutaka.jp/
[8] Naim, I., Tanveer, M. I., Gildea, D., Hoque, M. E.: Automated prediction and analysis of job interview performance: The role of what you say and how you say it, IEEE FG, 2015
[9] Rao S. B, P., Rasipuram, S., Das, R., Jayagopi, D. B.: Automatic assessment of communication skill innon-conventional interview settings: A comparative study, ICMI, pp. 221–229, 2017
[10] Nanaho Goda, Keitaro Ishihara, Tomoko Kojiri, Job Interview Support System Based on Analysis of Nonverbal Behavior, IEICE Technical Report,Vol.116,No.517,ET2016-98,25-30, 2017
[11] T. Barur, T. Ionut, G. Patrick, P. Kaska, A. Elisabeth.: A Job Interview Simulation: Social Cue-Based Interaction with A Virtual Character, IEEE International Conference on Social Computing (SocialCom2013), pp.220–227, 2013
[12] J. Matthew, B. Laura, F. Micael, J. Neil, A. Michael, J.Emily, W. Katherine, O. Dale, Morris, D, B.: Virtual Reality Job Interview Training for Veterans with Posttraumatic Stress Disorder, Journal of Vocational Rehabilitation 42, pp. 271–279, 2015
[13] H. Tanaka, S. Sakti, Graham. N, T. Toda, H. Negoro, H. Iwasawa, S. Nakamura: Automated Social Skills Trainer, IUI '15 Proceedings of the 20th International Conference on Intelligent User Interfaces, pp. 17–27, 2015
[14] Anderson, K., Andre, E., Baur, T., Bernardini, S., ´ Chollet, M., Chryssafidou, E., Damian, I., Ennis, C., Egges, A., Gebhard, P., Jones, H., Ochs, M., Pelachaud, C., Porayska-Pomsta, K., Rizzo, P., Sabouret, N.: The TARDIS framework: Intelligent virtual agents for social coaching in job interviews, ACE, pp. 476–491, 2013
[15] Damian, I., Baur, T., Lugrin, B., Gebhard, P., Mehlmann, G., Andre, E.: Games are better than books: In-situ ´ comparison of an interactive job interview game with conventional training, AIED, pp. 84–94, 2015
[16] Hoque, M. E., Courgeon, M., Martin, J.-C., Mutlu, B., Picard, R. W.: MACH: My automated conversation coach, UBICOMP, pp. 697–706 ,2013
[17] Langer, M., Konig, C. J., Gebhard, P., Andr ¨ e, E.: ´ Dear computer, teach me manners: Testing virtual employment interview training, International Journal of Selection and Assessment, Vol. 24, No. 4, pp. 312–323, 2016
[18] Baur, T., Damian, I., Gebhard, P., Porayska-Pomsta, K., Andre, E.: A job interview simulation: Social cue-based interaction ´ with a virtual character, SocialCom, pp. 220–227, 2013
[19] Callejas, Z., Ravenet, B., Ochs, M., Pelachaud, C.: A computational model of social attitudes for a virtual recruiter, AAMAS, pp. 93–100 ,2014
[20] Gebhard, P., Baur, T., Damian, I., Mehlmann, G., Wagner, J., Andre, E.: Exploring interaction strategies for virtual ´ characters to induce stress in simulated job interviews, AAMAS, pp. 661–668, 2014
[21] Lucas,G.M., Gratch,J., King,A., Morency,L.P. It's only a computer:Virtual humans increase willingness to disclose. Computers in Human Behavior, 37,pp.94-100, 2014.
[22] Lucas,G.M., Rizzo,A., Gratch,J., Scherer,S., Stratou,G., Boberg,J., Morency,L.P. Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. Frontiers in Robotics and AI, 4, 51. 2017.
[23] Tomoko Mashimo, Effects of advice expression on recipients' cognition, emotion, and motivation to improve behavior, Osaka University, Graduate School of Human Sciences, Proceedings. 45 P.53-P.66, 2020
[24] Kouki Miyazawa, Utterance Generation in Spoken Interaction Systems Based on the Human Dialogue Strategies, Waseda University, School of Human Sciences, Human Sciences Research 25 (1), 184-185, 2012
[25] MicrosoftAzure,https://azure.microsoft.com/ja-jp/services/cognitive-services/face/
[26] tf-pose-estimation, https://github.com/apulis/tf-pose-estimation
[27] Tobii Technology K.K., https://www.tobiipro.com/ja/
[28] Hayashi Fumitoshi, The Measurement of Individual Differences in Interpersonal Cognitive Structure. In Experimental Social Psychology 22, pp. 1-9, 1982.
[29] Tomoki Miyamoto, Daisuke Katagami, Yuka Shigemitsu,Mayumi Usami, Takahiro Tanaka, Hitoshi Kanamori,Yuki Yoshihara, and Kazuhiro Fujikaze, Impression Evaluation Focusing on the End-of-Sentence Style of Utterance in Driving Support Agent Based on Politeness Theory, Intelligence and Information (Journal of the Japanese Society for Intelligent Information and Fuzzy Studies) Vol. 31, No. 3, pp. 739 – 744, 2019