Analysis of Mood Changes and Facial Expressions during Cognitive Behavior Therapy through a Virtual Agent

Kazuhiro Shidara
shidara.kazuhiro.sc5@is.naist.jp
Nara Institute of Science and Technology, Japan

Hiroki Tanaka
hiroki-tan@is.naist.jp
Nara Institute of Science and Technology, Japan

Hiroyoshi Adachi
hadachi@psy.med.osaka-u.ac.jp
Health and Counseling Center, Osaka University, Japan

Daisuke Kanayama
kanayama@psy.med.osaka-u.ac.jp
Health and Counseling Center, Osaka University, Japan

Yukako Sakagami
sakagamiyukako@wellness.hss.osaka-u.ac.jp
Health and Counseling Center, Osaka University, Japan

Takashi Kudo
kudo@psy.med.osaka-u.ac.jp
Health and Counseling Center, Osaka University, Japan

Satoshi Nakamura
s-nakamura@is.naist.jp
Nara Institute of Science and Technology, Japan

ABSTRACT

In cognitive behavior therapy (CBT) with a virtual agent, facial expression processing is expected to be useful for dialogue response selection empathic dialogue. Unfortunately, its use in current works remains limited. One reason for this situation is the lack of research on the relationship between mood changes facial expressions through CBT-oriented interaction. This study confirms the improvement of negative moods through interaction with a virtual agent and identifying facial expressions that correlate with mood changes. Based on the cognitive restructuring of CBT, we created a fixed dialogue scenario and implemented it in a virtual agent. We recorded facial expressions during dialogues with 23 undergraduate and graduate students, calculated 17 types of action units (AUs), which are the units of facial movements, and performed a correlation analysis using the change rate of mood scores and the amount of the changes in the AUs. The mean mood improvement rate was 35%, and the mood improvements showed correlations with AU5 (r = -0.51), AU17 (r = 0.45), AU25 (r = -0.43), and AU45 (r = 0.45). These results imply that mood changes are reflected in facial expressions. The AUs identified in this study have the potential to be used for agent-interaction modeling.

CCS CONCEPTS

• **Applied computing → Health care information systems**.

KEYWORDS

cognitive behavior therapy; facial expression; virtual agents

ACM Reference Format:
Kazuhiro Shidara, Hiroki Tanaka, Hiroyoshi Adachi, Daisuke Kanayama, Yukako Sakagami, Takashi Kudo, and Satoshi Nakamura. 2020. Analysis of Mood Changes and Facial Expressions during Cognitive Behavior Therapy through a Virtual Agent. In *Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3395035.3425223

1 INTRODUCTION

Cognitive behavior therapy (CBT) examines biased negative thoughts and seeks solutions to current problems [3]. Conversational agents that adapt CBT methodology have been proposed, some of which have been effective [14]. Several types of agent dialogues have been proposed, such as text-based dialogues in the style of messaging apps [6, 7, 9, 17] and virtual agents that are animated from computer characters [11, 16]. Virtual agents have an advantage because they can provide a face-to-face multimodal interaction, which resembles CBT with human therapists. During actual psychotherapy, therapists understand their patient's conditions not only through words but also through such non-verbal communication as facial expressions, prosody, gestures, etc. [12]. References to patient's facial expressions have also been made in the literature on CBT in clinical practice [3]. Perhaps virtual agents can even cooperatively interact with users as humans do with each other [4]. Therefore, interacting with virtual agents has the potential to allow for health care interactions using non-verbal communication like facial expressions and voice, similar to human interaction. On the other hand, according to a 2019 survey on the use of conversational agents in health [14], few studies that consider non-verbal behaviors by referring to medical data-set. For example, agents have been proposed to provide social skills training based on the analysis of non-verbal behaviors for the autism spectrum disorders [18, 19]. Such cases

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-8002-7/20/10…$15.00
https://doi.org/10.1145/3395035.3425223

Table 1: Scenario of utterances spoken by the system (We translated the original Japanese into English).

Turn No.	Item	Sentence
1	Situation	Hello, my name is Mai, and I am a therapist. Come join me in training to face your worries. Is there anything that's been bothering you lately that's been difficult to deal with or to face? Please tell me something that you think is painful or burdensome.
2	Mood	How did your mood change at that time?
3	Mood score (0 to 100%)	How would you describe the intensity of your mood from 0 to 100?
4	Automatic thought	I want to know why you felt like that in that situation. What thoughts came to your mind when you faced that event?
5	Proof	I see... Such thoughts that surface when faced with certain events are called automatic thoughts. Since automatic thoughts are unconscious, they sometimes seem true. If your automatic thought is correct, what do you think it is based on?
6	Disproof	So, do you have another view about that situation? For example, does anything contradict your automatic thinking?
7	Disproof (confirmation)	Do you have any other thoughts? If so, please say them.
8	Socratic questioning* 1	Let's examine the basis of the automatic thought that we have considered so far as well as a way of thinking that is balanced from contrary facts. Let me ask you some questions. What is the worst possible consequence for this situation?
9	Socratic questioning 2	Looking at the other side, what is the end result when you act as you want?
10	Socratic questioning 3	Please describe the most realistic scenario from the predictions you just made.
11	Balanced thought	From that answer, let's create a new thought. How can you rethink this event?
12	Caring words 1	Good. Do you have anyone to hear from you about this problem?
13	Caring words 2	I see, can you manage that?
14	Mood after change (0 to 100%)	Well, sometimes it's easier to talk to someone. Please feel free to talk to me. How has the intensity of your original mood changed? Express it on a scale from 0 to 100.
[End]	Concluding a session	If it's different from the beginning, that suggests that you organized my thoughts well. That's it for today. Thank you for your hard work. Please feel free to call again.

* Questions that create awareness of inconsistency in thinking

Table 2: Action units (AUs) used for analysis

AU	Description	AU	Description
AU1	Inner brow raiser	AU14	Dimpler
AU2	Outer brow raiser	AU15	Lip corner depressor
AU4	Brow lowerer	AU17	Chin raiser
AU5	Upper lid raiser	AU20	Lip stretcher
AU6	Cheek raiser	AU23	Lip tighrener
AU7	Lid tightener	AU25	Lips part
AU9	Nose wrinkler	AU26	Jaw drop
AU10	Upper lip raiser	AU45	Eyes closed (Blink)
AU12	Lip corner puller		

are still rare, to the best of our knowledge, no agents have been created based on the analysis of the non-verbal features of CBT. In this study, we performed a one-day CBT using a virtual agent and analyzed the relationship between mood improvement and changes in facial expressions. In addition, we created a dialogue scenario based on the cognitive restructuring method [3] in CBT. This method often use a worksheet called the thought record, which is a seven-item question flow to restructure cognition. It has been widely used in both CBT with therapists and for personal use. The thought record was originally proposed by Beck et al [2]. In this study, an improved version by Greenburger et al. [8] was used. In previous studies, dialogue scenarios of conversational agents were either unavailable or only partially available to the public. One of our purposes is to provide a baseline for subsequent research.

Our paper makes the following three contributions: (1) presents a dialogue scenario that forms the basis of CBT-oriented dialogue agent research; (2) confirms that the agent effectively reduced negative moods in one dialogue with this scenario; (3) investigates facial movements displayed by the user and the areas of the face where the amount of the change is correlated with the improvement of the mood.

2 VIRTUAL AGENT DESIGN

MMDAgent [13] was used as the virtual agent in this paper. We used default parameters for its speech, such as speaking rate and voice pitch. This is a neutral facial expression setting, and it is assumed that the user's facial actions are not affected by the virtual agent's expression. This agent outputs spoken language, and the user inputs spoken natural language through a headset microphone. Table 1 shows the system side scenario, which was created under the supervision of a psychiatrist. In this experiment, we implemented a fixed dialogue scenario for the following reasons. In clinical practice, the treatment flow is structured to properly perform training, especially the question items and their order, which are determined in the cognitive restructuring method. A certain level of effectiveness exists, even if the agent's question in the dialogue is completely set. There are other benefits as well. The fixed dialogue scenario enables users to quantify moods before and after conversations and quantitatively evaluates them [15] and is widely used to people regardless of the presence or absence of symptoms such as depression. In addition to the seven-item thought record table and used in the cognitive restructuring method, the routine dialogue in our experiment included self-introduction, caring words, and Socratic questioning [3]. Socratic questioning enables users to become aware of and evaluate their own thoughts.

The purpose of the virtual agent is to acquire facial expressions in spoken dialogue, and the implemented scenario was aimed at effective CBT. Our hypothesis is that this scenario is effective in improving negative mood regardless of dialogue style. To examine this, we implemented the same scenario to a text-based agent with text input and output, and compared the effectiveness of CBT in the two dialogue styles. To implement a text-based agent, we used the text chat-bot creation feature of Slack, the messaging application [1].

3 DATA COLLECTION

3.1 Participants

We recruited 36 undergraduate and graduate students as participants (16 females), and 23 of them categorized as the virtual agent group and 13 as the text-based agent group. The research ethics committees of the Nara Institute of Science and Technology reviewed and approved this experiment. Written informed consent was obtained from all participants before the experiment. We also confirmed that participants had no severe depressive tendencies based on the Kessler Psychological Distress Scale (K6) [10]. The K6 scores ranged from 0 to 24, and the cutoff was 13+. The K6 score for the virtual agent group was (Mean = 7.18, SD = 4.22), the text-based agent group was (Mean = 6.23, SD = 4.40).

3.2 Experiment Procedure

The same procedure was applied for the two groups. To begin with, the participants read a leaflet explaining CBT [2]. This leaflet is open to the public and is designed for both clinical and general public uses. Then the first author explained how to use the system. The data were collected using a laptop PC (HP Probook) in a quiet room only occupied by the participant. We recorded virtual agent group's facial expressions from the built-in camera on the laptop. The completion time fluctuated based on the amount of the participant's speaking/typing.

3.3 Mood Score Rating

Participants were asked twice about their negative mood intensities as their "mood score". The first is the question on the 3rd turn, and the second is the question on the 14th turn. In this study, we analyzed the changes in their moods and the corresponding changes in facial expressions caused through the two ratings. We

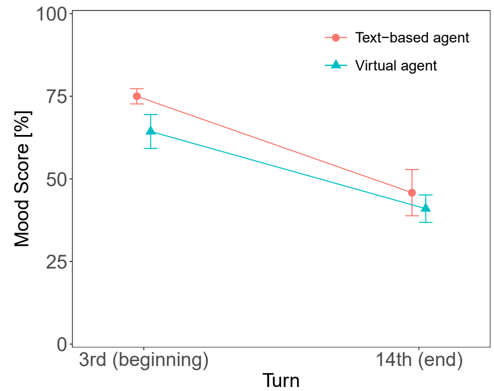

Figure 1: Mean and standard error of mood at the beginning score and mood score at the end of interactions with a virtual agent and text-based agent.

expected to describe the participants' moods with such labels as anxious, depressed, sad, inferior, fatigued, and so on. In this paper, we uniformly lumped all such feelings under the rubric of negative moods and only evaluated them by mood scores to focus on the changes themselves. Based on a previous work [15], we calculated the mood changes using the following formula:

$$Mood\ change = \frac{(Mood\ score\ at\ beginning) - (Mood\ score\ at\ end)}{(Mood\ score\ at\ beginning)} \tag{1}$$

3.4 Facial Data Extraction

Based on the Facial Action Coding System (FACS) [5], which is a quantitative method for describing facial movements, our analysis used action units (AUs), which are partial units of facial expressions. AUs were automatically extracted with OpenFace [1], a facial expression analysis tool. The 17 analyzed AUs are listed in Table 2. All the extracted AUs were represented as intensity information of continuous values from 0 to 5. In this study, the AUs data were pre-processed for analysis as follows:

a). Extracted the AUs from the videos.
b). OpenFace calculates the reliability of successful face recognition. Removed frames with the reliability of less than 70%.
c). Separated the agent and user turns and ignored the former.
d). Calculated the mean of the AUs in the frames: The mean AU of the frames was calculated for each extracted user turn.
e). Standardized the 14 mean values within individuals.
f). Calculated the facial expression differences through interaction by the following equation:

$$Change\ of\ AU = (AU\ at\ beginning) - (AU\ at\ end) \tag{2}$$

We analyzed the correlation between the mood changes and the changes of AUs.

[1] https://slack.com/
[2] https://www.cbtjp.net/downloads/skillup/pdf/

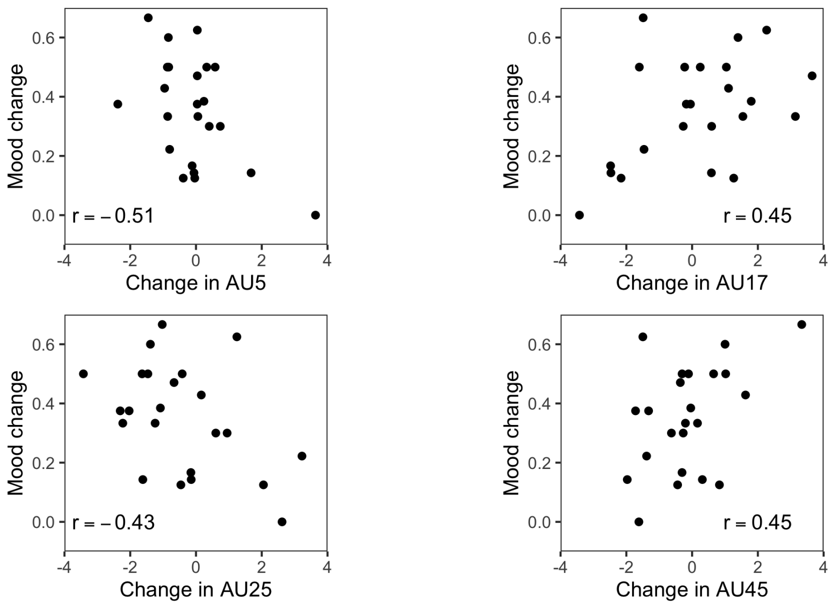

Figure 2: Scatter diagram between Mood Changes vs. AU5 (top left), A17 (top right), AU25 (bottom left) and AU45 (bottom right).

4 RESULTS

4.1 Mood Changes through the Interaction

Figure 1 plots the means and the standard errors of the mood scores. The reported mood scores changed significantly between the 3rd turn and the 14th turn with the paired t-test in both groups (virtual agent, Mean= 0.35, SD = 0.18, t = 7.72, p < 0.001; text-based agent, Mean = 0.37, SD = 0.29, t = 4.41, p < 0.001). Therefore, there was no interaction effect due to the two dialogue styles.

4.2 Facial Actions that are Correlated with Moods

As a result of Pearson's non-correlation test, the following changes of AUs were significantly correlated with mood changes at α = 0.05: AU5 (r = -0.51, t = -2.68, p = 0.01), AU17 (r = 0.45, t = 2.32, p = 0.03), AU25 (r = -0.43, t = -2.24, p = 0.04), and AU45 (r = 0.45, t = 2.32, p = 0.03). Figure 2 shows scatter diagrams between these changes in AUs and moods. AU5 and AU45 are movements in the upper half of the face, showing squinting; AU17 and AU25 are the movements in the lower half of the face, showing that the chin is raised.

5 DISCUSSION

These results imply that CBT with a conversational agent can change the mood of participants, and such changes appear as facial expressions in proportion to mood improvement. It was difficult to interpret the reason for these movements, therefore we reviewed the recorded videos. Then, those who had relatively large mood changes appeared to think more deeply when answering the mood score at the 14th turn than at the 3rd turn. Therefore, we assumed that the deeper thinking attitude was reflected in the movements of eyelids raising and mouth closing. In this study, individual AUs were analyzed. However, there is insufficient information to conclude that these AUs are due to mood changes. In order to recognize the mood changes more reliably, it is considered effective to analyze not only facial expressions but also multimodal behavioral indicators such as voice and gestures.

6 CONCLUSIONS

We created a fixed dialogue scenario and used it for all our participants. Our scenario, which significantly improved their negative moods, is suitable as a baseline for CBT research with a conversational agent. This is also the first study to analyze how facial movements, which are related to mood changes, are affected by interaction with the agent. As a result, we identified candidates for facial expressions that change with mood changes. This result is expected to make an important contribution to the research of dialogue systems that recognize facial movements as well as to the basic research of human facial expression analysis. However, it has not been fully verified that the correlation between the mood changes and changes of AUs is due to factors other than mood, except for the factors of dialogue style. Future work must investigate this point. In this study, we analyzed facial expressions on the premise that virtual agents make users express emotions more than text-based agents. However, it is not yet clear whether virtual agents are more suited to promoting expressing emotions. Analyzing facial expressions between different dialogue styles is another subsequent step in our future work.

ACKNOWLEDGMENTS

This work was supported by JST CREST Grant Number JPMJCR19A5, Japan.

REFERENCES

[1] Tadas Baltrušaitis, Peter Robinson, and Louis-Philippe Morency. 2016. Openface: an open source facial behavior analysis toolkit. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 1–10.

[2] AT Beck, AJ Rush, BF Shaw, and G Emery. 1979. Cognitive therapy of depression Guilford Press. *New York* (1979).

[3] Judith S Beck and AT Beck. 2011. Cognitive behavior therapy. *New York: Basics and beyond. Guilford Publication* (2011).

[4] David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, et al. 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 1061–1068.

[5] Paul Ekman and Wallace V Friesen. 1978. *Facial action coding systems*. Consulting Psychologists Press.

[6] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e19.

[7] Russell Fulmer, Angela Joerin, Breanna Gentile, Lysanne Lakerink, and Michiel Rauws. 2018. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health* 5, 4 (2018), e64.

[8] Dennis Greenberger and Christine A Padesky. 1995. *Mind over Mood: a cognitive therapy treatment manual for clients*.

[9] Becky Inkster, Shubhankar Sarda, and Vinod Subramanian. 2018. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth* 6, 11 (2018), e12106.

[10] Ronald C Kessler, Gavin Andrews, Lisa J Colpe, Eva Hiripi, Daniel K Mroczek, S-LT Normand, Ellen E Walters, and Alan M Zaslavsky. 2002. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological medicine* 32, 6 (2002), 959–976.

[11] Everlyne Kimani, Timothy Bickmore, Ha Trinh, and Paola Pedrelli. 2019. You'll be Great: Virtual Agent-based Cognitive Restructuring to Reduce Public Speaking Anxiety. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 641–647.

[12] Sander L Koole and Wolfgang Tschacher. 2016. Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance. *Frontiers in psychology* 7 (2016), 862.

[13] Akinobu Lee, Keiichiro Oura, and Keiichi Tokuda. 2013. MMDAgent—A fully open-source toolkit for voice interaction systems. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 8382–8385.

[14] Joao Luis Zeni Montenegro, Cristiano André da Costa, and Rodrigo da Rosa Righi. 2019. Survey of conversational agents in health. *Expert Systems with Applications* 129 (2019), 56–67.

[15] Jacqueline B Persons and David D Burns. 1985. Mechanisms of action of cognitive therapy: The relative contributions of technical and interpersonal interventions. *Cognitive therapy and Research* 9, 5 (1985), 539–551.

[16] Lazlo Ring, Timothy Bickmore, and Paola Pedrelli. 2016. An affectively aware virtual therapist for depression counseling. In *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) workshop on Computing and Mental Health*. 01951–12.

[17] Shinichiro Suganuma, Daisuke Sakamoto, and Haruhiko Shimoyama. 2018. An embodied conversational agent for unguided internet-based cognitive behavior therapy in preventative mental health: feasibility and acceptability pilot trial. *JMIR mental health* 5, 3 (2018), e10454.

[18] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PloS one* 12, 8 (2017).

[19] Hiroki Tanaka, Sakti Sakriani, Graham Neubig, Tomoki Toda, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2016. Teaching social communication skills through human-agent interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 6, 2 (2016), 1–26.