



Fourteenth International Conference on Intelligent Virtual Agents (IVA 2014)

Proceedings of the Workshop on

Affective Agents

Organising Committee

Lazlo Ring, Northeastern University, USA
Iolanda Leite, Yale University, USA
João Dias, INESC-ID and Instituto Superior Técnico, Portugal

Program Committee

Joost Broekens, TU Delft, NL
Frank Dignum, Utrecht University, NL
Sidney D'Mello, University of Notre Dame, USA
Dirk Heylen, Twente University, NL
Kate Hone, Brunel University, UK
Ian Horswill, Northwestern University, USA
Eva Hudlicka, University of Massachusetts, USA
Arvid Kappas, Jacobs University, Germany
James Lester, North Carolina State University, USA
Margot Lhommet, Northeastern University, USA
Stacy Marsella, Northeastern University, USA
Catherine Pelachaud, Paristech Telecom, FR
Laurel Riek, University of Notre Dame, USA
Stefan Scherer, USC Institute for Creative Technologies, USA
Daniel Schulman, Veterans Health Administration, USA
Adriana Tapus, ENSTA-ParisTech, France
Janneke van der Zwaan, TU Delft, NL

Preface

Affective computing has been on the rise in the last decade, with numerous advances in how we model, detect and understand a users affective state. Many researchers in the virtual agents and social robotics communities are exploring the integration of affective computing into their systems. However, the majority of this research focuses on only one aspect of affective computing, such as the detection, modeling, or expression of emotion. The goal of this workshop is to create collaborations between the experts in each of these areas to share their latest findings so as to shape the picture of how all of these systems can be integrated together to create an Affective Agent that senses, models, expresses and responds to a user's affective state in real time.

26st, August, 2014

Contents

- 1 **M. Lim, M. E. Foster, S. Janarthanam, A. Deshmukh, H. Hastie & R. Aylett, *Let's Go for a Treasure Hunt*** 1
- 2 **D. Cernea, C. Weber, A. Kerren & A. Ebert, *Group Affective Tone Awareness and Regulation through Virtual Agents*** 9
- 3 **M. Bruijnes, S. Wapperom, R. Akker, & D. Heylen, *A Virtual Suspect Agent's Response Model*** 17
- 4 **J. D. McDaniel & M. Si, *Length of Smile Apex as Indicator of Faked Expression*** 25

Let's Go for a Treasure Hunt

Mei Yii Lim, Mary Ellen Foster, Srinivasan Janarthanam,
Amol Deshmukh, Helen Hastie, and Ruth Aylett

School Of Mathematical and Computer Science
Heriot-Watt University,
EH14 4AS, Edinburgh, Scotland, UK
{m.lim,m.e.foster,s.chandrasekaran_janarthanam}@hw.ac.uk
{a.deshmukh,h.hastie,r.s.aylett}@hw.ac.uk

Abstract. This paper presents a study designed to explore the effect of feedback on perception of an embodied agent as well as the overall performance and experience of primary school children aged 12-13 carrying out a treasure hunt activity. We use an embodied agent to compare three experimental conditions: no feedback, neutral feedback, and affective feedback. What the students think about the embodied agent and how they feel about the task under the different conditions will be elicited through a questionnaire upon completion of the treasure hunt activity. Moreover, how each condition affects the students' performance will be analysed.

1 Introduction

Emotions play an important role in human-human interaction [1]. Agents that exhibit human-like emotions have now become a commonplace in the domain of human-computer interaction. Starting from the pioneering work of [2] and [3], emotional agents now exist in various applications to serve different purposes including but not limited to military [4], health [5], commerce [6], tourism [7], video games [8] and education [9, 10, 11, 12]. In education, emotional expressions have been incorporated into embodied teaching agents with the aim of improving learning experience in users. Although inclusion of emotional expressions into virtual tutors rarely lead to negative interaction, positive effect was not always achieved on learning experience [13]. This might be due to the fact that learning task requires concentration and if an agent offers assistance at inappropriate time, the result is more of a distraction than facilitation.

It is essential that we understand the impact of emotions in embodied agents upon users in order to establish a successful agent-human interaction. In order to investigate the impact of emotional expressions on users' learning experience, it is not sufficient to simply ask whether emotional agents are better or worse than unemotional agents [14]. The more relevant issues are: (1) what kind of emotional expression has an effect on users; (2) what elements of the user's attitude and/or performance are affected; and (3) what is the impact of different forms of emotional expression. In this paper, we present an experiment to investigate

how feedback—none, neutral, or affective—affects a child’s perception, experience and performance in a real-world treasure hunt activity. This work takes place in the context of the EU project EMOTE¹ (EMbOodied-perceptive Tutors for Empathy-based learning) which aims to develop virtual tutors that have the perceptive and expressive capabilities to engage in empathic interactions with learners in school environments, grounded in psychological theories of emotion in social interaction and pedagogical models for learning facilitation.

2 The Treasure Hunt

2.1 The Experiment

The treasure hunt activity requires a child to apply his/her map reading skills and is aimed at primary school children aged 12-13. There will be three experimental conditions: no feedback, neutral feedback and affective feedback. In the no feedback condition, students will be given paper maps and instructions, and will not interact with an embodied agent at all during the treasure hunt. In the other two conditions, students will be given Android tablets running an application which displays a digital version of the paper map, along with an embodied agent which will present the instructions and pose the questions. This agent will also provide the students with feedback on the correctness of their answers to the questions posed during the treasure hunt; depending on the experimental condition, the feedback will be either neutral or affective. The no feedback condition was designed using paper maps and instructions due to non-availability of Android tablets for a section of students who were to take part in this annual school exercise. Although, this condition was performed without tablets, we consider it as a baseline to the two tablet conditions.

In total, 36 students will participate in this study. They will carry out the treasure hunt in pairs, resulting in 6 groups per condition. Prior to the treasure hunt, all students will have a short interactive session with a robot called Susie. The robot will introduce the treasure hunt and conduct a short question and answer session to check the students’ readiness for the activity. The robot will be controlled by a wizard in the neighbouring room, and will therefore be capable of taking a few questions from the students if necessary. The main aim of this session is to allow the students to interact and familiarise themselves with the robot, which will then appear as an embodied virtual agent on the tablet for the feedback conditions. The virtual agent looks and sounds the same to the robot. It uses the same text-to-speech (TTS) engine as the robot.

2.2 Objectives

Through this treasure hunt activity, we would like to explore the effect of feedback on the students’ perception of an embodied agent as well as their overall

¹ <http://www.emote-project.eu/>

experience and performance in carrying out the task at hand. Applying the two-tiered method for evaluating affective interfaces [15], we start by verifying that the students notice the expression or non-expression of emotions and that the perceived emotions are those we intended the agent to portray. Failing to effectively interpret the emotional expressions of the agent will lower the validity of our study as it would be unclear that any effects found are due to the manipulation of emotional expression. In this study we restrict the emotional display to only three basic expressions (neutral, happy and sad) to ensure that the children understand the affective information being communicated.

The feedback includes both emotional facial expressions and utterances. In the affective condition, a happy expression will be displayed accompanied by utterances such as “brilliant, very good, fantastic” when students answer a question correctly, while a sad expression will be displayed accompanied by utterances such as “Oh no, I’m sorry” when they answer incorrectly; in the latter case, the correct answer will also be provided. In the neutral condition, the agent will always display a neutral expression and reply with “correct” or “incorrect” utterances. Figure 1 shows the three expressions used in this study.

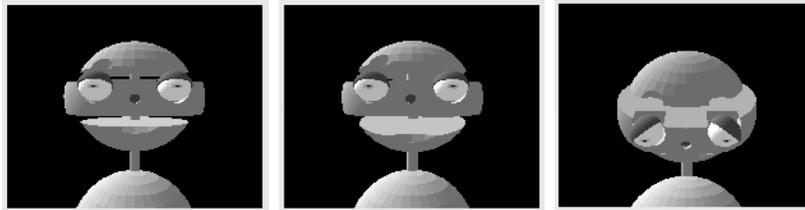


Fig. 1. Neutral, Happy and Sad Expressions

Upon verifying the ability of the students in interpreting the emotional expressions correctly, we would then like to find out how different types of feedback affect the students’ interaction with the agent. In other words, if an agent praises a child when they make good progress, how does this affect the child? Hence, we seek the answers to the following questions:

- Is the affective agent being perceived as more friendly, kind, pleasant and helpful?
- Does affective feedback make them enjoy the interaction more?
- Does affective feedback improve their performance?
- Is the neutral agent more reassuring?
- Which version of the agent is rated more highly by the students as an interaction partner?
- Does the agent—whether neutral or affective—actually help the students in task performance?
- Is there a difference between boys and girls perception of the agent—neutral or affective?

2.3 Treasure Hunt Application

We have designed and implemented a treasure hunt Android application for the above study. In order to compare the three experimental conditions, we have kept the features of the application to be as close to the paper version as possible, except for the addition of the embodied character Susie. All images, fonts and layout are comparable between the two versions. The application (Figure 2) displays a map corresponding to its paper counterpart (Figure 3) and presents a sequence of the same steps as in the paper version to be carried out by the students.



Fig. 2. The Treasure Hunt Application Start Screen

As the screen of the tablet is smaller than a piece of A4-size paper, the map comes with a drag and zoom functionalities to enable the students to explore it as they would with the paper version; note that the map cannot be zoomed to larger than 100% its actual size. Each step starts with the virtual character presenting a task and questions to the user through speech. Subtitles are displayed on screen in case the students missed what Susie was saying, and the students can also replay the speech at any point if necessary. Each task requires the students to walk a few yards making use of their map skills. At the end of each walk, the students have to confirm their arrival.

The system will then re-present relevant questions related to the task with multiple choice answers and the students are required to select an answer from

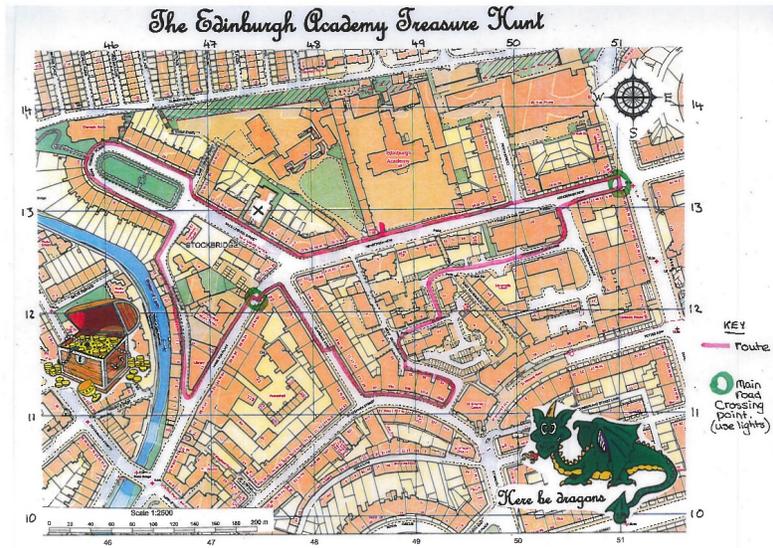


Fig. 3. The Paper Map

the given choices (Figure 4). Depending on whether the answer is correct or not, the system responds with appropriate feedback—neutral or affective. In the paper version, the students are also presented with multiple choice answers of which they have to circle the correct one. In both the paper and the tablet conditions, the students are also given a chance to win extra prizes by answering additional questions at the top of the paper questionnaire or through an ‘Extra Prize’ link on the top right corner of the tablet screen.

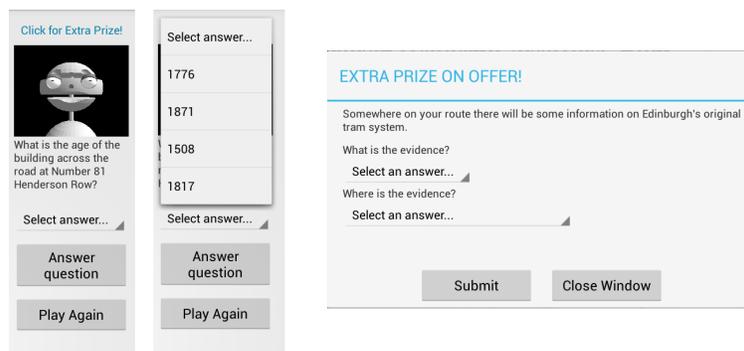


Fig. 4. Question and Answer Interface; The Extra Prize Dialog Box

2.4 Data Collection

Following the treasure hunt, the students will answer a short questionnaire. It focuses specifically on the children’s perception of the embodied agent and their overall experience of the treasure hunt activity, applying the combination of Godspeed likeability items [16] and the Smileyometer, an instrument used to measure enjoyment and fun [17] aiming to make the task of answering the questionnaire more interesting for the target group. The Smileyometer uses pictorial representations of different kinds of happy faces to depict the diverse level of satisfaction according to 5-point Likert scale as shown in Figure 5.

The treasure hunt was ...

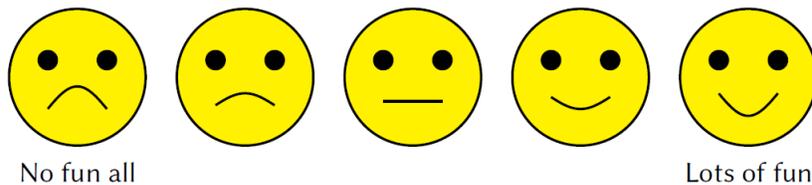


Fig. 5. Example Question with the Smileyometer

During the treasure hunt, as the students complete each task and answer a question in the tablet conditions, the information is logged. The information stored includes a timestamp of task completion, the task ID, the student’s answer as well as the agent’s feedback, enabling the teachers to discuss the students’ performance when they are back in the classroom. Additionally, time stamped GPS data is also collected for all participants including those who carry out the treasure hunt on paper. This is done by running a GPS logging application on a mobile phone attached to the clipboard they are carrying. The time stamped GPS data will allow us to investigate how the use of technology in comparison to paper based version affected the overall experience and timing of solving the map reading task in addition to investigating how the use of affective and non-affective feedback strategies affected the interactions on the tablet.

3 Conclusion and Future Work

The study is scheduled for the third week of June 2014. By the time of this workshop we will have analysed and deduced reasonable answers to our research questions in section 2.2 which hopefully will provide insights to our future design of an empathic tutor.

Acknowledgements

This work was partially supported by the European Commission (EC) and was funded by the EU FP7 ICT-317923 project EMOTE. The authors are solely responsible for the content of this publication. It does not represent the opinion of the EC, and the EC is not responsible for any use that might be made of data appearing therein.

References

- [1] Damasio, A.: *Descartes' Error: Emotion, Reason and the Human Brain*. Grosset/Putnam Press, New York (1994)
- [2] Bates, J.: The role of emotion in believable agents. *Communications of the ACM* **37**(7) (Jul 1994) 122–125
- [3] Picard, R.W.: *Affective Computing*. MIT Press (1997)
- [4] Gratch, J., Marsella, S.: A domain-independent framework for modeling emotion. *Journal of Cognitive Systems Research* **5**(4) (2004) 269–306
- [5] Bickmore, T., Picard, R.: Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer Human Interaction (TOCHI)* **12**(2) (2005) 193–327
- [6] Gong, L.: Is happy better than sad even if they are both non-adaptive? effects of emotional expressions of talking-head interface agents. *International Journal of Human Computer Studies* **65**(3) (2007) 183–191
- [7] Lim, M.Y.: *Emotions, Behaviour and Belief Regulation in An Intelligent Guide with Attitude*. PhD thesis, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, Edinburgh (2007)
- [8] Isbister, K.: *Better Game Characters by Design: A Psychological Approach*. Morgan Kaufmann (2006)
- [9] Okonkwo, C., Vassileva, J.: Affective pedagogical agents and user persuasion. In Stephanidis, C., ed.: *Proceedings of the 4th International Conference on Universal Access in Human Computer Interaction, Beijing, China, Springer* (2001) 5–10
- [10] Prendinger, H., Mayer, S., Mori, J., Ishizuka, M.: Persona effect revisited. using bio-signals to measure and reflect the impact of character-based interfaces. In Rist, T., Aylett, R., Ballin, D., Rickel, J., eds.: *Fourth International Working Conference On Intelligent Virtual Agents (IVA 03)*, Kloster Irsee, Germany, Springer (2003) 283–291
- [11] Dias, J., Paiva, A.: Feeling and reasoning: A computational model for emotional agents. In: *12th Portuguese Conference on Artificial Intelligence (EPIA 2005)*, Portugal, Springer (2005) 127–140
- [12] Maldonado, H., Lee, J., Brave, S., Nass, C., Nakajima, H., Yamada, R., Iwamura, K., Morishima, Y.: We learn better together: enhancing elearning with emotional characters. In Koschmann, T., Suthers, D., Chan, T., eds.: *Computer Supported Collaborative Learning 2005: The Next 10 Years!* Lawrence Erlbaum Associates, Mahwah, NJ (2005) 408–417
- [13] Beale, R., Creed, C.: Affective interaction: How emotional agents affect users. *Human-Computer Studies* **67** (2009) 755–776
- [14] Dehn, D., Van Mulken, S.: The impact of animated interface agents: a review of empirical research. *International Journal of Human Computer Studies* **52**(1) (2000) 1–22

- [15] Höök, K.: User-centred design and evaluation of affective interfaces. From Brows to Trust: Evaluating Embodied Conversational Agents **7** (2004) 127–160
- [16] Bartneck, C., Croft, E., Kulić, D.: Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* **1**(1) (2009) 71–81
- [17] Read, J., Macfarlane, S.: Endurability, engagement and expectations: Measuring children’s fun. In: *Interaction Design and Children*, Shaker Publishing, Shaker Publishing (2002) 1–23

Group Affective Tone Awareness and Regulation through Virtual Agents

Daniel Cernea^{1,2}, Christopher Weber³, Andreas Kerren², and Achim Ebert¹

¹ University of Kaiserslautern, Computer Graphics and HCI Group
P.O. Box 3049, D-67653 Kaiserslautern, Germany

² Linnaeus University, Computer Science Department, ISOVIS Group
Vejudes Plats 7, SE-35195 Växjö, Sweden

³ UC Davis, Department of Computer Science, CA 95616, United States
{cernea,ebert}@cs.uni-kl.de, chrweber@ucdavis.edu, andreas.kerren@lnu.se

Abstract. It happens increasingly often that experts need to collaborate in order to exchange ideas, views and opinions on their path towards understanding. However, every collaboration process is inherently fragile and involves a large set of human subjective aspects, including social interaction, personality, and emotions. In this paper we present Pogat, an affective virtual agent designed to support the collaboration process around displays by increasing user awareness of the group affective tone. A positive group affective tone, where all the participants of a group experience emotions of a positive valence, has been linked to fostering creativity in groups and supporting the entire collaboration process. At the same time, a negative or inexistent group affective tone can suggest negative emotions in some of the group members, emotions that can lead to an inefficient or even obstructed collaboration. A study of our approach suggests that Pogat can increase the awareness of the overall affective state of the group as well as positively affect the efficiency of groups in collaborative scenarios.

Keywords: affective virtual agents, group affective tone, awareness

1 Introduction

It is not seldom that gaining insight into large, compound datasets requires the knowledge and experience of a diverse group of users. But while certain advantages can be achieved by harnessing the expertise of multiple users through collaboration, one needs to also consider the subjective human aspects that influence communication and cooperation. As such, a group of experts can only maintain their collaboration as long as subjective human aspects, like personality, emotions and social interactions, do not affect it negatively. One measure employed to express the subjective coherence of a group is the *group affective tone* (GAT) [10]. It is defined as the presence of homogeneous emotional states throughout the entire group, i.e., all group members present an affective state of similar valence orientation and value (e.g., all group members have very positive emotions).

The group affective tone of a team can have multiple values. In configurations where the members of a group have diverse emotional states and valence, a GAT cannot be defined. However, in cases where all members have either a positive or a negative emotional state, the corresponding positive GAT (PGAT) or a negative GAT (NGAT) can be defined. More importantly, PGAT has been linked on multiple occasions to increased effectiveness and creativity levels in group settings [6, 14].

As our current research focuses on increasing GAT awareness in groups, we explored the literature for visual representations of affect. Sadly, there is only limited work on representing GAT or supporting collaboration with affective virtual agents. Potential representations for emotional states involve abstract visualizations [17, 5], interface widgets [16], as well as a range of affective icons and agents [9, 13, 16]. More importantly for our approach are the affective virtual agents used in collaboration-relevant contexts like frustration management [12], emotional self-awareness [2], empathy [7], and negotiations [1, 18]. Even more focused on collaboration, in [8] the Virtual Messenger system is highlighted which employs affective 3D animated avatars in order to capture and express the emotions of users in a communication setting.

In the following section, we highlight the specifics of Pogat, an affective virtual agent for increasing GAT awareness and aiding the development of PGAT in collaborative scenarios. Next, we present a user study capturing the effects Pogat has on groups that try to solve a task collaboratively on a tabletop display. Finally, we discuss advantages and potential pitfalls of our approach and offer our conclusions.

2 Awareness and Collaboration

In order to determine the GAT for a group, we need to have a method for interpreting the emotional states of each group member in real-time. While this is achievable through a set of technologies, for our context where users would interact and collaborate around a large display or tabletop, we focused on employing lightweight and portable electroencephalographic (EEG) devices for obtaining the current affective states of each participant (see Figure 1). Each member of a collaborative session is equipped with an EEG headset that is capable of capturing electrical signals from the user’s scalp. These brain signals are interpreted as emotional states with the help of the Emotiv software framework and decomposed into values of emotional valence through the use of Russell’s circumplex model of affect [19]. Note that the focus of this paper is not on the acquisition of the user emotional states with the Emotiv EPOC device. Our approach in this domain is similar to the ones presented and detailed in [3, 4], where an detection accuracy of up to 80% has been obtained when compared to user self-reports.

Once the emotional readings are obtained for each participant, the data is sent to the Pogat virtual agent system. The decision for employing an affective virtual agent for expressing GAT was influenced by an agent’s natural ability to mirror user emotions, as well as the ability of an individual to better perceive



Fig. 1. Image of the Emotiv EEG neuroheadset (left). User wearing the EEG device while interacting with a computer (right).

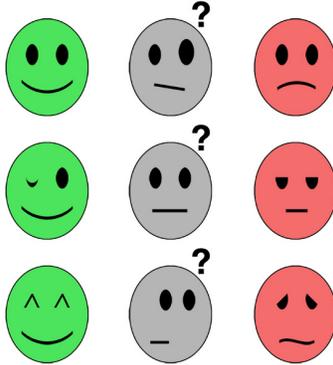


Fig. 2. Representations of the virtual agent encoding the valence of the group affective tone: group affective readings with a positive valence are expressed through corresponding facial expressions and green coloring of the agent (left); heterogeneous group affective readings in terms of valence are represented through facial expressions that suggest confusion and a desaturated color (center); a negative group affective tone is represented through corresponding facial expressions and a red coloring of the agent (right). The facial expressions of the agent are inspired by a subset of *emoji* faces.

emotional states through facial expressions than abstract representations. The Pogat system is comprised of two modules: one for analyzing the individual emotional states of the group members and extracting the GAT, and one for managing the representation and feedback offered by the affective virtual agent. For the computation of the GAT, the system checks the emotional valence of each participant by accessing the affective information derived by the EEG device and mapped in the normalized domain of valence $[-1, 1]$. This results in N readings V_N , where N is the number of group members and V_i is the current valence

of member i in the group, $-1 \leq V_N \leq 1$. If all valence readings have similar values, i.e. $V_i - V_j \leq k$ for any i, j in $[0, N]$ and k is a threshold with the default value of 0.25, then a GAT exists. Homogeneous readings in the positive space are interpreted as the presence of a positive group affective tone (PGAT), and homogeneous negative emotional states as negative group affective tone (NGAT).

These findings are represented by the second module by an emoticon-like agent that exploits the visual channel to offer feedback by modulating its facial expressions based on the current GAT of the group: expressions suggesting happiness correspond to the presence of a PGAT, expressions related to sadness correspond to a NGAT. And finally, neutral expression or expressions of wonderment correspond to a current lack of GAT (i.e., the group members have very different emotional states in terms of valence). To further enforce user awareness, Pogat modulates the color of the virtual agent by mapping it to the three potential GAT states (see Figure 2). The transitions between the represented GAT states are gradual, allowing for three seconds in which the agent slowly changes his expression and color.

3 User Study

We executed a user study in order to inspect the effects that our affective virtual agent would have on the GAT awareness and subsequently how this would affect group performance and interaction. Our study involved 12 participants divided in groups of three, with an average age of 22.5 and an equal male-female distribution. The experiment focused on finding particular subgraph structures in a time-series visualization running on a tabletop (see Figure 3). More precisely, the groups were asked to find specific patterns of user browsing behavior in a tabletop visualization of large datasets storing the browsing history of multiple users over multiple days.

After a brief introduction of the tabletop system and the Pogat agent, each group was engaged in six search tasks resulting in 24 executed tasks in total. The tasks of each group were divided into three categories: without the help of the Pogat affective agent, with the presence of the Pogat affective agent on the tabletop and with the presence of the Pogat agent that was also offering text-based suggestions for achieving a PGAT. For the third category, the agent was programmed to not only react and present the current GAT state, but also to offer a suggestion on how to achieve a positive group affective tone in cases when NGAT or the lack of emotional consensus (i.e., some group members presented a deviating and negative emotional valence) was present. The text-based messages of the agent were selected randomly from a predefined dataset and focused on tasks and actions shown to lead to a PGAT in certain conditions: taking a break or executing a short fun task, reevaluating previous decisions, or increasing group interaction and communication [15]. Note that the order of the tasks and appearance of the agent were randomized for every group, and that the timer measuring the efficiency during a particular task was stopped while the users were executing one of the agent-proposed suggestions.



Fig. 3. Users collaborating around a tabletop application while wearing BCI devices that interpret their affective states. The GAT virtual agent inspects these affective states and offers real-time feedback about the group affective tone, as well as suggestions for developing a PGAT.

The results of our study are highlighted in Figure 4. They suggest that employing a virtual agent for offering feedback about the GAT can significantly reduce the average task completion time by up to 26%. Further, the average times spent by each group without a GAT or in a NGAT present a drop in the two cases where the Pogat agent was used. Additionally, although the completion times are similar in the last two cases, it seems that employing an agent capable of offering concrete suggestions towards regulating the GAT improves the amount of time spent by a group in a no-GAT state, and increases the period of experienced PGAT.

Furthermore, an increased level of communication both about the task at hand and the group affective tone was noticeable in our study. Participants seemed to be more interested in addressing issues revolving around potential controversial group decisions (e.g., "I think we can combine these two filters for finding the longer active time. Does everyone agree with this?"). At the same time, post-task inquiries have shown that group members were mostly positive towards the use of an affective virtual agent for highlighting GAT, with 10 participants considering such a system as advantageous in supporting collaboration around large displays.

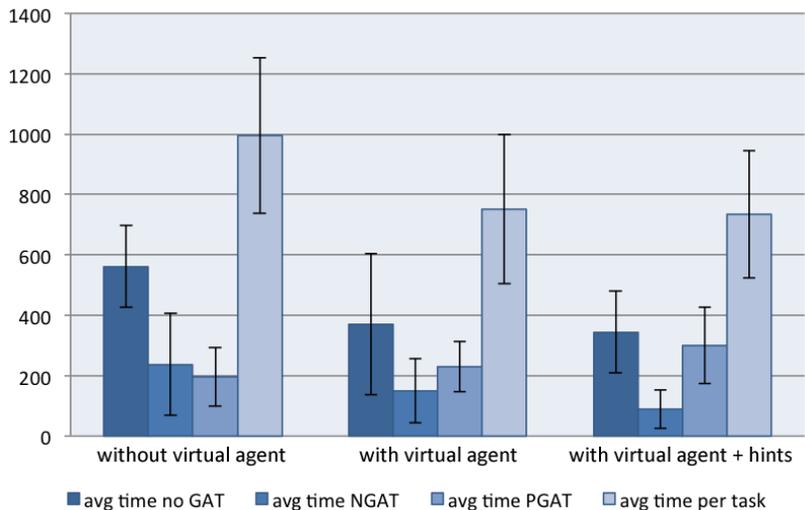


Fig. 4. Average times (in seconds) and standard deviation values for the three categories of tasks the groups had to solve: without the aid of the virtual agent (left), with the support of the virtual agent (center) and with the virtual agent offering also text-based hints (right). The four bars in each category encode (left to right) average time that the groups spent with no-GAT, negative GAT and positive GAT, as well as the average time that was necessary to complete the task. This average time is identical to the sum of no-GAT, NGAT, and PGAT.

4 Discussion

Maintaining good communication and frictionless collaboration is vital in a wide range of domains and multi-user systems. The awareness of the group affective tone, as shown in our study, has the potential to empower group members to take control over a degrading or erratic collaboration and guide it toward a more efficient configurations. Furthermore, affective agents like Pogat can, through their human-like representation and hint-based communication, have a principal role in managing feelings of frustration, supporting a positive group mood and even aiding conflict management in a team.

More importantly, as affective agents can be employed in a wide range of collaborative application (e.g., medicine, entertainment, emergency management, or visual analytics), it becomes clear that their effect is particularly valuable in sustaining the collaboration in contexts where *unanimous decisions* are of utmost importance. This is further supported by the link between a person’s ability of recognizing emotions in others and her/him making a good decision [11]. Consider, for example, architects working in a collaborative setting. When decisions can have such a wide impact like in this field, it is required that all team members agree on decisions and their results. As such, one team member who does not express a different point of view on a certain topic can have serious reper-

cussions. Yet, in most cases, such a repressed action results in emotions that can be perceived by our system, thus, it makes the group aware of a discrepancy in the team.

Besides the importance of a unanimous group decision depending on the domain and complexity of the task, collaborative scenarios are also defined by the size of the group. As such, a negative or inexistent GAT can also have varying relevance for small and large groups, as the impact of a single person in a larger group can be less than the one of a person in a smaller group. However, this does not constitute a rule, as some domains require all participants to reach a common solution or conclusion (e.g. medicine, architecture, etc.). Thus, Pogat is aimed mostly at collaborative scenarios where every potential disagreement needs to be expressed and analyzed in order to ensure precision and safety.

Because the Pogat system raises awareness of the GAT and not the emotional levels of select individuals, it also addresses a set of *privacy issues* by offering a convoluted view and avoids segregating a person as the source of a problem in the team. At the same time, one has to consider that certain group members might feel uncomfortable with sharing their affective states, even in a summative fashion.

5 Conclusion

In this paper we have focused on presenting Pogat, an affective virtual agent for supporting collaboration through increased group affective tone awareness. Besides allowing group members to be aware of the current emotional state of the team, Pogat also supports the transition of a group to a more positive GAT. The emotion acquisition in our system is done by employing a set of mobile EEG headsets and extracting the real-time emotional valence of each user. On the other side, the Pogat affective agent increases GAT awareness in the team by modulating its representation through facial expressions, colors and text-based hints. Our study suggests that our virtual agent helped users to increase their GAT awareness and improve their efficiency on the proposed collaborative tasks.

In future research we plan to extend the interaction modalities of the Pogat agent as well as further inspect the efficiency of various techniques for manipulating group affective tone to support collaboration.

References

1. Bartneck, C.: Interacting with an embodied emotional character. Proc. of the International Conference on Designing Pleasurable Products and Interfaces. ACM Press, Pittsburgh, USA, 55–60 (2003)
2. Burleson, W., Picard, R.W.: Affective agents: sustaining motivation to learn through failure and a state of stuck. Workshop on Social and Emotional Intelligence in Learning Environments (2004)
3. Cernea, D., Olech, P.-S., Ebert, A., Kerren, A.: EEG-Based Measurement of Subjective Parameters in Evaluations. Proc. of the 14th International Conference on

- Human-Computer Interaction (HCII 2011), poster paper, volume 174 of CCIS, Springer, Orlando, Florida, USA, 279–283 (2011)
4. Cernea, D., Olech, P.-S., Ebert, A., Kerren, A.: Measuring Subjectivity - Supporting Evaluations with the Emotiv EPOC Neuroheadset. *Journal for Artificial Intelligence (KI)*, volume 26, number 2, 177–182 (2012)
 5. Cernea, D., Weber, C., Ebert, A., Kerren, A.: Emotion Scents – A Method of Representing User Emotions on GUI Widgets. *Proc. of the SPIE 2013 Conference on Visualization and Data Analysis (VDA 2013)*, volume 8654, IS&T/SPIE, Burlingame, CA, USA (2013)
 6. Cummings, A.: Contextual characteristics and employee creativity: Affect at work. *Proc. 13th Annual Conference, Society for Industrial Organizational Psychology*. Dallas, USA (1998)
 7. Egges, A., Kshirsagar, S., Magnenat-Thalmann, N.: Generic personality and emotion simulation for conversational agents: Research Articles. *Comput. Animat. Virtual Worlds* 15.1, 1–13 (2004)
 8. Fabri, M., Moore, D.J., Hobbs, D.J.: Empathy and enjoyment in instant messaging. *Proc. of 19th British HCI Group Annual Conference (HCI2005)*, Edinburgh, UK, 4–9 (2005)
 9. Garcia, O., Favela, J., Machorro, R.: Emotional awareness in collaborative systems. *String Processing and Information Retrieval Symposium (1999)*, International Workshop on Groupware, 296–303 (1999)
 10. George, J.M.: Group affective tone. *Handbook of work group psychology*. Chichester, UK: Wiley, 77–93 (1996)
 11. Goleman, D.: *Emotional Intelligence. Why it can matter more than IQ*. Bantam Books, New York (1995)
 12. Hone, K., Akhtar, F., Saffu, M.: Affective agents to reduce user frustration: the role of agent embodiment. *Proc. of Human-Computer Interaction (HCI2003)*, Bath, UK (2003)
 13. Huisman, G., van Hout, M., van Dijk, E., van der Geest, T., Heylen, D.: LEMtool: measuring emotions in visual interfaces. *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*, ACM Press, New York, NY, USA, 351–360 (2013)
 14. Isen, A.M., Daubman, K.A., Nowicki, G.P.: Positive affect facilitates creative problem solving. *Journal of Personality and Social Psychology*, volume 52, 1122–1131 (1987)
 15. Kelly, J.R., Spoor, J.R.: Affective Influence in Groups *Proc. of the 8th Annual Sydney Symposium of Social Psychology*, Sydney, Australia (2005)
 16. Liu, Y., Sourina, O., Nguyen, M.K.: Real-Time EEG-Based Human Emotion Recognition and Visualization. *Proc. of the 2010 International Conference on Cyberworlds (CW)*, 262–269 (2010)
 17. McDuff, D., Karlson, A., Kapoor, A., Roseway, A., Czerwinski, M.: AffectAura: an intelligent system for emotional memory. *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)*, ACM Press, New York, NY, USA, 849–858 (2012)
 18. de Melo, C.M., Carnevale, P., Gratch, J.: The effect of expression of anger and happiness in computer agents on negotiations with humans. *Proc. of 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3 (AAMAS '11)*, Vol. 3, 937–944 (2011)
 19. Russell, J.A.: A circumplex model of affect. *Journal of Personality and Social Psychology*, volume 39, 1161–1178 (1980)

A Virtual Suspect Agent’s Response Model

Merijn Bruijnes, Sjoerd Wapperom, Riëks op den Akker, and Dirk Heylen

`m.bruijnes@utwente.nl`

Human Media Interaction, University of Twente
PO Box 217, 7500 AE, Enschede, The Netherlands

Abstract. We develop a computational interpersonal affective response model for virtual characters that act as suspect in a serious game for training interviewing (interrogation) skills to police officers. We implemented a model that calculates the responses of the virtual suspect based on theory and observation. We describe the aspects of the move (question asked) by the police interviewer that we distinguish and how the suspect responds to the move. This response is dependent on static personality characteristics of the suspect character (persona) and on the dynamic state of the interaction. We evaluated it by means of our test, the “Guess who you are talking to?” test, showing the response model can portray a personality in a recognizable manner.

Keywords: Response Model, Virtual Agent, Affective Agent, Police Interview, Social Simulation

1 Introduction

We work towards a virtual agent that can play a suspect in a serious game that can be used by police students to hone their skills in police interviewing. A virtual agent needs three main components to be able to have a meaningful interaction. The actions of the user have to be sensed and interpreted (e.g. the user says “Confess, criminal!” which is interpreted in the abstract terms dominant and aggressive behaviour). This interpretation provides the input to a response model that provides the reasoning of the agent (e.g. the user is dominant and aggressive which makes me sad and angry). A response model should take into account the specific role that the agent plays. In this case that is a suspect with all the tactics and psychological manoeuvring that is involved. A response model based on human behaviour can be used to make the behaviour of a virtual agent more believable to humans. Based on the state of the response model the agent can select the most appropriate behaviour in its repertoire (e.g. the abstract state of the response model is sad and angry, so make a sad face and say “You’re not nice!”). The human responds to the agent and the cycle continues. In this paper we present a response model for such a virtual suspect agent.

Realistic agent behaviour can elicit learning in a user by experiencing the interaction. Architectures for social agents (e.g. [7, 11, 15]) often place emphasis on the reasoning (goals, planning, actions), emotion (appraisals, mood, emotion),

and dialogue (grammar, utterances) of an agent. All this to increase the ‘positive things’ in an interaction with the user, affiliation, cooperation, respect, coordination, understanding, etc. However, for a learning application it can be beneficial to have an agent decrease the positive things in an interaction to facilitate learning by making mistakes. A virtual agent can allow the user to make mistakes by being non-cooperative. However, the agent needs to do more than simply refuse to give in or show behaviour that the user was tasked to prevent [17]. The agent should consider the goals that fit the role it is enacting and the goals of the tutoring application in which it serves [3]. In a training application it is important for the system to have the ability to explain its reasoning [6]. Such ‘explainable intelligence’ can lead to learning by reflecting on the interaction [8]. Our model can provide the information needed to explain its behaviour. During the interaction the model has states and state transitions, a log of these provides information on the interaction that the user had. The user can use this information to evaluate his interaction as it provides insight into why the interaction went the way it went. For example, the user could compare his intentions with the way the agent interpreted his intentions.

We developed a response model that can ‘play’ a suspect that has a ‘personality’ (a persona). It simulates a persona and models the interpersonal aspects of an interaction in an abstract manner. It calculates the interpersonal properties that the response of the suspect should have, based on the interpretation of the contribution by the user.

1.1 Related Work

Several other researchers looked at building computational models of the mind of agents such as suspects, that is agents that are not fully cooperative in interaction. Roque and Traum [14, 17] distinguish three levels of compliancy: compliant, reticent and adversarial. “When characters are compliant, they provide information when asked, but fall short of Gricean cooperativity because they don’t provide helpful information that was implicated rather than explicitly solicited. When characters are reticent, they provide neutral information, but will evade any questions about important or sensitive information. When characters are adversarial, they provide deceptive or untruthful answers.” [17](p67). In [12], Olsen describes a system that can teach police students to build rapport while maintaining professionalism, listen to verbal cues and detect important changes in both verbal and non-verbal behaviour. A list of 400 predefined questions are available for the police officer to chose from. The simulated suspect responses are given based on the question and the internal state of the suspect. The internal state consists of the mood of the suspect (angry, denial or compliance) and the rapport between the suspect and user. Luciew et al. [10] build an interview and interrogation immersive learning simulation, specifically to train police officers in interviewing children who were victims of sexual abuse and interrogate suspects on that matter (i.e. two prototype systems were developed). In this system the behaviour of the agent is dependent largely on the proficiency of the user in detecting non-verbal cues and reporting them outside the interaction. Topic of

the questions seems to be the only direct influence the user has during the interaction on the behaviour of the agent.

Reisenzein et al. [13] discuss how computational modelling of emotion benefits from the exchange of ideas and practices from psychology and computer science. They propose emotion theories should be deconstructed into their basic assumptions to be able to construct a more unified or standardized conceptual system or implementation. We are interested in interpersonal and social workings of an interaction (in a police interview) and do not focus on emotion. However, the idea of deconstructing social and interpersonal theories into their basic assumptions has beneficial results. In the next section we describe what we include in the response model for the suspect agent based on observed interactions in police interviews. The interpersonal concepts we include were selected by deconstructing the social theories that describe a police interview into the basic concepts from these theories.

1.2 Interactions in Police Interviews

Police interviewing is a skill that revolves around making an often uncooperative suspect cooperate. The Dutch National Police uses a theory of interpersonal stance (Leary’s rose) that consists of the concepts of dominance and affiliation [9]. Students of the Police Academy get the opportunity to practice their interview skills with a professional suspect actor in role-playing exercises after studying the theory of interpersonal stance. The Dutch Police Interview Training corpus (DPIT-corpus) is a corpus of such role-played police interviews [1]. We analysed the DPIT-corpus (in [4]) to get insight into the social behaviour of police officers and suspects in the police interview setting. We collected many terms that people use to describe the interactions in the corpus. A factor analysis revealed factors that could be interpreted as relating to the theories of *interpersonal stance* [9], *face* [2], and *rappport* [16] and the meta-concepts *information* and *strategy*. These theories provide a way to describe the interaction in a police interview. Each of these theories and meta-concepts is a collection of concepts (see Table 1) and all these concepts are relevant in police interviews. Therefore, we argue that these concepts are necessary to include in a response model for a virtual suspect that captures the social interactions of that suspect in a police interview. Next, we present the response model that we constructed for a virtual suspect.

2 Suspect Response Model

To present our response model, we use the abstract interview simulation that is used in the testing of the model as an illustration. We start with a description of the the static variables that make up a persona in our model and the variables that serve as input to the model. Next, we present the instance of the model that holds the ‘current response model state’ and how this state is updated based on the input, personality, and state. We finish with a description of the possible outputs of the response model based on the updated state.

Table 1. Concepts within the theories *stance*, *face*, and *rapport* and the meta-concepts *information* and *strategy* that were found relevant in police interviews [4].

Stance	Face	Rapport	Information	Strategy
Friendly (Dominant-Together)	Autonomy+	Coordination	Questioning	Confront
Aggressive (Dominant-Opposed)	Approval+	Attention	Give info	Surround
Withdrawn (Submissive-Opposed)	Autonomy-	Positivity	Lie	Evade
Dependent (Submissive-Together)	Approval-		Withhold info	Annoy
			Frame/topic	

2.1 Persona Specification

The persona the response model portrays consists of a set of static variables that influence the calculations that update the state and the response of the model. A persona consists of five settings based on interpersonal stance, rapport, face-threatening topics, and information (see Fig 1): 1) A preferred interpersonal stance that might be considered as a ‘personality’ and can have the values: **Friendly**, **Aggressive**, **Withdrawn**, or **Dependent**. It influences how fast interpersonal stance, mood, and rapport change. 2) Dominance and affiliation settings state the initial stance of the suspect. For example, an aggressive suspect has positive dominance (dominant) and negative affiliation (opposed). 3) The sensitivity to rapport states how effective rapport building is with this persona. 4) The attitude the suspect has towards being met with an opposed or aggressive stance means how strongly he reacts to negative action by the police and how easily he turns to aggression himself. 5) Finally, the suspect’s sensitivity to internal and external pressure determine whether he will lie about guilt sensitive topics or not and what approach would be best to make the suspect break. Internal pressure rises with feelings of guilt. External pressure rises when the police officer puts pressure on the suspect, for example by showing proof of guilt. To illustrate the model we use a persona that is ‘aggressive, dominant, sensitive to rapport, very sensitive to being opposed, and low sensitivity to pressure’.

2.2 Interaction with the Response Model

The response model receives input from (automatic or manual) interpreters of the contribution to the interaction by the user. We call this set of input-variables the *Question Frame* (QF). The QF consists of nine aspects that describe the question being posed (see Fig 1: Question Frame): 1) The interpersonal stance [9] of the police officer during this contribution, can be: **Friendly**, **Aggressive**, **Withdrawn**, or **Dependent**. 2) The question type is based on the meta-concept *information* and can be: **Open**, **Yes/No**, **Probing**, **Leading**, **Forced Choice**, or **Statement**. 3) Topic threat describes how face-threatening the topic for the suspect [2]. This can be: **Low**, **Medium**, **High**, or **Guilt Indication**. Low, medium, or high relate to the threat to topics not related to the crime. The last indicates an utterance with which the suspect is related to the crime, for example

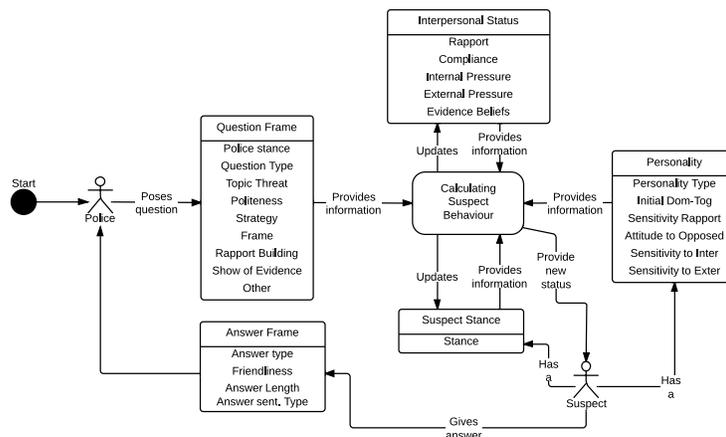


Fig. 1. The flow of an interview and the information in the system. See main text for details.

“You were seen at the gas station that was robbed yesterday!”. 4) Politeness is related to the politeness strategy used to mitigate a face-threat [2] and can be: **Direct, Approval Oriented, Autonomy Oriented, or Off Record.** 5) Strategy is based on the meta-concept *strategy* and can be: **Being Kind, Being Equal, Emotional Appeal, Intimidating, Direct Pressure, or Rational Convincing.** 6) Dutch police officers go through two phases during the interview: a **person related frame** that covers the personal life of the suspect and a **case related frame** that covers topics related to the case. 7) Rapport building [16] can be done by showing: **Attention, Positivity, and Coordination.** The amount of rapport the suspect experiences with the user is updated with every contribution of the user. 8) Showing evidence can pressure the suspect into confessing. It can be: **None, Low, or High.** 9) The ‘Other’ attribute is used for special occasions: **Confronting a Lie, Repeating the Question, or Accusing.** For example, the user says “*I know it’s hard to talk about, but it would help me if you tell me if you were at the crime scene*” which is interpreted as “**Friendly, High Topic Threat, . . . , Autonomy Oriented politeness**”.

The instantiated response model holds the state of the suspect and the state of the interaction. It consists of the variables: the current rapport the suspect experiences with the police officer, his current stance towards the police officer, the current state of compliance of the suspect (**Compliant or Aggressive**), his internal and external pressure, his beliefs about the amount of evidence against him, and the static personality traits (see Fig 1). For our example persona, this is initially “**Low Rapport, Aggressive Stance, Aggressive compliance, Low Pressure, and Low Evidence Believes**” based on his personality.

The response model’s state is updated when a new QF comes in. The **rapport** between the two increases if a rapport building action is performed. Rapport

decreases if no rapport building action is performed. The reduction is bigger when no rapport is build during the person related frame. The reduction is biggest with an intimidating strategy. Next, the new **Stance** of the suspect is calculated, taking into account: suspect’s old stance and preferred stance, and the police officer’s rapport building, topic threat, politeness and applied strategies. The ‘togetherness’ of the suspect increases if the police officer takes an dominance stance that is opposite from the preferred dominance stance of the suspect (moving the suspect towards a **Friendly** or **Dependent** stance). The ‘togetherness’ increases if rapport is being build, the topic is not threatening, the strategy is **Being Kind**, **Being Equal**, or **Emotional Appeal**. The ‘dominance’ of the suspect increases if the police officer uses a threatening topic, strategy, or stance. The size of increases/decreases varies depending on the personality (the sensitivity to: rapport, opposed behaviour, and pressure). The **Compliance** is updated based on the previous state of compliance, the new stance of the suspect, and the strategy employed by the police officer. The compliance can have two variables: **Compliant** or **Aggressive**. Both receive a score based on the input, moderated by personality, and the value with the highest score wins. For example, an aggressive personality scores **Aggressive** stronger than a non-aggressive personality when confronted with an **Intimidating** strategy. Next, the the **Internal and External Pressure** are calculated based on the sensitivity to pressure of the suspect, police officer’s strategy, and the optional fields: **Confronting a Lie** and **Repeating the Question**. The internal pressure increases when the police officer employs a friendly strategy like **Emotional Appeal**, where external pressure rises most with strong strategies like **Intimidating** or information related tactics like **Confronting a Lie**. The pressure is dropped to zero when the suspect tells the truth (see next paragraph). Finally, the suspect’s **Evidence Beliefs** increases if new evidence has been provided by the police and when the suspect tells the truth about a guilt indicative topic. For our example, the initial state is updated towards “**Higher Rapport, less Aggressive Stance, more compliance, Low Pressure, and Low Evidence Believes**”. The user is being friendly and the response model reflects this, even if the persona is very unfriendly.

The response model provides the interpersonal properties the response should have in the form of an *Answer Frame* (AF) (Fig 1). This frame contains four aspects that describe the answer of the suspect: 1) The **Answer Type** is related to the information strategy used by the suspect and can be: **Truth, Lie, Avoid, or Aggression**. 2) **Friendliness** is related to stance and can be: **Friendly, Neutral, or Unfriendly**. 3) **Answer Length** is also related to the information strategy (**Long, Short, One Word, or Silence**). 4) **Answer Sentence Type** is related to the question type being posed and the way the suspect wishes to answer to this type and can be: **Open Telling, Counter Question, Aggressive Expression, Yes/No, Play Dumb, Probing Answer, or Ignore**. The example response is “**Aggressive answer type, and an Unfriendly, Short, and an Aggressive Expression**. The agent can use the information in the AF and the state of

the response model to select the most appropriate behaviour in its repertoire. The user can respond to this by asking another question and the cycle continues.

3 Method for Evaluation of Response Models

We want to know whether our response model can portray a persona in a recognizable and consistent way using our “Guess who you are talking to?” test (see [5]). Participants interact with the response model and have to guess which of a selection of personas is portrayed by the system. This interaction is done in the (abstract) terms of the response model. However, this comes at a cost. The participants need to be instructed on the abstract factors that the model uses and the personas that are portrayed by the model. Three personas were created, based on personas from the DPIT-corpus [1, 4]. Each persona was introduced in a short text. The participants have at least two sessions of interactions with the response model, once with one of the personas and once with a random response generator (not based on a persona or response model). During each session they are asked to indicate with which of the personas they think they are interacting.

3.1 Results of Evaluation

For our evaluation, 48 participants (42 male, mean age 24.8 with SD 3.7) took part in the study. A total of 39 (81.25%) participants guessed correctly with which persona they were interacting after eight interactions. Participants who were correct were (significantly: $Z = -2.001, p < 0.1$) more confident (4.41) compared to the participants who were incorrect (3.67) (rated on a 5-point Likert scale (1=strongly disagree, 5=strongly agree)). The realism rating was similar: 3.90 for correct compared to 3.89 for incorrect. In the interactions where the responses of the system were random we might expect that each of the personas would be chosen an equal number of times (33%). However, the distribution of choices for the personas was 62.5%, 20.8%, and 16.7%. The average confidence level for interactions with personas was significantly higher 4.27 (SD = 0.76) compared to 3.46 (SD = 0.77) for the random interactions ($Z = -4.2, p < 0.00$). The average level of realism for personas was significantly higher 3.90 (SD = 0.52) compared to 3.35 for random rounds (SD = 0.89) ($Z = -3.7, p = 0.001$).

4 Conclusion

The results of this “Guess who you are talking to” test give an indication that our response model generates responses to user actions in such a way that the user is able to recognize a persona. This gives evidence of the validity of the response model and it promises that the model can be used in the implementation of believable virtual suspect characters with various personal characteristics as we encountered in our police interview corpus.

Acknowledgements This publication was supported by the Dutch national program COMMIT.

References

1. op den Akker, R., Bruijnes, M., Peters, R., Krikke, T.: Interpersonal stance in police interviews: content analysis. *Computational Linguistics in the Netherlands Journal* 3, 193–216 (2013)
2. Brown, P., Levinson, S.C.: *Politeness: Some universals in language usage*. Cambridge University Press, Cambridge (1987)
3. Bruijnes, M., Kolkmeier, J., op den Akker, H., Linssen, J., Theune, M., Heylen, D.: Keeping up stories: design considerations for a police interview training game. In: *Proceedings of the Social Believability in Games Workshop (SBG2013)*. p. 14. CTIT, University of Twente, Enschede, The Netherlands (2013)
4. Bruijnes, M., Linssen, J., op den Akker, R., Theune, M., Wapperom, S., Broekema, C., Heylen, D.: Social behaviour in police interviews: Relating data to theories. In: *Conflict and negotiation: Social research and machine intelligence* (2014)
5. Bruijnes, M., Wapperom, S., op den Akker, R., Heylen, D.: A method to evaluate response models. In: *IVA2014* (in press)
6. Core, M.G., Lane, H.C., Van Lent, M., Gomboc, D., Solomon, S., Rosenberg, M.: Building explainable artificial intelligence systems. In: *Proceedings of the National Conference on Artificial Intelligence*. vol. 21 (2006)
7. Dias, J., Mascarenhas, S., Paiva, A.: Fatima modular: Towards an agent architecture with a generic appraisal framework. In: *Proceedings of the International Workshop on Standards for Emotion Modeling* (2011)
8. Kooops, M., Hoevenaar, M.: Conceptual change during a serious game: Using a Lemniscate Model to compare strategies in a physics game. *Simulation & Gaming* (2012)
9. Leary, T.: *Interpersonal Diagnosis of Personality: Functional Theory and Methodology for Personality Evaluation*. Ronald Press, New York (1957)
10. Luciew, D., Mulkern, J., Punako, R.: Finding the truth: Interview and interrogation training simulations. In: *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (2011)
11. Marsella, S.C., Gratch, J.: Ema: A process model of appraisal dynamics. *Cognitive Systems Research* 10(1), 70–90 (2009)
12. Olsen, D.: Interview and interrogation training using a computer-simulated subject. In: *The Interservice/Industry Training, Simulation & Education Conference* (1997)
13. Reizenzein, R., Hudlicka, E., Dastani, M., Gratch, J., Hindriks, K., Lorini, E., Meyer, J.: Computational modeling of emotion: Toward improving the inter- and intradisciplinary exchange. *Affective Computing, IEEE Transactions on* 4(3), 246–266 (2013)
14. Roque, A., Traum, D.: A model of compliance and emotion for potentially adversarial dialogue agents. In: *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*. pp. 35–38 (2007)
15. Steunebrink, B., Vergunst, N., Mol, C., Dignum, F.P.M., Dastani, M., Meyer, J.: A generic architecture for a companion robot. In: *Proceedings of the 5th International Conference on Informatics in Control, Automation and Robotics (ICINCO08)* (2008)
16. Tickle-Degnen, L., Rosenthal, R.: The nature of rapport and its nonverbal correlates. *Psychological inquiry* 1(4), 285–293 (1990)
17. Traum, D.: Non-cooperative and deceptive virtual agents. *IEEE Intelligent Systems* 27(6), 66–69 (2012)

Length of Smile Apex as Indicator of Faked Expression

J. Dean McDaniel¹ and Mei Si²

¹Computer Science Department, ²Cognitive Science Department,
Rensselaer Polytechnic Institute, Troy, USA
{mcdanj2, sim}@rpi.edu

Abstract. Facial expressions are important cues of people’s emotions, attitudes, and intentions. Smiling is one of the most common facial expressions, often associated with a welcoming, positive attitude. In social interactions, people sometimes fake their smiles for this effect. Research has shown that one difference between faked and genuine smiles is that faked ones often have longer apexes. In this work, we explore creating fake and genuine smiles for virtual humans. We systematically varied the apex length and examined the effect of this manipulation on smiles of different lengths and on both male and female faces. Using mTurk, 40 subjects rated these smiles on their genuineness and fakeness. Contrary to previous findings, our results suggest that smiles of longer apex time are perceived as more genuine and less fake than smiles of shorter apex for virtual characters. This paper presents the design of the experiment followed by the results and discussion.

Keywords: emotions · facial expressions · perception of emotions · smile

1 Introduction

Virtual human characters have received increasing attention in recent years. People can interact with them, either conversationally or physically, to get information, practice social interaction skills, receive training, or be entertained [5], [7], [9, 10]. Often, we want our virtual human characters to appear genuine, sincere, and welcoming to the user.

Smiling is one of the most common facial expressions and is often associated with a welcoming and positive attitude. Factors related to how people smile have become an increasingly important research topic because of its observed use in both positive and aversive environments [1, 2]. Because smiles are easily discernible from other expressions, one may present a smile when lying or when being insincere [2]. This leads to the distinction between at least two types of smiles: genuine and fake. According to Ekman and Friesen, fake expressions arise when people learn to interrupt their natural emotional response and instead present a voluntary, masked expression [2].

Further, Ekman and Friesen pointed out that genuine and fake smiles differ in onset, apex, and offset timing [2]. In a faked smile, the apex is usually too

long, making the person appear to intentionally hold the expression. Thus, the onset time falls short, and the smile appears on the face abruptly. The offset time of a fake smile is also in some way irregular, indicating the person has stopped intentionally holding the expression [2].

In this work, we examine the effect of varying the duration of smile apex with respect to its onset duration for virtual humans. We created the smiles using the VHuman Toolkit from the Institute for Creative Technologies (ICT) at the University of Southern California [4]. We tested using two smiling virtual human characters (male and female) and for three different lengths of the smile: 3, 5, and 7 seconds. We chose to study longer smiles because people often exhibit longer smiles and other conversational agents, such as robots, may have slower-moving faces. Contradicting to our initial hypothesis based on Ekman and Friesen’s findings, the results indicate that smiles with longer apex time are perceived as more genuine and less fake than smiles of shorter apex.

In the next sections, we first summarize related works on genuine and fake smiles, then present our empirical study examining the effects of different smiling factors on participants’ perception of genuineness and fakeness of the virtual human. We discuss the implication of the results followed by a plan of future work.

2 Related Work

Ekman and Friesen examined the facial muscles used in smiling and identified two major muscles are being used: zygomatic major, extending the lip corners, and orbicularis oculi, raising the cheek and tightening the lower eyelid [2]. Along with Ancoli, they found a significant correlation between happiness and the frequency, duration, and intensity of the zygomatic major facial muscle’s movements [3].

In this work, we concentrate on studying the impact of the duration of the smiles for expressing genuineness in virtual humans because it is one of the most studied factors. The duration of a smile, like other expressions and behaviors, can be separated into onset, apex, and offset timing [2], [6]. A smile’s onset time is the duration from the start of the smile to its apex. The apex duration is when the smile is at its most intense, and the offset of the smile is the span of time from apex until all evidence of the smile is absent from the face.

Through analyzing data collected from observing how human subjects smile, Ekman and Friesen found that genuine and fake smiles differ in onset, apex, and offset timing [2]. Fake smiles usually have longer apexes and shorter onset and offset times that make the expression appear suddenly and remain for a prolonged time. The faked smiles then disappear from the face in a similar, quick manner [2].

Ekman and Friesen’s finding has been confirmed in a few research studies using virtual characters. Ochs, Niewiadomski, and Pelachaud [8] let participants create smiles on a virtual agent, and the majority used shorter onset and offset times for fake smiles and longer times for genuine smiles. Krumhuber and Kappas [6] manipulated the durations of onset, apex, and offset times and found that

smiles with longer onset and offset (closer to half of a second) were judged as significantly more genuine than their shorter counterparts (closer to one-tenth of a second). Similarly, they found that a smile lost authenticity the longer its apex was held, with apexes closer to 1 second being judged as significantly more genuine than smiles with apexes closer to 5 seconds.

In Krumhuber and Kappa’s work, the smiles were brief and presented to the subjects without a dialogue context [6]. In this work, we want to further examine the effect of apex length on the genuineness of a virtual character’s smile when the smiles are longer and when the virtual characters are initializing a dialogue with the user.

We encoded the virtual characters’ smiles using the Facial Action Coding System (FACS), which was invented by Ekman and Friesen [2] and has been widely used for labeling facial expressions. Ekman and Friesen mapped the zygomatic major muscle to action unit AU12, the lip corner puller. They noted that when used extremely, AU12 could cause change similar to AU6, the cheek raiser. Orbicularis oculi was mapped partially to AU6 and partially to AU7, the lower eyelid tightener [2]. In our study, these action units were manipulated together to create different smiles.

3 Experimental Design

3.1 Participants

The participant sample was taken from registered workers on Amazon Mechanical Turk (MTurk) who have completed 100 or more tasks through the crowdsourcing system prior to this study and who have an approval rating of 90% or greater. The sample size is 40 adults (19 men, 21 women; mean age = 31.95 years, range = 20 to 66 years).

3.2 Materials and Procedure

We created a series of smiles on virtual human characters by the use of the Virtual Human Toolkit developed at the Institute for Creative Technologies (ICT) at the University of Southern California [4]. Two characters were employed for the study, the default male character Brad and the default female character Rachel, both pictured in Figure 1. Each character was programmed to give smiles with a shorter apex time than onset, intended to be genuine, and smiles with a longer apex time than onset, intended to be fake. Each type of smile lasted either 3, 5, or 7 seconds long. We chose longer smile durations than previous studies because we plan on replicating the work on a social robot, which will need to present longer smiles in order to sync with other bodily movement. The onset, apex, and offset times for each smile are listed in Table 1. We created smiles with identical durations for both the male and female characters. In total, 12 videos of smiles were created.

The action units employed for both genuine and fake smiles were units 6, 7, and 12, as suggested by Ekman and Friesen [2]: the cheek raiser, the lower eyelid

Table 1. Smile onset, apex, and offset times, in seconds.

Total	Onset	Apex	Offset
3	1.2	0.6	1.2
3	0.5	2.0	0.5
5	2.1	0.8	2.1
5	1.0	3.0	1.0
7	3.0	1.0	3.0
7	1.5	4.0	1.5

**Fig. 1.** The male and female virtual human characters from the VHuman Toolkit

tightener, and lip corner puller, respectively. When the character begins to smile, all action units are gradually updated in intensity to reach the predefined apex intensity by the end of the specified onset time. We linearly interpolated the intensity during onset. After the character holds the smile at apex, the action units begin to gradually decrease to return to a neutral state. Similarly, we linearly interpolated the intensity during offset. Figure 2 shows an example of how the intensities of these action units are updated over time using a 3 second smile intended to be genuine—that is, with a shorter apex time than onset time.

To put the smiles into context, in each video the male or female character showed their smiles along with simple verbal statements. They first say, “My name is,” followed by the character’s name and a smile, then, “Have a good day,” followed by a smile. The two smiles are identical, i.e. they have the same duration and are both genuine or both fake. Videos of the smiles were administered to human participants who rated the smiles on how genuine, welcoming, and felt the smiles appeared, as well as how false, fake, and forced they appeared. The rating questions for each video were randomly ordered to prevent any ordering effect.

For participating in this study, participants completed a survey in which they watched 12 videos of virtual characters smiling in random order and rated the smiles on 5-point Likert scales for the following metrics: genuine, welcoming, felt, fake, false, and forced.

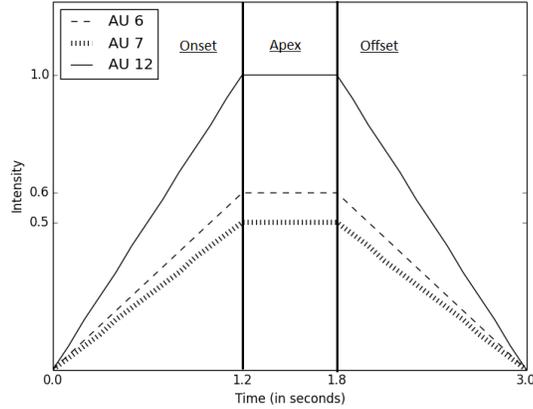


Fig. 2. The intensity of action units in generated smiles over time for a 3 second smile with shorter apex than onset Time

4 Results

For both virtual characters, scores on the 5-point Likert scales were summed into two overall scores: “genuineness” for genuine, welcoming, and felt ratings and “fakeness” for fake, false, and forced ratings. Two-way repeated-measures ANOVAs were performed using SPSS. The independent variables are smile apex length with two levels (shorter apex than onset, longer apex than onset) and the overall duration of the smile with three levels (3 seconds, 5 seconds, and 7 seconds). Four ANOVA tests were conducted: the male’s smiles rated as genuine, his smiles rated as fake, and likewise for the female’s smiles. We used an alpha level of .05 for all statistical tests. The interaction effect between the length of apex and the duration of the smile was significant in all of the ANOVA tests. We plotted the interaction effects in Figure 3.

The F ratios for main effects are reported in Table 2. We found that for both the male and female faces, apex length is a significant factor for the subject’s ratings of how genuine and fake the smiles appear—the main effect of apex length is significant in all four ANOVA. However, contrary to our initial expectation, the longer the apex length is, the more genuine and less fake the subjects rated the videos to be.

Similar to smile apex, overall duration of the smile is a significant factor for all ANOVA tests, except for the ratings of fakeness with the male smile. In general, smiles lasting 7 seconds were rated as most fake, and smiles lasting 5 seconds were perceived as most genuine.

Post-hoc comparison using Fisher’s test show that the mean ratings of smile genuineness were significantly higher in smiles of longer apex time for both the male and the female face than for the smiles with shorter apex time (3.18 vs. 2.93

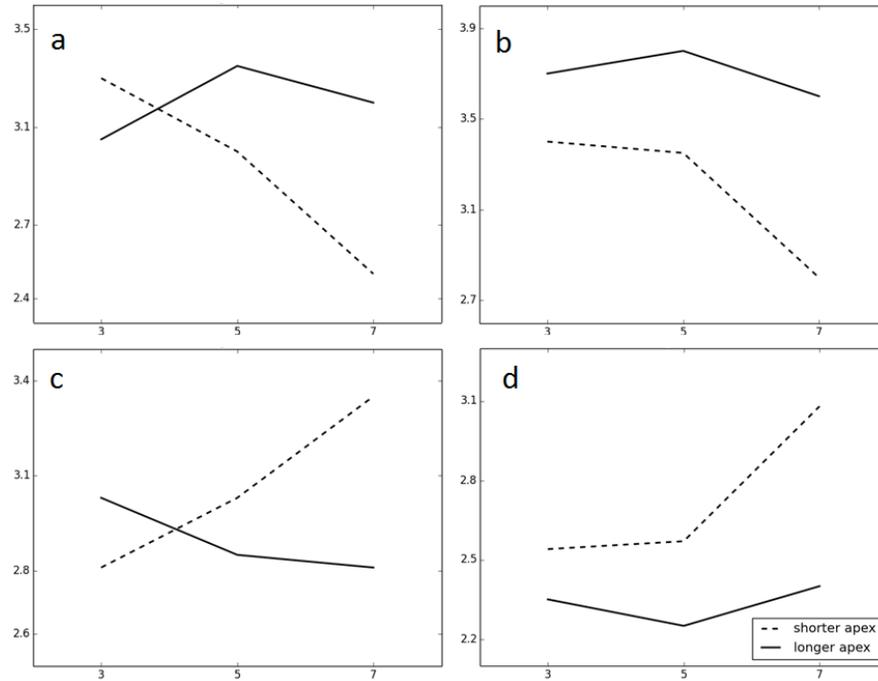


Fig. 3. Interaction effect between apex Length and duration for ratings: a) genuineness of the male face, b) genuineness of the female face, c) fakeness of the male face, d) fakeness of the female face

Table 2. *F* ratios for main effects and interaction effect across all ANOVA

	male genuineness	female genuineness	male fakeness	female fakeness
apex length	$F(1, 110) = 15.17^{**}$	$F(1, 112) = 82.28^{**}$	$F(1, 107) = 5.94^*$	$F(1, 114) = 24.59^{**}$
duration	$F(2, 220) = 7.38^{**}$	$F(2, 224) = 12.05^{**}$	$F(2, 214) = 1.74$	$F(2, 228) = 5.99^{**}$
interaction	$F(2, 220) = 16.07^{**}$	$F(2, 224) = 8.00^{**}$	$F(2, 214) = 8.81^{**}$	$F(2, 228) = 4.21^*$

* Significant at the .05 level. ** Significant at the .01 level.

for the male face; 3.69 vs. 3.20 for the female face). Consistently, mean ratings of smile fakeness were significantly higher in the smiles with shorter apex time for both the male and the female faces than the smiles with a longer apex time (3.07 vs 2.91 for the male face; 2.72 vs 2.34 for the female face).

In terms of the overall durations of the smiles, for both the male and female characters, ratings of smile genuineness were significantly higher in smiles lasting 3 or 5 seconds than smiles lasting 7 seconds in smiles on both faces (3.15 and 3.17 vs. 2.84 for the male face; 3.56 and 3.59 vs. 3.20 for the female face). For both characters, a comparison between smiles of 3 and 5 seconds yielded a non-significant difference.

Ratings of fakeness for the female smile were significantly higher in smiles lasting 7 seconds than 3 seconds or 5 seconds (2.72 vs. 2.44 and 2.43). Smiles of 3 and 5 seconds are not different in terms of their ratings of fakeness. For the male face, smile duration does not affect the ratings of fakeness.

5 Discussion and Future Work

Contrary to Ekman’s previous studies, smiles with a longer apex time than onset seems more genuine and less fake to our subjects. This trend is more significant for smiles that last 7 seconds than shorter smiles of 3 and 5 seconds, and it is also consistent across two different virtual human faces—one being male, and the other female.

On the other hand, this somewhat agrees with findings by Ekman and Friesen in which participants “who saw just the facial expressions judged people to be most truthful when they were, in fact, lying” [1]. Although in this study we put the character’s smiles in the context of simple verbal statements, we suspect the statements were too simple to provide the participants a context for properly judging the smiles.

Practically, the results of this study suggests that when virtual human characters are used in a relatively simple social context and when their smiles last longer than 3 seconds, they should display smiles with a longer apex in order to appear more genuine to users.

Future work will examine more closely the distinction between genuine and fake smiles. We plan to study the phenomenon again when the smiles are embedded in a much richer social context, including a social robot capable of displaying facial expressions. In addition, we also plan to study manipulation of action unit classification and laterality. For example, Ekman and Friesen indicated that when people present fake smiles, they consciously raise their lip corners but usually fail to conceal other involuntary facial cues [2].

6 Conclusion

Having appropriate facial expressions is important for virtual humans. In this work, we investigated the difference between genuine and fake smiles, which has

been extensively studied in human facial expressions. Expression apex time was shown to have a significant effect on participants' observations and rating of smile genuineness and fakeness. The results from this study indicate that in a simple, social interaction context with relatively long smile durations (> 3 seconds), smiles with a longer apex time than onset are perceived as more genuine for virtual humans. Future work has been planned for further study of this phenomenon.

References

1. Ekman, P., Friesen, W.V.: Detecting Deception from Body or Face. *Journal of Personality and Social Psychology*, vol. 29(3), 288–298 (1974)
2. Ekman, P., Friesen, W.V.: Felt, False, and Miserable Smiles. *Journal of Nonverbal Behavior*, vol. 6(4), 235–252 (1982)
3. Ekman, P., Friesen, W.V., Ancoli, S.: Facial Signs of Emotional Experience. *Journal of Personality and Social Psychology*, vol. 39(6), 1125–1134 (1980)
4. Harthold, A., Traum, D., Marsella, S., Shapiro, A., Stratou, G., Leuski, A., Morency, L., Gratch, J.: All Together Now: Introducing the Virtual Human Toolkit. In: *International Conference on Intelligent Virtual Humans*. Edinburgh (2013)
5. Johnson, K., Raiji, A., Stevens, A., Lind, D.S., Lok, B.: The Validity of a Virtual Human Experience for Interpersonal Skills Education. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1049–1058. ACM (2007)
6. Krumhuber, E., Kappas, A.: Moving Smiles: The Role of Dynamic Components for the Perception of the Genuineness of Smiles. *Journal of Nonverbal Behavior*, vol. 29(1), 3–24 (2005)
7. Niewiadomski, R., Pelachaud, C.: Model of Facial Expressions Management for an Embodied Conversational Agent. In: *2nd International Conference on Affective Computing and Intelligent Interaction (ACII)*. Lisbon, Portugal (2007)
8. Ochs, M., Niewiadomski, R., Pelachaud, C.: How Should a Virtual Agent Smile? Morphological and Dynamic Characteristics of Virtual Agent's Smiles. In: *Intelligent Virtual Agent Conference (IVA)*. Philadelphia (2010)
9. Rehm, M., André, E.: Catch Me if You Can: Exploring Lying Agents in Social Settings. In: *AAMAS*. Academic Press Inc (2005)
10. Traum, D., Gratch, J., Marsella S., Lee, J., Hartholt, A.: Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal Virtual Agents. In: *8th International Conference on Intelligent Virtual Agents*. Tokyo, Japan (2008)