Evaluating Affective Feedback of the 3D Agent Max in a Competitive Cards Game[*]

Christian Becker[1], Helmut Prendinger[2], Mitsuru Ishizuka[3], Ipke Wachsmuth[1]

[1] Faculty of Technology, University of Bielefeld, 33549 Bielefeld, Germany
[2] National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan
[3] Graduate School of Information Science and Technology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Abstract. Within the field of Embodied Conversational Agents (ECAs), the simulation of emotions has been suggested as a means to enhance the believability of ECAs and also to effectively contribute to the goal of more intuitive human–computer interfaces. Although various emotion models have been proposed, results demonstrating the appropriateness of displaying particular emotions within ECA applications are scarce or even inconsistent. Worse, questionnaire methods often seem insufficient to evaluate the impact of emotions expressed by ECAs on users. Therefore we propose to analyze non-conscious physiological feedback (bio-signals) of users within a clearly arranged dynamic interaction scenario where various emotional reactions are likely to be evoked. In addition to its diagnostic purpose, physiological user information is also analyzed online to trigger empathic reactions of the ECA during game play, thus increasing the level of social engagement. To evaluate the appropriateness of different types of affective and empathic feedback, we implemented a cards game called *Skip-Bo*, where the user plays against an expressive 3D humanoid agent called *Max*, which was designed at the University of Bielefeld [6] and is based on the emotion simulation system of [2]. Work performed at the University of Tokyo and NII provided a real-time system for empathic (agent) feedback that allows one to derive user emotions from skin conductance and electromyography [13]. The findings of our study indicate that within a competitive gaming scenario, the absence of negative agent emotions is conceived as stress-inducing and irritating, and that the integration of empathic feedback supports the acceptance of Max as a co-equal humanoid opponent.

1 Introduction and Motivation

Embodied Conversational Agents are computer-generated, humanoid characters that are able to conduct a natural face-to-face dialogue with human users [12], whereby the types of communication channels range from purely textual input to multi-modal speech–gesture interfaces. Since most autonomous ECAs are designed to assist human users [6, 7], it is often assumed that the implementation of

---

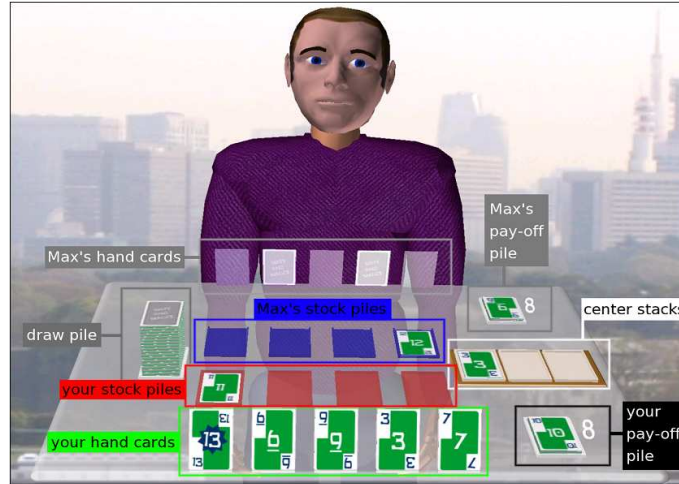[*] This paper builds on the joint project outline described in [3].

Fig. 1. The card game *Skip-Bo* as an interaction scenario for an Empathic Max.

negative emotions is unnecessary or even contra-productive. In those scenarios, the ECA typically plays the role of some subordinate "butler" who is expected to be always kind and friendly while serving the needs of the human user. In the context of affective gaming, however, the simulation and expression of negative emotions might increase the believability and naturalness of the synthetic opponent dramatically as most games are designed to evoke various kinds of emotions (positive and negative) to intensify the game experience.

Moreover, research in affective computing is offering promising results on interpreting human physiological information as emotions [11]. In accord with the two-dimensional (arousal, valence) model of [8], we claim that all emotions can be characterized in terms of judged valence (positive or negative) and arousal. Since skin conductance increases with a person's level of overall arousal and stress, and electromyography correlates with negatively valenced emotions, named emotions can be identified in the arousal–valence space. A real-time system based on [8] is described in [13]. In our work, the behavior of the ECA is modulated by both its own and the human interlocutor's emotional state. Consequently, the agent may be experienced as a more sensible and sociable interaction partner.

As our ECA, we use a 3D agent called *Max* that has been developed by the Artificial Intelligence Group at the University of Bielefeld [6] (see Fig. 1). Max has basic abilities for multi-modal interaction such as synchronized auditory speech, and facial and bodily gestures, and is controlled by a cognitively motivated architecture that enables him to conduct deliberative as well as reactive behavior. Max uses a concurrent emotion simulation system based on dimensional emotion theories [2]. The gaming scenario described in this paper provides the emotion simulation system with a clearly defined goal (to win the game), and it may thus derive a power relationship between the human player and itself in any given (game) state. This information enables the agent to dis-

tinguish between the emotion categories "fear" (low dominance) and "anger" (high dominance), and adapt the behavior of Max accordingly.

The rest of the paper is organized as follows. Section 2 describes the setup of our empirical study and in Section 3, we present our results based on questionnaires and the analysis of the physiological data recorded during game play. Section 4 concludes the paper.

2 Empirical Study

As an affective gaming scenario, the classical cards game *Skip-Bo* has been implemented as a face-to-face interaction scenario between a human player and the Max agent (see Fig. 1). In the game, players have the conflictive goal of getting rid of the eight cards on the pay-off piles to the right side of the table by playing them to the shared white center stacks. As on these center stacks the order of cards from one to twelve is relevant, the hand and stock cards must be used strategically by the players to achieve this overall goal, and win the game.

The 'physical' objects necessary to play the game were modelled as 3D objects and enriched by semantic information, so that intuitive point-and-click interaction by the human player as well as natural gestural interaction by Max (e.g. moving cards on the table) were easily realized [9].

Max displays different types of facial emotions within the pleasure-arousal-dominance space that reflect his current emotional state [2]. Speech was not seen as important in the cards game setting and has therefore not been implemented. However, when in a negative affective state, Max utters a variety of grunts and moans. Moreover, he continuously simulated breathing and eye-blinking, giving the user the impression of interacting with a life-like agent.

Visual and auditory feedback was also given whenever the human player was selecting or moving cards. Moreover, the Max agent gave visual feedback to the user by dynamically looking at the objects (cards) selected by himself or the user for a short period of time, and then looking straight ahead again in the direction of the user. Max also performs a simple type of turn–taking by nodding whenever he completed his move. These behavior are intended to increase the user's perception of interacting with an agent that is aware of its environment and the actual state of the game.

2.1 Design

In order to assess the effect of simulated emotions and empathic feedback in the context of human–computer interaction we designed the following four conditions within the proposed gaming scenario (see [4] for a similar set of conditions):

1. *Non-Emotional* condition: Max does not display emotional behavior.
2. *Self-Centered Emotional* condition: Max appraises his own game play only, and displays e.g. (facial) happiness when he is able to move cards.

3. *Negative Empathic* condition: Max shows self-centered emotional behavior and responds to those user actions that thwart his own goal of winning the game. Consequently, he will e.g. display distress when the user performs a good move or is detected to be positively aroused. This condition implements a form of 'negative' empathy.
4. *Positive Empathic* condition: Here, Max is also self-centered emotional, but user actions are appraised 'positively' so that he is "happy for" the user's game progress. If the user can be assumed to be distressed, Max will display "sorriness" for the user. This condition implements 'positive' empathy.

Note that Max follows a competitive playing strategy in all conditions.

2.2 Subjects

The study included 14 male and 18 female subjects. All but one subject were Japanese, and two of them had never played a card game before. The age of the subjects ranged from 22 to 55 years and the average age was 30 years. Subjects were given a monetary reward of 500 Yen for participation. They were told in advance that they would get an extra reward if they won against Max. Subjects were randomly assigned to the four conditions (eight in each condition).

2.3 Procedure

Subjects received written instructions of the card game (in Japanese) with a screenshot of the starting condition before they entered the room with the experimental setup. Subjects entered the room individually and were seated in front of an 50 inch plasma display with attached loudspeakers on both sides (see Fig. 2). They were briefed about the experiment, in particular that they would play a competitive game. Then, subjects could play a short introductory game against a non-

Fig. 2. Experimental setup.

emotional Max, which allowed them to get used to the mouse based point-and-click interface, and also provided subjects the possibility to ask clarifying questions about the game. Every subject won this first game easily.

Next, the biometrical sensors of the ProComp Infinity encoder [15] were attached to the subject and the subject was assured that these sensors were not harmful. Upon consent, a skin conductance (SC) sensor was attached to the
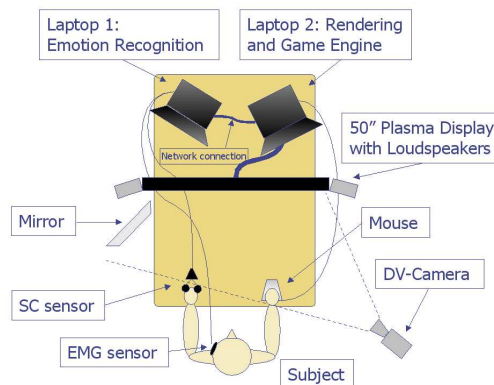
index finger and the small finger of the non-dominant hand. The electromyography (EMG) sensor was attached to the subject's left (mirror-oriented) cheek to measure the activity of the masseter muscle. Then a relaxation phase of three minutes started, with Max leaving the display and the subject being advised not to speak. This phase was necessary to obtain a baseline for the normalization of the bio-signals, since values may greatly vary depending on subject.

From now on, the experimenter remained visually separated from the subject (behind the screen) only to supervise the experiment. After the baseline was set, Max re-entered to the screen and the subject was asked to start the game. After the game was completed, the subjects were asked to fill in an questionnaire in English presented on the screen, together with a Japanese translation on hard-copy. The questionnaire contained 25 questions that were related to the participant's subjective experience while playing the game.

The whole interaction was recorded with a digital video camera positioned to the right behind the subject. In order to capture both the interaction on the screen as well as the human player's facial expression, a mirror was set up to acquire in indirect image of the human players face. Each game lasted for about ten minutes. A protocol of the development of the game, the acquired physiological data, and the video data were recorded for later analysis.

3 Results

Both questionnaires and biometrical data were evaluated to estimate the impact of different forms of emotional agent behavior (or their absence) on human users. Our findings will be presented in the following sections.

3.1 Questionnaire Results

The questionnaire contained twenty-five questions, which can be grouped into the following categories: (i) *Overall Appraisal*: Seven questions about the experimental condition, including questions about whether subjects liked playing the game or how they felt during game play; (ii) *Affective Qualities of Max*: Twelve questions related to the emotionality, personality, and empathic capability of Max; (iii) *Life-Likeness of Max*: Six questions about user judgements of the human-likeness of Max' behavior and and outward appearance.

Questions were rated on a 7 point Likert scale. Due to space limitations, only selected questions will be discussed. (The full set of questions and results can be obtained from the authors.) With respect to the first group of questions (Overall Appraisal), all but two subjects liked to play the game and everyone wanted to play it again. A nearly significant effect of the two empathic conditions in comparison with the Non-Emotional and Self-Centered Emotional conditions could be found. Subjects in the empathic conditions tended to feel less lonely ($t(30) = 1.66$; $p = 0.053$).[4]

---

[4] The level of statistical significance is set to 0.05.

The second group of questions (Affective Qualities of Max) – while not providing results of statistical significance – showed that subjects had a tendency to perceive Max as hiding his 'true feelings' in the Non-Emotional and Self-Centered Emotional conditions and showing his true feelings in both empathic conditions ($t(30) = -1.49$; $p = 0.073$). Also, Max was experienced as more caring about the human players' feelings when playing a positive empathic manner then when playing in a negative empathic manner ($t(14) = -1.6$; $p = 0.068$).

Concerning the third group of questions (Life-Likeness of Max), the agent was more perceived as "a human being" when playing in an empathic way, opposed to playing in a non-emotional or self-centered emotional way ($t(30) = -3.42$; $p = 0.001$). Moreover, Max' outward appearance was judged as more attractive when reacting empathically as compared to the Non-Emotional and Self-Centered Emotional conditions ($t(30) = -2.2$; $p = 0.018$).

3.2 Results of Biometrical Data Analysis

In order to analyze the recorded physiological data (skin conductance and electromyography), we focused on game situations where emotional reactions in the human player as well as Max were likely to occur. Specifically, we assumed emotional reactions whenever either of the players was able to play at least two pay-off pile cards in a row, which are moves towards winning the game, and found eighty-seven such situations.

Determining the exact duration of emotions is a notoriously hard problem [10]. We chose to analyze periods of ten seconds, consisting of five seconds before the last card was played, and the succeeding five seconds. For those periods and each of the four experimental conditions, the arithmetic means (averages) were calculated for both normalized skin conductance and normalized electromyography values (see Fig. 3). For each data set (each subject and each signal type), normalization was performed by first subtracting the average baseline value from the current signal value and dividing the resulting value by the range of values applicable to each subject (maximum minus minimum).

Regarding skin conductance, we found a significant difference between the Negative Empathic condition and the Positive Empathic condition ($t(48) = -3.48$; $p = 0.0006$), as well as between the Non-Emotional condition and the Negative Empathic condition ($t(44) = 1.81$; $p = 0.04$). Moreover, the Self-Centered Emotional and Positive Empathic conditions were statistically significantly different ($t(38) = -3.1$; $p = 0.002$).

As high skin conductance is an indicator of stress [5], the human player was seemingly most stressed in the Positive Empathic condition where Max was mostly "happy for" for the human player's success and giving positive empathic feedback by smiling back whatever the emotional or game state of the user was. Although counter-intuitive at first sight, it is important to note that in the setting of a competitive game, positive empathic behavior is quite unnatural and may thus induce user stress. The relatively high stress level in the Non-Emotional condition further supports our argumentation that inappropriate behavior (relative to an interaction task) leads to higher stress levels. These
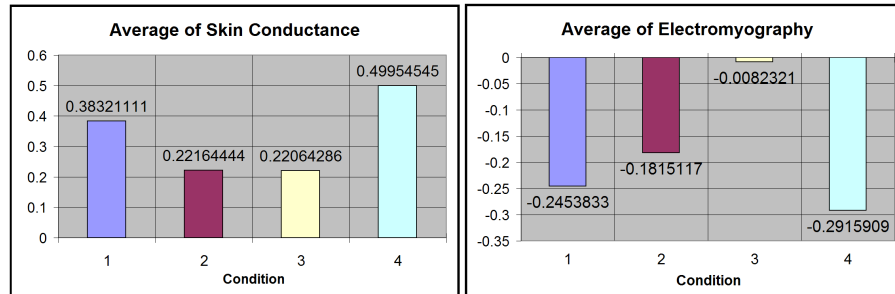
Fig. 3. The average values of normalized skin conductance and electromyography data within dedicated segments of the interaction. Explanation: 1 (Non-Emotional), 2 (Self-Centered Emotional), 3 (Negative Empathic), 4 (Positive Empathic).

results are consistent with the corresponding questionnaire item, asking whether Max' behavior is seen as irritating. Here, Max was perceived as most irritating in Non-Emotional condition, followed by the Positive Empathic condition.

Regarding EMG, only the Negative Empathic condition differs significantly from all other conditions, whereby all p-values are smaller than 0.012. (Observe that in Fig. 3 all values are below zero, meaning that the baseline period was not experienced relaxing in terms of muscle tension.) High values of electromyography are primarily an indicator of negative valence [1]. The highest value was achieved in the Negative Empathic condition, where Max was designed to evoke negative emotions in the human player by showing negative emotions, e.g. a sarcastic smile to the user's (assumed) frustration. Notably, the lowest EMG values can be observed in the Positive Empathic condition where Max performed a "calm down" gesture whenever the human player was assumed to be frustrated or angry. The same tendency can be found in the questionnaire as the subjects were judging Max as most caring in the Positive Empathic condition and the least caring in the Negative Empathic condition.

4 Conclusions

To our knowledge, this is the first study that systematically investigates the impact of different types of emotional behavior on human physiology. Emotions were displayed by the highly sophisticated 3D agent Max that was integrated into a realistic game setting. While previous similar studies only considered positive emphatic response [13, 14], the current experiment also analyzes the utility of displaying negative emotions. Our chief finding is that – within a competitive game scenario – the absence of negative empathy is conceived as stressful (derived from SC) and irritating, as it might also be experienced when playing against a human player. A complementary result is that negatively emphatic behavior induces negatively valenced emotions (derived from EMG) in the user. While emotion simulation systems such as [2] or emotion recognition systems such as [13] cannot be validated directly, the use of physiological information seems to be a promising approach to evaluate their effects on human users.

Acknowledgements

References

1. J. L. Andreassi. *Psychophysiology. Human Behavior & Physiological Response.* Lawrence Erlbaum Associates, Mahwah, NJ, 4 edition, 2000.
2. C. Becker, S. Kopp, and I. Wachsmuth. Simulating the emotion dynamics of a multimodal conversational agent. In *Proceedings Tutorial and Research Workshop on Affective Dialogue Systems (ADS-04)*, LNAI 3068, pages 154–165, Berlin Heidelberg, 2004. Springer.
3. C. Becker, H. Prendinger, M. Ishizuka, and I. Wachsmuth. Empathy for Max (Preliminary project report). In *The 2005 International Conference on Active Media Technology (AMT-05)*, pages 541–545. IEEE, 2005.
4. S. Brave, C. Nass, and K. Hutchinson. Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62(2):161–178, 2005.
5. J. Healey and R. Picard. Detecting stress during real-world driving tasks,.
6. S. Kopp, B. Jung, N. Lessmann, and I. Wachsmuth. Max – a multimodal assistant in virtual reality construction. *KI Zeitschift (German Magazine of Artificial Intelligence), Special Issue on Embodied Conversational Agents*, 2003.
7. S. Kopp, P. Tepper, and J. Cassell. Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the International Conference on Multimodal Interfaces (ICMI-04)*, 2004.
8. P. J. Lang. The emotion probe: Studies of motivation and attention. *American Psychologist*, 50(5):372–385, 1995.
9. M. E. Latoschik, P. Biermann, and I. Wachsmuth. Knowledge in the loop: Semantics representation for multimodal simulative environments. In *Proceedings of the 5th International Symposium on Smart Graphics*, 2005.
10. R. W. Levenson. Autonomic specifity and emotion. In R. J. Davidson, K. R. Scherer, and H. H. Goldsmith, editors, *Handbook of Affective Sciences*, pages 212–224. Oxford University Press, Oxford, 2003.
11. R. W. Picard. *Affective Computing.* The MIT Press, Cambridge, MA, 1997.
12. H. Prendinger and M. Ishizuka, editors. *Life-Like Characters. Tools, Affective Functions, and Applications.* Cognitive Technologies. Springer Verlag, Berlin Heidelberg, 2004.
13. H. Prendinger and M. Ishizuka. The Empathic Companion: A character-based interface that addresses users' affective states. *International Journal of Applied Artificial Intelligence*, 19(3):267–285, 2005.
14. H. Prendinger, J. Mori, and M. Ishizuka. Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International Journal of Human-Computer Studies*, 62(2):231–245, 2005.
15. Thought Technology Ltd., 2002. URL: `http://www.thoughttechnology.com`.