Developing Positive Attitudes Towards Cooperative Problem Solving by Linking Socio-emotional and Cognitive Intentions

Masato Kuno
Kyoto University
Kyoto, Japan
Email: kuno@ii.ist.i.kyoto-u.ac.jp

Yoshimasa Ohmoto
Shizuoka University
Shizuoka, Japan
Email: ohmoto-y@inf.shizuoka.ac.jp

Toyoaki Nishida
Fukuchiyama University
Fukuchiyama, Japan
Email: toyoaki.nishida@fukuchiyama.ac.jp

***Abstract*—We focus on problem-solving situations in which people cooperate with virtual interactive agents. The final goal is to achieve high-quality problem solving through a problem-solving process in which people recognise agents as effective collaborators and actively cooperate with the agents. It is known that trust, which is the basis of cooperation, has both affective and competent aspects. However, because the impression of agents tends to focus on the ability side, it is necessary to make the people recognise that the ability and emotional sides of agents are not separate but are integrated. In the proposed method, we apply the Alternate Estimation by representing Global and Local (AEGL) goal-oriented behaviour model, which adjusts the behaviour of an agent that show the agent's ability side and the emotional side by estimating the causes of the human behaviour. We demonstrated how both behaviours change consistently through interaction with people. In this study, we designed an experimental agent model to realise the proposed model. The people and agents are asked to perform cooperative decision-making tasks by exchanging opinions and adapting to each other's behaviour. The results suggest that the relationship between the ability and affective functioning of the agents eases tension and that people feel more comfortable talking to the agents.**

Keywords–human-agent interaction; cooperative problem.

I. INTRODUCTION

Today, agents are developed to cooperate for solving problems. When solving complex problems with no optimal solution, if people think alone, their view is narrow, and they will not be able to generate a solution. If people speak with other people and exchange opinions, their opinions stimulate them to produce various ideas. However, not everyone can be a good collaborator.

Kwon et al. argued that a situation in which self-disclosure and intimacy with collaborators are triggered is a precondition for a sense of community and that a sense of community is the final product of successful collaborative learning [1]. If people feel uncomfortable interacting with a partner who is unfriendly and difficult to talk to, they are not likely to interact willingly.

On the other hand, if they find a partner who is likeable and easy to talk to, they are expected to cooperate with him/her without resistance and to consider his/her opinion to be valid. They will be able to compare opinions efficiently and have a broader perspective regarding problem solving. As a result, good quality satisfactory problem solving can be achieved.

However, such a partner does not always exist in the scene of problem solving. Therefore, an agent who can be called on as a partner at any time and who is easy to talk to and likeable is expected to be available as a collaborator.

The final goal of this study is to realise good, satisfactory problem solving through a process in which people recognise agents as effective partners and actively cooperate with them. However, it is difficult to induce people to actively cooperate even in situations in which people must cooperate with other people.

Chi et al. categorised student engagement behaviours into four patterns: interactive, constructive, active, and passive [2]. This classification suggests that a gap exists between the passive and interactive states, in which students produce knowledge by talking with others. Therefore, to induce people's willingness to cooperate, it is necessary for people to feel that it is easy to talk to agents and to be friendly with them to overcome the gap. This study focuses on this aspect.

In collaborative learning where learners discuss and exchange opinions with each other to solve problems, social interaction is important for the learners' positive attitude towards cooperation [3]. The following two types of social interactions are typical during collaborative learning [4] [5]:

- Cognitive interaction: Discussions related to the task itself or the metacognition of the collaborators; and
- Socio-emotional interaction: Shared emotions about the task and pronounced expressions of positive and negative emotions.

First, cognitive interaction with other learners implies an active exchange of ideas for the solution of the problem. Smooth cognitive interactions can stimulate discussion of the problem and enhance people's evaluations of the agents' abilities. A previous study has shown that a group of learners who share their thoughts and understanding through cognitive interaction engage in a deeper level of the learning process than a group of learners who do not actively share their thoughts and understanding [6]. A study by Maltz et al. also demonstrated that people continue to accept system suggestions when they trust the system's capabilities [7]. Thus, if the partner makes appropriate and accurate statements about the task, the learner trusts his/her partner's ability, and the learner's positive attitude towards cooperation is expected to be induced.

Second, socio-emotional interaction is related to the expression of emotions in a social context. In other words, the interaction aims to 'build trust and belonging by getting to know each other'. Kwon et al. argued that socio-emotional interaction has the effect of smoothing out the behaviour of the members and protecting them from friction [1]. Kreijns et al. further argued that socio-emotional interaction facilitates overall interaction and increases the efficiency of cooperative learning [8]. In conclusion, socio-emotional interaction may increase the familiarity with other learners and induce smooth and low-resistance interaction. In addition, socio-emotional interaction is thought to have the effect of increasing the learners' positive attitudes by synergistically inducing total interaction including cognitive interaction.

From the above discussion, cognitive and socio-emotional interactions have the effect of giving people positive impressions about the competence and familiarity of the other learner, respectively. In this study, we induce agents to perform cognitive and socio-emotional behaviours to that people can recognise agents as collaborators and interact with them without resistance.

In the case of people, it is obvious that they have emotions in addition to their abilities. However, the fact that agents have emotions (including intention) is not obvious, since because their abilities are often emphasised due to their strong associations with machines. In this regard, Dennett proposed an idea called 'intentional stance', which refers to the idea that robots and agents have intentions when people interact with them [9]. Dennett stated that people do not usually think that robots and agents have intentions. A comparison of interactions between people and agents that induce intentional stance and those that do not induce intentional stance demonstrates that people who interact with the former interact more actively, even in situations unrelated to the task [10].

In conclusion, it is possible that the two functions of an agent are understood by people separately. If these two functions are not understood as a whole, people cannot perceive an agent's intention consistently. As a result, people cannot perceive the agent's abilities and emotions towards the inconsistent behaviour of the agent, and thus, they cannot induce positive attitudes towards the cooperation of the agent. We aim to demonstrate how agents behave with consistent intentions by inducing people to perceive that their cognitive and socio-emotional behaviours are related to each other. For this purpose, the cognitive and socio-emotional intentions of the agents are represented by the model, and both intentions are updated with consistency. Thus, we propose generating both cognitive and socio-emotional behaviours of agents using the AEGL model developed by Omoto et al. [11]. The AEGL model has the characteristic of combining the intentions of both people and agents, estimating their intentions alternately. Because cognitive and socio-emotional behaviours of people are consistent, it is expected that the cognitive and socio-emotional intentions of agents, which are updated based on the estimated intentions of people, are also consistent. In other words, it is expected that people can perceive how the two intentions are integrated and combined with people's intentions.

In this study, as a first step towards the realisation of the above-proposed method, the cognitive and socio-emotional behaviours of agents are generated in parallel. In the generation of socio-emotional behaviours, we suggest the relationship between the agents' socio-emotional behaviours and cognitive behaviours so that humans can recognise the relationship between them. Specifically, the AEGL model is not used, but a simple model that mimics the AEGL model is used to generate agents' behaviours. We then asked the agents and people to solve problems by cooperating and observed how the relationship between the agents' cognitive and socio-emotional behaviours affects people.

The goal of this study is to induce people to feel that agents are easy to talk with and are familiar to them. That is, we aimed to reduce the resistance that people feel to cooperating with an agent and to increase the subjective liking of the agent. These effects support people's willingness to cooperate with agents. When people feel comfortable and familiar with agents, they are more likely to speak with the agents and accept their opinions.

This paper is organised as follows. Section 2 introduces the related work. Section 3 provides an overview of the proposed method. Section 4 describes the experiments conducted in this study. Section 5 presents the results of the experiment. Section 6 describes the results of the experiment and future tasks. Section 7 concludes the work.

II. Related Work

Previous studies have shown that treating emotional interactions in addition to task-oriented interactions in dialogue systems and agents can increase people's satisfaction [12] and induce positive perceptions of interactions [13].

For example, Kumar et al. investigated the effect of having tutor agents support students in their studies while performing socio-emotional behaviours in addition to cognitive behaviours [14]. The tutor agents work on socio-emotional behaviours using interaction strategies based on the three categories of 'showing solidarity', 'showing tension release', and 'agreeing'. The rules for generating the behaviours are predefined, and the cognitive and socio-emotional behaviours of the agents were triggered using different interaction strategies with the input of task progress and interaction states. The tutor agents generated cognitive and socio-emotional behaviours separately, and the students did not perceive these behaviours to be linked to each other. Thus, although the questionnaire results indicated that the agents who performed the socio-emotional behaviour were friendlier than those who did not perform the socio-emotional behaviour, we did not find that the students had the impression that the agents were easy to talk to or friendly, and we did not obtain any indicators that revealed students' positive attitudes towards cooperation.

In contrast, in this study, agents used the same intention model to generate these behaviours, and the agents always behaved according to their intentions during the interaction. Furthermore, the parameters of one intention model were used as input to the other, and the goal was to make the participants perceive that the agent's cognitive and socio-emotional behaviours do not follow different interaction strategies, but that they are acting based on consistent intentions.

A study that considers the generation of behaviours based on the estimated intentions of people is the work by Zhou et al. [15]. Zhou et al. proposed a neural model that can detect emotions in people's speech and generate conversations by learning from a large set of conversational data. In the study by Zhou et al., emotions were detected by the trained model and the agent's optimal emotion was expressed based on the trained model. However, this study used the AEGL model to present the agents' goal orientation, assumed the same two-layered intention model for people and agents, and alternately updated the intention model based on the real-time cognitive and socio-emotional behaviours of the people during the interaction. Then, people can observe how the agents

 ISBN: 978-1-61208-761-0

change their behaviour based on the intentions and emotions of the people during the interaction and can predict the agents' intentions. By making people strongly aware of the cognitive and socio-emotional intentions of agents, people can expect consistent intentions.

III. PROPOSED MODEL

The goal of this study is to induce the effect of the ease of talking and familiarity with an agent to support people's willingness to cooperate with the agent. For this purpose, we attempted to induce people to recognise the relationship between agents' abilities and emotional functions. The proposed model differs from previous studies in that both cognitive and socio-emotional behaviours are simply output in parallel, and both types of behaviour are output using the AEGL model developed by Omoto et al. [11]. The AEGL model, in which the intentions of people and agents are alternately estimated and combined, is used to output the agent's cognitive and socio-emotional behaviours in parallel so that people can perceive how the two intentions are combined with the people's intentions. In the following, we describe the details of the proposed model.

A. AEGL Model for Cognitive and Socio-Emotional Behaviour Generation

First, an overview of the AEGL model developed by Omoto et al. is presented in Figure 1. In the AEGL model, the intention of the people is inferred from their verbal and non-verbal behaviours. However, various intentions are inferred from the observed behaviour of the people. In the AEGL model, the relationship between behaviour and intention is represented by two different levels of concreteness: global purpose and local objective. That is, the two-layered relational intention model, with global purpose and local objective, is used to infer human intentions. The local objective is the categorisation of the actual observed behaviours and represents a temporal objective. For example, the observed behaviours, such as 'laugh' and 'eye contact', can be categorised into the category 'synchronize'. Therefore, 'laugh' and 'eye contact' are related to the local objective called 'synchronise'. However, the global purpose expresses a longer-term purpose. For example, temporary objectives, such as 'synchronise' and 'show attention', are thought to lead to the long-term purpose of 'showing acceptance'. Therefore, 'synchronise' and 'show attention' are related to the global purpose called 'showing acceptance'. In this way, the task-specific global purpose and local objective are represented as nodes, and each node has its own parameters.

Omoto et al. assumed the above intention models for both agents and people and updated the parameters of both intention models alternately with people's behaviours as input. Figure 1 presents an overview of the update, and Figure 2 reveals the details of the update.

1) First, the agent outputs the behaviour based on the parameters of the local objective.
2) Next, the agent observes the behaviour of the people, updates the parameters of the local objective in the people's intention model, and then updates the parameters of the global purpose of the people.
3) The parameters of the people's global purpose and the agent's global purpose are merged, and the agent updates the parameters of the local objective based on the parameters of the global purpose.
4) The agent then outputs its next behaviour based on the parameters of its updated local objective.

Omoto et al. stated that, by making people observe an agent's trial and error behaviour, people can infer the unobservable internal state of an agent [10]. By doing so, the agent's intentions are inferred by people, and their intentional stance is induced. Therefore, by outputting the cognitive and socio-emotional behaviours of agents using the AEGL model, it is possible to induce people to estimate the process by which agents produce their task behaviour and emotional expressions. Moreover, it can make people recognise that the agent has intentions regarding competence and emotional aspects.

In this study, we apply the feature of the AEGL model of inducing people to estimate the internal state of the agent and attempt to induce people to estimate the association between the agent's cognitive and socio-emotional intentions. An overview of the proposed model is illustrated in Figure 2. In the proposed model, cognitive and socio-emotional intentions are inferred from people's behaviour, and the next cognitive and socio-emotional behaviours of agents are determined in parallel. Both actions are output as a single behaviour of an agent. Because task-related behaviours and emotional behaviours of people are consistent, the cognitive and socio-emotional intentions of the agents, which are updated based on the estimated people's intentions, are also considered consistent (dotted arrows in Figure 2).

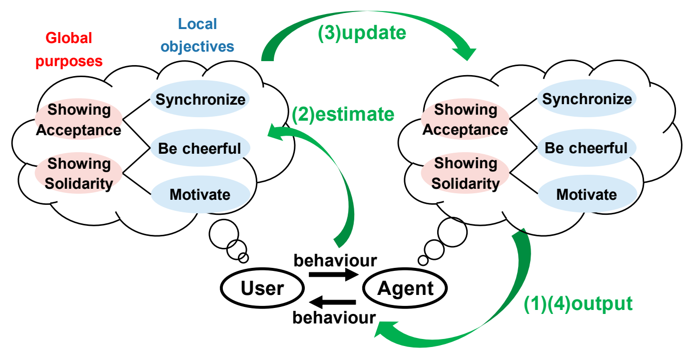

Figure 1. AEGL model.

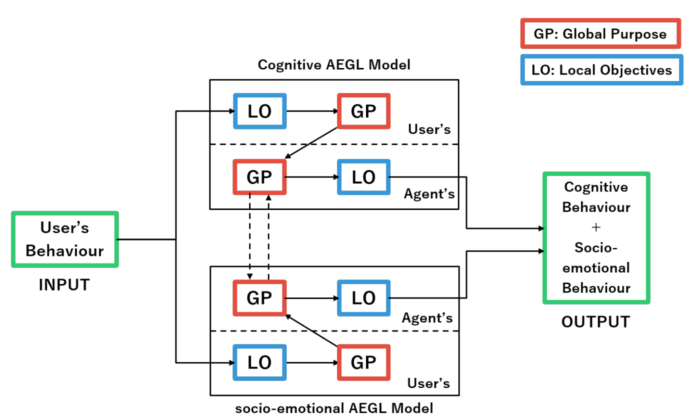

Figure 2. Agent model in this study.

 ISBN: 978-1-61208-761-0

B. Experimental Socio-emotional Behaviour Generation Model

The goal of this study is to facilitate people to perceive the relationship between agents' cognitive and socio-emotional behaviours so that they feel that it is easy to talk to and be familiar with the agents. In this study, we designed an experimental agent model that mimics the proposed model, and we investigated the effects of the relationship between agents' cognitive and socio-emotional behaviours on people. Specifically, the agent's cognitive behaviour, to which the AEGL model has already been applied in Ohmoto's study, is generated using the rule-based model, which is described later for simplicity. The socio-emotional behaviours of agents to which the AEGL model has not been applied in the previous studies are generated based on a simple behavioural model that mimics the AEGL model, which is also described later. Such a simple experimental agent model can also present the structure of the interaction in which cognitive and socio-emotional behaviours are performed in parallel during the interaction. Thus, we can induce people to feel that the agents perform each behaviour based on consistent intention, and people can recognise the relationship between each behaviour. The cognitive behaviours of the agents are generated from verbal and non-verbal behaviours of people based on predetermined rules. Cognitive behaviours, such as 'proposal', 'dividing labour', and so on, are triggered by the rule base on the observed task-related behaviours of the people.

The socio-emotional behaviours of the agents are generated using the behaviour generation model depicted in Figure 3. This model differs from the AEGL model in two respects. One aspect is that only the agent's intention model is assumed, and the other is that the local objective layer does not exist. In this model, the parameters related to the emotional state of people are set, and the connection between the parameters and their verbal and non-verbal behaviours is predetermined. Thus, based on the verbal and non-verbal behaviour of people during the interaction, parameters related to people's emotional state, such as 'nervous' and 'favourability', are updated. The socio-emotional intentions, such as 'showing tension release' and 'showing acceptance', are selected according to the values of these parameters, and the specific behaviours, such as 'praise' and 'acknowledge', are output for each intention. The five basic behaviours to express intentions other than 'seeing users attitude' are based on Kwon's classification of socio-emotional interactions [1].

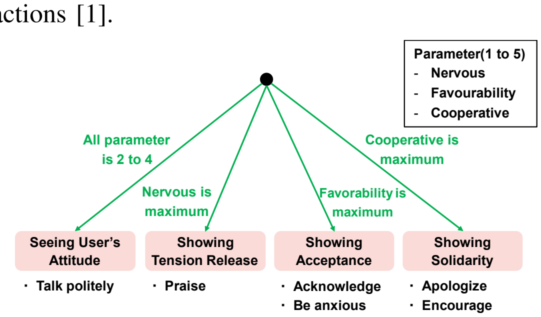

Figure 3. Socio-emotional behaviour generation model in this experiment.

C. Role of the Agent

In the proposed model, the following two agents are used to develop a positive attitude towards cooperation. The cooperative agent takes the same position as the people and interacts with them to engage in collaborative problem solving. This agent generates behaviours using the AEGL model. The teacher agent knows the details of the task that neither the people nor the cooperative agent knows and offers knowledge in response to their questions. This agent is not directly involved in solving the problem but leads the way in ensuring the task goes smoothly. In this study, we used two agents based on the work by Ohmoto et al. [16]. One of the reasons was to reduce the psychological resistance to interaction by having people observe interactions between agents. By doing so, we aim to induce people to learn how to interact with agents and reduce their psychological resistance to the interaction. The other reason is that this study is based on a cooperative learning situation in which learners in the same position work together to solve problems. Therefore, we aim to promote an equal discussion between the people and cooperative agents and to smooth the progress of the task by having a separate agent as a teacher who maintains the knowledge of the task. In the next section, we provide an overview of the evaluation experiments using this experimental model.

IV. EXPERIMENT

For the realisation of the proposed model, we designed the experimental model and conducted experiments in which people and agents were asked to perform cooperative problem solving. The purpose of this experiment is to generate both cognitive and socio-emotional behaviours of the agents in parallel and to observe the effects of the link between these behaviours on people. For this purpose, we adopted experimental tasks that require cooperation and assistance between people and agents. The task is performed in such a way that they help each other and show socio-emotional behaviours, such as gratitude and apology. In this way, we can evaluate the degree of familiarity and ease of talking of the agents based on their behaviour and physiological indices.

A. Task

1) Task Overview: For the task, we use a tower defence game. The player and agent communicate with each other to place a tower in position to prevent an enemy attack. The game was developed using Unity, and the player can move the character in the virtual world using a controller. The player interacts with the agent in the virtual world by speaking. Interactions involve three parties: the player, cooperative agent, and teacher agent. In the experimental group, the socio-emotional behaviour of the cooperative agent throughout the task is generated by the experimental model in Section 3-B. In the control group, the cooperative agent does not perform any socio-emotional behaviour. The experimental model of Section 3-B is not used for cognitive behaviour, but both groups generate rule-based behaviours based on the goal of completing the game. The game overview is illustrated in Figure 4.

2) Rule: The player works with the cooperative agent to discuss and determine how to place the towers to allow the player to defend his/her position against the enemy. The placement of the towers is costly and must be within the cost limitations. Players need to discuss and consider the placement of towers that can efficiently defeat enemies, considering trade-offs, such as the tower attack power versus cost. Players can

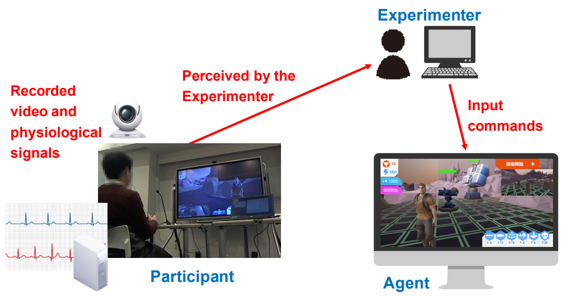

Figure 4. Experimental setup.

also strengthen and repair towers. The player and cooperative agent can communicate with the teacher agent and ask questions about the effects of the tower and types of enemies. The player and two agents talk to each other using voice chat.

To succeed in this task, the player and cooperative agent must work well together. For example, differences exist in the ability to strengthen and repair towers, and the players and agent must discuss and choose an action depending on the situation. When socio-emotional behaviours are displayed, such as thanking the partner for his/her help, praising the partner for his/her skill, and apologising for the failure of the tower, we believe that the player will become more familiar with the agent and more willing to interact and cooperate more actively.

B. Wizard-of-Oz

The experimental model used in this study was realised using the Wizard-of-Oz (WoZ) method. The experimenter observed the verbal and non-verbal behaviours of the players during the task and manually increased or decreased the parameters of the players associated with the behaviours based on predefined rules. For example, the ‘cooperative’ parameter increased if the player acted to encourage the cooperative agent, and the ‘showing acceptance’ parameter decreased if the player rejected the suggestion of the cooperative agent. The experimenter manually selected the behaviour of the cooperative agent from among the behaviours expressing socio-emotional intentions according to the highest parameters. The physical behaviours of the agents were performed manually by the experimenter using a controller, and the real-time speech of the experimenters was changed to be perceived as the agents’ speech using a voice changer. The reason for not using the recorded audio is to respond immediately to changes in the player’s behaviour.

C. Participants

In this experiment, 15 students who were not involved in information engineering were subjects. Participants ranged in age from 19 to 32 years old with an average age of 22.53 (variance of 10.65), including 12 males and three females.

D. Experimental Setup

Eight participants were set up as the experimental group, and seven participants were in the control group. An overview of the experimental setup is illustrated in Figure 4. Participants interacted with the agents through a monitor and controlled their avatars in the game using an Xbox controller. The perceptions of the participants’ actions and statements and the manipulations of the agents were assessed by the experimenters based on the above model for each group of agents. The video images of the participants’ upper body during the experiment were captured using a web camera, and the time series of the heart rate and Skin Conductance Response (SCR) were obtained using a Polymate biometric analyser.

After the game, a questionnaire was conducted to obtain the participants’ subjective evaluations. The physiological indices obtained in this study were missing the heart rate data and SCR data for two participants, and for another three participants, the SCR data exhibited little response, so we concluded that they were not appropriate for use in the analysis. Therefore, considering the small number of participants, we included participants whose data were correctly obtained for each analysis.

The analyses in Sections 5-C, 5-D, and 5-E were performed on eight subjects in the experimental group and seven subjects in the control group. The analysis in Section 5-A was performed on six subjects in the experimental group and seven subjects in the control group. The analysis in Section 5-B was performed on five subjects in the experimental group and five subjects in the control group.

V. Results

A. Cardiac Sympathetic Index and Cardiac Vagal Index

To estimate the internal state of the participants during the task, the Cardiac Sympathetic Index (CSI) and Cardiac Vagal Index (CVI) were calculated from the participants’ heart rate data. The CSI and CVI are indices designed by Toichi et al. [17]. The long-axis component L and the short-axis component T were calculated from the distribution of the heart rate intervals in the Lorenz plot analysis, where T/L is CSI, and $log(L\times T)$ is CVI. These are indicators that can detect the heightened sympathetic and parasympathetic nerves. Hayashi et al. found that the stress state is higher when the sympathetic nervous system is high, and the relaxed state is higher when the parasympathetic nervous system is high [18]. In this analysis, we evaluated the participants’ stress state in terms of their ease of talking with the agents and their resistance to cooperation.

We hypothesised that the change in the internal state of the participants would be more pronounced after the speech of the cooperative agent. Thus, we calculated the CSI and CVI of the participants for 30 seconds after normal speech (excluding socio-emotional speech) by the cooperative agent and analysed them separately as statistical data. The agent’s socio-emotional speech was ‘acknowledged’, ‘apologies’, ‘be anxious’, ‘encourage’, and ‘praise’. First, the average CSI for 30 seconds after normal speech for the entire task was 1.58 for the experimental group and 1.78 for the control group, with Welch’s t-test showing a significant difference between the groups at p = 0.0001 (t = -4.38, p = 1.29*e-05). The average CVI is 5.33 for the experimental group and 5.12 for the control group, with Welch’s t-test showing a significant difference between the groups at p = 0.001 (t = 3.67, p = 0.00025; Figure 5).

Therefore, to capture the temporal variation of CSI and CVI values, we calculated the average CSI and CVI 30 seconds after the cooperative agent’s speech during the first and last 5 minutes of the task for each participant (Figures 6 and 7). A two-way analysis of variance (ANOVA) for the averages of CSI indicates no significant differences between groups or over time. We also analysed the CVI and found a significant interaction at the significance level of $p < 0.05$ ($F = 7.34$, $p = 0.02$), and an effect of temporal variation exists in the control group at a significance level of $p < 0.05$ ($F = 5.98$, $p = 0.033$).

The above analysis suggests that participants' internal states were particularly affected after the speech of the cooperative agent. The experimental group exhibited lower post-speech CSI values and higher CVI values throughout the task (i.e., the participants were at a relatively low level of tension and excitement and were relaxed). In the control group, the CVI value after the speech decreased and approached a tense state as the task progressed, whereas no decrease was found in the experimental group. This suggests that the socio-emotional speech of the cooperative agent suppresses the participants' tension and enhances their relaxation, removing some of their psychological resistance to cooperating with the cooperative agent, whereas under normal circumstances, the participants' tension and excitement levels increase as the task progresses.

Finally, to investigate whether the socio-emotional speech of the cooperative agent directly affects the internal state of the participants, we calculated the CSI and CVI values for 30 seconds after the socio-emotional and normal speech of the cooperative agents in the experimental group and compared them. The average CSI is 1.71 after the socio-emotional speech and 1.64 after the normal speech, and the paired t-test was performed with no significant difference found ($t = 0.25$, $p = 0.81$). The average CVI is 5.60 after socio-emotional speech and 5.34 after normal speech. A similar test was performed, and no significant difference was found ($t = 0.80$, $p = 0.44$).

The results suggest that, although socio-emotional speech does not affect participants immediately, the accumulation of socio-emotional speech may change the influence of the agent's normal speech. In other words, by performing cognitive and socio-emotional behaviours in parallel, the socio-emotional behaviour supports the cognitive behaviour in suppressing the participants' tension.

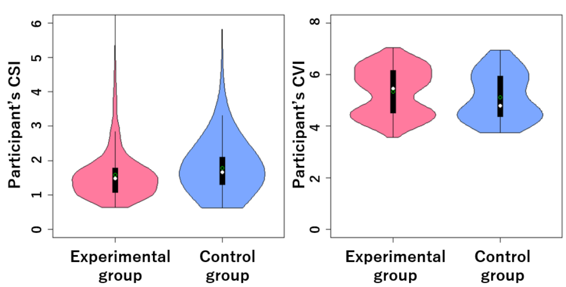

Figure 5. CSI, CVI average for 30 seconds after cooperative agent's speech for the entire task.

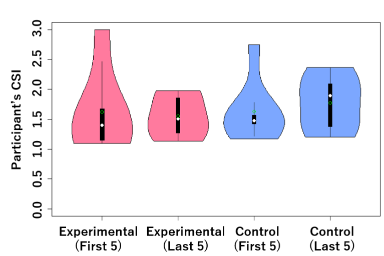

Figure 6. CSI average over time for 30 seconds after the cooperative agent's speech in the first and last five minutes in each group.

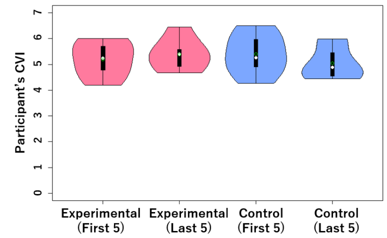

Figure 7. CVI average over time for 30 seconds after cooperative agent's speech in the first and last five minutes in each group.

B. Skin Conductance Response

To estimate the internal state of the participants during the task, we measured their SCR. The SCR is an electrical measure of sweating caused by mental tension and excitement, and it is expected that SCR can be used to estimate mental stress and emotion. Lin et al. showed that higher subjective stress results in a higher SCR value [19]. In this study, we focused on the effect of the cooperative agent's speech on the stress state of the participants by examining the change in the SCR value 30 seconds after the cooperative agent's speech during the task. In terms of the participants' stress state, we assessed the degree of resistance to cooperation and the ease of talking with the agents.

There is a delay in the reaction of the SCR and a delay in returning to the original value after a reaction. Therefore, it is reasonable to examine how the values vary concerning a certain threshold in the SCR analysis. For each participant, we calculated the average of the SCR for 30 seconds after each speech of the cooperative agent and set the average of the bottom 20% of values in the speech as the baseline for that participant. Then, we analysed the variation in SCR values 30 seconds after the speech based on a baseline +0.5 threshold.

First, we analysed the SCR response rate of the cooperative agent's normal speech, considering the speech to be responsive if the SCR value exceeded the threshold within 30 seconds after the speech. Thus, we found 330 responsive and 91 nonresponsive speech results in the experimental group, and 364 and 77 in the control group, respectively. We performed the chi-square ($\chi 2$) test to compare the response rates between groups, finding no significant differences in the response rates ($\chi 2 = 2.11$, $p = 0.15$). To investigate the response of the SCR in each group, we focused on the speech with an SCR. We first calculated the number of seconds that the SCR value exceeded the threshold within 30 seconds after speech and then compared the averages between groups (Figure 8). The average of the experimental group is 7.72, and the average of the control group is 8.48, which is the result of Welch's t-test. The significance level is $p < 0.05$, with the control group having a significantly longer time ($t = -2.01$, $p = 0.045$).

This result indicates that the control group tends to have a longer reaction time than the experimental group. This suggests that participants in the control group tend to pay too much attention to the cooperative agent's speech and become tense. However, the participants in the experimental group were not too tense and were able to relax in response to the agent's speech.

C. Participants' Speech

We measured the participants' socio-emotional speech and compared the speech with that of the experimental and control

Copyright (c) IARIA, 2020. ISBN: 978-1-61208-761-0

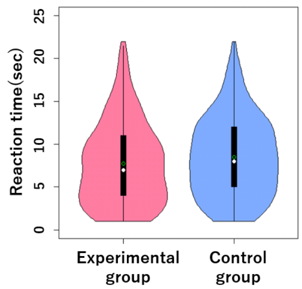

Figure 8. Results of analyses on SCR.

groups. Based on the categorisation of socio-emotional interactions by Kwon et al., we measured five types of speech: 'acknowledge', 'apologies', 'be anxious', 'encourage', and 'praise' [1]. The annotations were done manually by the experimenter after observing the video of the experiment. The chi-square ($\chi 2$) test was used in the evaluation statistics.

We measured the number of socio-emotional and other normal speech in the task for each group and compared the proportion of socio-emotional speech between groups. The results are listed in Table I. The results reveal that the number of instances of socio-emotional is high in the experimental group at a significance level of $p < 0.001$ ($\chi 2$ = 12.05, p = 0.00052).

Next, participants' cognitive speech other than socio-emotional speech was measured for 5 minutes after the start of the task, 5 minutes before the end of the task, and in the middle of the other tasks, and the number of instances of speech between the experimental and control groups was compared. The two-way ANOVA was used as a statistical measure.

The results of that average are presented in Table II. The results indicate no significant difference between the groups (F = 3.85, p = 0.072) and no significant interaction (F = 3.00, p = 0.067), but the cognitive speech of the experimental group increased in the middle part of the task. The mean mid-task time (s) was 1376.875 for the experimental group and 1321 for the control group, and Welch's t-test found no significant difference (t = 3.17, p = 0.10).

These results suggest that the familiarity and ease of talking to the cooperative agent felt by the participants were exhibited in the participants' behaviour in the form of increased socio-emotional speech. Furthermore, the recognition of the link between the cognitive and socio-emotional speech of the cooperative agent may have led to an increase in the participants' cognitive speech.

TABLE I. NUMBER OF SPEECHES OF PARTICIPANTS IN THE TASK

	Socio-emotional	Normal
Experimental group	40	538
Control group	7	381

TABLE II. NUMBER OF COGNITIVE SPEECHES OF PARTICIPANTS IN THE TASK

	First 5	Middle	Last 5
Experimental group	14.25	70.50	16.00
Control group	13.14	46.00	14.71

D. Participants' Speech Latency

The latency between the end of the cooperative agent's speech and the start of the participant's speech was measured, and the speech latency was compared between the experimental and control groups. We assessed the participants' familiarity and ease of talking with the cooperative agent by observing whether participants respond to the cooperative agent's speech in a fast-paced way.

A two-way ANOVA was used to evaluate the participants' speech latency. To capture the temporal changes in speech latency, we focused on the first and last 5 minutes of the task. The latency averages were calculated for each participant in each situation. The results are displayed in Figure 9. The results of the analysis reveal that the average of the experimental group is 1.51 seconds for the first 5 minutes and 0.95 seconds for the last 5 minutes. The average of the control group is 1.73 seconds for the first 5 minutes and 1.38 seconds for the last 5 minutes. A two-way ANOVA was applied, corresponding to each participant. A main effect between groups was found at a significance level of $p < 0.05$ (F = 4.68, p = 0.0498), and the experimental group exhibited a shorter speech latency than the control group. The main effect of the temporal change was found at a significance level of $p < 0.01$ (F = 12.27, p = 0.0039).

As the task progressed, the speech latency became shorter in both groups, but it was particularly short in the experimental group. We assumed that the long speech latency indicates that the participants found it challenging to communicate with the cooperative agent and that they did not consider the suggestions and opinions of the cooperative agent to be valid. The short speech latency indicates that they were actively attempting to communicate and cooperate with the cooperative agent.

In both groups, as the task progressed, the speech latency decreased because participants felt that the cooperative agent was easier to communicate with and more effective as a partner; thus, the speech latency decreased. However, in the experimental group, the latency of speech was shorter because participants felt more familiar with the agent and felt it was easier to talk to the cooperative agent due to the link between the cooperative agent's cognitive and socio-emotional behaviours.

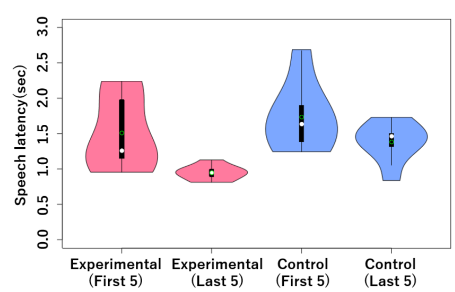

Figure 9. Speech latency in the first and last five minutes in each group.

E. Subjective Evaluation by Participants

A questionnaire with a seven-point Likert scale was conducted after the experiment to investigate the participants' subjective evaluations. Participants were asked to evaluate all statements on a scale of 1 (not true) to 7 (true). We assessed the answers to the following 12 statements about the cooperative agent.

- Q1: I took a liking to the agent.
- Q2: The agent was reliable.
- Q3: I felt easy to talk with the agent.
- Q4: The behavior of the agent was natural.
- Q5: I found the agent's behaviour human-like.
- Q6: I felt the value of the cooperation with the agent.
- Q7: I was willing to the cooperation with the agent.
- Q8: I could understand the way of thinking of the agent.
- Q9: The agent understands my way of thinking.
- Q10: I felt accepted by the agent.
- Q11: I felt relieved by the agent.
- Q12: I felt solidarity with the agent.

The answers to each statement were analysed, and the results of some of the statements are summarised in Figure 10. In the following section, we describe the content of each statement and the results of the answers.

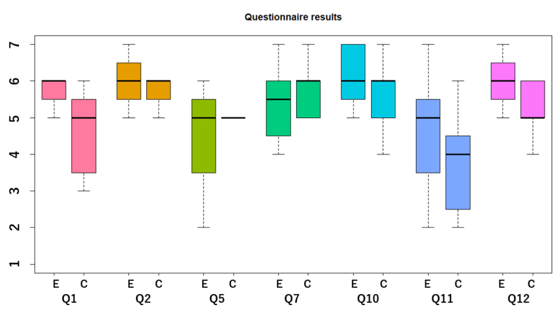

Figure 10. Questionnaire results (left: experimental group, right: control group).

- Q1: I took a liking to the cooperative agent.
  We assessed the participants' subjective liking for the cooperative agent. We performed the Mann-Whitney U test on the answer results and found a significance level of $p < 0.05$, resulting in high favourability in the experimental group.
- Q2: The cooperative agent was reliable and Q7: I was willing to the cooperation with the cooperative agent.
  We assessed the subjective trust in the cooperative agent and the participants' positive attitude. There was no significant difference in the answer results between the groups.
- Q5: I found the cooperative agent's behavior human-like.
  We used a voice changer for the cooperative agent's speech; however, if the participants were aware of WoZ, their perception of the cooperative agent's humanity could have been greatly enhanced. The actual answers obtained in both groups are close to the median, which suggests that the participants were not aware of the use of the WoZ technique using the voice changer.
- Q10: I felt accepted by the cooperative agent. /Q11: I felt relieved by the cooperative agent. /Q12: I felt solidarity with the cooperative agent.
  These statements assessed the participants' subjective perceptions of each socio-emotional intention expressed by the cooperative agent. As a result, although no significant difference was found between groups, the experimental group exceeded the average of the control group.

VI. Discussion

A. *Effect of Link of Socio-emotional and Cognitive Behaviour*

Sections 5-A and 5-B demonstrate how the link between the cognitive and socio-emotional behaviours of the cooperative agent affect the internal state of the participants. The results reveal that the participants tend to be less tense and more relaxed in response to the cooperative agents' speech. This tendency supports the participants' willingness to co-operate with the cooperative agent and that the participants' psychological resistance to the cooperation with the agent was reduced. Furthermore, the results in Section 5-A suggest that the socio-emotional behaviour of the cooperative agents did not directly affect the participants but that cognitive behaviour, based on socio-emotional behaviour, was effective in reducing the participants' tension. This suggests that it is important to perform both behaviours in parallel.

The above psychological changes may have led to the changes in the participants' behaviour observed in Sections 5-C and 5-D. The increase in the participants' socio-emotional speech is thought to be due to a decrease in the participants' psychological resistance to sharing and expressing their emotions as a result of feeling more familiar with agents and because it was easier to talk to the cooperative agent. Furthermore, the participants' cognitive speech also increased in the middle part of the task, which suggests that familiarity with the cooperative agent may have affected the participants' willingness to cooperate. The decrease in the latency of the participants' speech may be because they became less stressed when interacting with the cooperative agent and thus responded more quickly.

The results of the questionnaire demonstrated that the participants in the experimental group were more favourable towards the cooperative agents. As a result of the participants' recognition of the link between the cognitive and socio-emotional behaviours of the cooperative agent and because they felt less tension and burden when interacting with the cooperative agent, their subjective favourability towards the cooperative agent was likely to increase. However, this study failed to develop more cooperative attitudes, such as trust and active cooperation, among participants towards the cooperative agent.

B. *Constructing an Ideal Proposal Model*

In this study, the analyses of the CSI, CVI, and SCR indicated that the participants in the experimental group tended to be more relaxed. According to the questionnaire results, the participants in the experimental group more strongly perceived the socio-emotional intentions of showing tension release, acceptance, and solidarity expressed by the cooperative agent. Thus, the socio-emotional intentions adopted in the experimental model used in this study were relatively well

 ISBN: 978-1-61208-761-0

conveyed to the participants. Considering the purpose of the proposed model, which is to induce participants to feel the intentionality of the cooperative agent using the AEGL model, the types of adopted intentions were appropriate. However, room for improvement still exists in terms of the failure to develop the participants' positive attitudes, and it is necessary to continue to investigate optimal intentions.

In the experimental model, we aimed to generate the cognitive and socio-emotional behaviours of the cooperative agent in parallel, so that the participants perceive the link between the behaviours and experience familiarity and the ease of talking with the cooperative agent. However, we could not induce the participants' trust in the cooperative agent and strong positive attitudes towards it. Therefore, there is room to devise more effective ways to link the cognitive and socio-emotional behaviours of the cooperative agents. In this study, we did not implement the link in the model. However, it is possible to induce participants to feel a stronger consistency of the agent's behaviour by implementing the model using the intention parameter state to update the other intention parameters.

VII. Conclusion and Future Work

The final goal of this study was to induce people's positive attitudes towards cooperation with agents in cooperative problem-solving situations. To achieve the final goal, we aimed to induce people to feel that agents are easy to talk with and that they are familiar with the agents, which supports their positive attitude towards cooperation. Therefore, we proposed an agent model in which cognitive and socio-emotional behaviours are output in parallel using the AEGL model. The AEGL model was used to demonstrate that the agents behave based on consistent intentions and to induce people to recognise the integrity of their behaviours and to induce positive attitudes. As a first step towards the realisation of the proposed model, we designed an experimental agent model to simulate the link between cognitive and socio-emotional behaviours and aimed to make the agents feel easy to talk with and friendly to people. For the task, we used a tower defence game.

As a result, we found a change in the psychological state of people who became less nervous about the agents' speech and a change in the behaviour of people who were presumed to feel more comfortable talking to the agents. In addition, based on the socio-emotional speech of the agents, the positive effects of cognitive speech on people were increased.

The next task is to create a method to link the cognitive and socio-emotional behaviours of agents based on the findings of this experiment. Further consistency between cognitive and socio-emotional intentions in the model would further allow people to recognise the agent as a single entity with intentions and induce people's willingness to cooperate.

Acknowledgment

This work was supported by JST CREST Grant Number JPMJCR17A5, Japan.

References

[1] K. Kwon, Y.-H. Liu, and L. P. Johnson, "Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators," Computers & Education, vol. 78, 2014, pp. 185–200.

[2] M. T. Chi and R. Wylie, "The ICAP framework: Linking cognitive engagement to active learning outcomes," Educational Psychologist, vol. 49, no. 4, 2014, pp. 219–243.

[3] S. Järvelä, H. Järvenoja, J. Malmberg, J. Isohätälä, and M. Sobocinski, "How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement?" Learning and Instruction, vol. 43, 2016, pp. 39–51.

[4] P. Dillenbourg, M. Baker, A. Blaye, and C. OMalley, "The evolution of research on collaborative learning," In P. Reimann and H. Spada, Learning in humans and machines. Towards an interdisciplinary learning science, 1995, pp. 189–211.

[5] R. Toni and L.-G. Lisa, "Socially shared regulation in collaborative groups: An analysis of the interplay between quality of social regulation and group processes," Cognition and Instruction, vol. 29, no. 4, 2011, pp. 375–415.

[6] P. Näykki, J. Isohätälä, S. Järvelä, J. Pöysä-Tarhonen, and P. Häkkinen, "Facilitating socio-cognitive and socio-emotional monitoring in collaborative learning with a regulation macro script–an exploratory study," International Journal of Computer-Supported Collaborative Learning, vol. 12, no. 3, 2017, pp. 251–279.

[7] M. Maltz and J. Meyer, "Cue utilization in a visually demanding task," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 44, 2000, pp. 283–283.

[8] K. Kreijns, P. A. Kirschner, W. Jochems, and H. Van Buuren, "Determining sociability, social space, and social presence in (a) synchronous collaborative groups," Cyber Psychology & Behavior, vol. 7, no. 2, 2004, pp. 155–172.

[9] D. C. Dennett, "The intentional stance," MIT Press, 1989.

[10] Y. Ohmoto, T. Suyama, and T. Nishida, "A method to alternate the estimation of global purposes and local objectives to induce and maintain the intentional stance," in Proceedings of the Fourth International Conference on Human Agent Interaction, 2016, pp. 379–385.

[11] Y. Ohmoto, T. Suyama, and T. Nishida, "Extended method to alternate the estimation of global purposes and local objectives in multiple human-agent interaction," The Eleventh International Conference on Advances in Computer-Human Interactions, 2018, pp. 212–217.

[12] 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, vol. 62, no. 2, 2005, pp. 231–245.

[13] H. Prendinger and M. Ishizuka, "The empathic companion: A character-based interface that addresses users'affective states," Applied Artificial Intelligence, vol. 19, no. 3-4, 2005, pp. 267–285.

[14] R. Kumar, H. Ai, J. L. Beuth, and C. P. Rosé, "Socially capable conversational tutors can be effective in collaborative learning situations," in International Conference on Intelligent Tutoring Systems. Springer, 2010, pp. 156–164.

[15] H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu, "Emotional chatting machine: Emotional conversation generation with internal and external memory," in Thirty-Second AAAI Conference on Artificial Intelligence, 2018, pp. 730–738.

[16] Y. Ohmoto, J. Karasaki, and T. Nishida, "Inducing and maintaining the intentional stance by showing interactions between multiple agents," in Proceedings of the 18th International Conference on Intelligent Virtual Agents, 2018, pp. 203–210.

[17] M. Toichi, T. Sugiura, T. Murai, and A. Sengoku, "A new method of assessing cardiac autonomic function and its comparison with spectral analysis and coefficient of variation of r–r interval," Journal of the Autonomic Nervous System, vol. 62, no. 1-2, 1997, pp. 79–84.

[18] R. Hayashi and S. Kato, "Proposal of design policy of therapy robots system based on relaxation state," in 2017 IEEE 6th Global Conference on Consumer Electronics (GCCE). IEEE, 2017, pp. 1–3.

[19] T. Lin, M. Omata, W. Hu, and A. Imamiya, "Do physiological data relate to traditional usability indexes?" in Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future, 2005, pp. 1–10.

 ISBN: 978-1-61208-761-0