The Effect of Alternating Propagation of Local Objective and Global Purpose by a Network-Connected Two-Layer Model of Emphasizing Factors

Yoshimasa Ohmoto(ohmoto@i.kyoto-u.ac.jp)
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 Japan

Asami Matsumoto (matsumoto@ii.ist.i.kyoto-u.ac.jp)
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 Japan

Toyoaki Nishida (nishida@i.kyoto-u.ac.jp)
Department of Intelligence Science and Technology
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 Japan

Abstract

When performing a collaborative task that includes different, but related, subordinate tasks, we must consider the consistency and coordination between objectives of the subordinate tasks (local objectives) and the purposes of the collaboration (global purposes). In this study, we propose a method for alternatingly propagating local objectives and global purposes using a two-layer model. We investigated the effects of the consistency and coordination on human stress levels during a collaborative task. We conducted an experiment that compared the results of the collaborative task using two different types of agents: the proposed method and a previously reported method. Questionnaires confirm that the proposed agent significantly improved impressions of consistency and naturalness. The results of our electrocardiogram analysis confirm that the participants' stress levels increased throughout the task when they interacted with the traditional agent.

Keywords: verbal and nonverbal behavior; human-agent interaction; decision-making support.

Introduction

Embodied conversational agents (ECAs) have recently been integrated into real-life situations. Examples of ECAs include robots in the real world and virtual agents in augmented reality environments. However, these agents are normally regarded as automated machines to perform tasks or human-centered interfaces to naturally seek information (e.g., (Shneiderman & Maes, 1997)). We intend to develop an ECA that is regarded as a social partner and supports humans in their actual life, rather than one that is just a multimodal user interface. Such ECAs are not specialized for a particular task, but are designed to respond to broad collaborative tasks.

There are various types of collaborative tasks. To facilitate tasks that include different but related subordinate tasks, it is important to consider the consistency and coordination between the objectives of the subordinate tasks (local objectives) and the purposes of the entire task (global purposes). We call this type of task a "project style task". Project style tasks in our daily life include rearranging a room, organizing a party, or setting up a social event. In a project style task, each of the collaborative members has a vague final image of the task, and they perform a subordinate task in response to this image. At the same time, the member modifies the final image according to the status of the subordinate task.

In this study, we propose a method that alternately propagates local objectives and global purposes during a collaboration through interactions. Using an agent implemented by the proposed method, we investigated the influence of consistency and coordination on a user's stress levels and impressions, when performing a project style task with the agent.

Previous studies (Ohmoto, Miyake, & Nishida, 2012; Ohmoto, Kataoka, & Nishida, 2014) supported interactive decision-making using a method for dynamically estimated emphasizing points (DEEP) based on verbal reactions, body movements, and physiological indices. This method was applied to interactive decision-making. The emphasizing points are factors that influence decision-making. Those studies found that the dynamic estimation of emphasizing points was helpful for the participants in the decision-making process.

In this study, we extended this method to a project style task. We separately described the emphasizing points of local objectives and global purposes, and have defined the network connections between them. Through the network connection, each local objective and global purpose alternately propagates the results of the emphasizing point estimates.

The paper is organized as follows. In Section 2, we briefly introduce some related work. Section 3 contains an outline of the proposed method for alternately propagating local objectives and global purposes. Section 4 contains a description of our experiment that compared experimental and control groups, and presents our results. In Section 5, we discuss the achievements of this research and some future work. We give our conclusion in Section 6.

Related Works

The agents which collaboratively perform various tasks have been proposed in many researches, such as subordinate support agents when people perform tasks on their own initiative, automated attentive agents which automatically perform tasks in line with human's wishes.

The typical way to speculate human's intentions is human-agent communication through a dialogue. Kitamura et al. (Kitamura et al., 2008), for example, developed a system that matches users' queries with search targets by communicating with users throughout the interview. Most of the researches considered that people had reliable demands and needs and they tried to search them. However, especially in collaborative tasks, people's demands and needs are ambiguous and they are interactively changed through the tasks.

Tohyama et al. (Tohyama, Omori, Oka, & Morikawa, 2003) proposed a model of a learning agent that changes its own actions through observations and predictions of other agents' actions in a cooperative task. The learning agent acquired several models of others and selected the appropriate one. The agent changed its behavior by using selected model when the other agent changed strategy. They considered that the communication partners changed their intentions, but the consistency and coordination between them have to be managed by a main person who took the initiative because the cooperative learning agent subordinately followed others.

There are some methods and concepts to be mutually directable when performing tasks (Allen, Guinn, & Horvtz, 1999; Klein, Woods, Bradshaw, Hoffman, & Feltovich, 2004). Mixed-initiative, for example, refers to a flexible interaction strategy, where each agent can contribute to the task what it does best. Furthermore, in the most general cases, the agents' roles are not determined in advance, but opportunistically negotiated between them as the problem is being solved. In many cases, they just provide a division of roles among interaction members, so the consistency and coordination for performing the task are still managed by a main person.

It is known that assigned goals when performing the task affect personal goals and task performance (Meyer & Gellatly, 1988). In the project style task, there are two different goals; global purposes and local objectives. The global purposes are related to final product images. The local objectives is important for the feasibility. Both of them are interlinked and changed through the interaction in the task. Therefore, to facilitate project style tasks, it is important to consider the consistency and coordination between local objectives and global purposes based on their changes. In this study, we extended the method to speculate participant's intentions for the project style task, and propose a method to keep the consistency and coordination through human-agent interaction in the task.

Alternating Propagation of Local Objective and Global Purpose

When performing a project style task, we must consider the consistency and coordination between local objectives and global purposes. To achieve consistency and coordination, we propose a method that alternately propagates the estimates of local objectives or global purposes into another human-agent interaction. In this section, we outline the process of propagating the estimates and describe how to proceed with the interaction.

Outline of Alternating Propagation

When we perform a project style task, we achieve a large collaboration using a combination of different, but related, smaller tasks. The project style task can be divided into two main phases: the local and global phases. In a local phase, we perform subordinate tasks and mainly consider how to achieve each task's local objective. In a global phase, we integrate the results of the sub-tasks to comply with the global purposes. We propose a method that propagates the influence of the local objectives into the global purposes, and vice versa, using a relational network that is designed in advance.

We cannot directly observe other people's intentions. Therefore, researchers have investigated a method that dynamically estimates "task specific" emphasizing points (DEEP) based on verbal reactions, body movements, and physiological indices. Emphasizing points are useful clues to a person's intentions. In this study, we extended this method to a network-connected two-layer model of emphasizing factors.

The network is composed of two layers: a layer in which local objectives are speculated from emphasizing points related to each subordinate task (local layer), and the layer in which global purposes are speculated from emphasizing points related to the whole collaborative task (global layer). The parameters of the nodes of the emphasizing points in each layer are updated in a correspondent phase based on the DEEP method. After that, the parameters of the local objective nodes or global purpose nodes (referred to as "intention nodes") are speculated based on the parameters of the network-connected emphasizing point nodes. Until this step, the speculating procedure is the same as in previous studies. That is, the system determines a plan of action based on the speculated local objectives or global purposes.

In this study, we connect the local objective nodes to emphasizing point nodes in the global layer, and the global purpose nodes to emphasizing point nodes in local layer. In a correspondent phase, we update the parameters of the intention nodes. This method determines a plan of action before the system is implemented, and the value of the nodes are propagated to the network-connected emphasizing point nodes in another layer. In another layer, the parameters of the intention nodes are updated based on the propagation. Their values are also propagated into the original layer, which updates the parameters of the intention nodes in the original layer. In this way, the value of each intention node is influenced by the values of nodes in another layer. We expect that this is useful when considering the consistency and coordination between local objectives and global purposes.

Outline of the interaction process using the proposed method

Figure 1 shows the interaction process in a collaborative task using the proposed method. The top shows the updating pro-

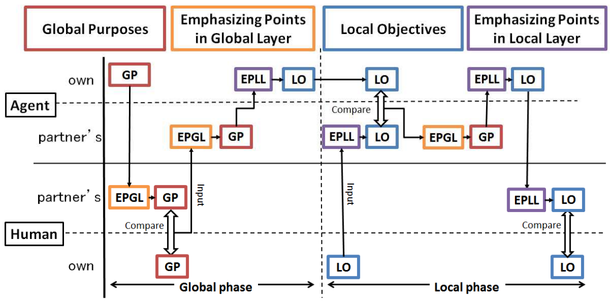

Figure 1: Overview of the interaction process using the proposed method

cess, and the bottom shows the speculated thinking process. This is an example. The left side shows a global phase, and the right shows a local phase. We briefly explain the interaction process below. In the global phase, when the interaction agent provides a target image of a final product, a person compares the provided image with their own. If there are some differences, they convey the differences to the agent using verbal and nonverbal information. The agent estimates their truthful structure of emphasizing points in the global layer using this communication, and speculates their global purposes. In the next step, the agent develops an action plan for achieving the speculated global purposes, and compares it with previous actual actions in the subordinate tasks. After that, the agent updates the structure of the emphasizing points in the local layer and the local objectives. In response to the result, the agent refines the estimated structure of the emphasizing points in the global layer and the global purposes.

In the local phase, when a person provides instructions to modify the agent's behavior when performing the task, the agent estimates and updates their emphasizing points in the local layer using the instructions, and speculates their local objectives. The agent compares the speculated local objectives with their own. If there are some differences, the agent estimates the influence of the differences on the emphasizing points in the global layer, and propagates the differences into the layer. After that, the agent speculates the global purposes and produces the next action plan for achieving the updated global purposes. The agent estimates the truthful structure of the emphasizing points in local layer based on the action plan, and updates their local objectives.

By repeatedly conducting the processes in a project style task, the agent can dynamically estimate the person's intentions and achieve consistency and coordination between local objectives and global purposes.

Evaluation Experiment

In this experiment, a participant performed a project style collaboration task with an interactive agent, to investigate how the proposed method affects a user's impressions of the agent and cognitive loads. We used two types of agents to evaluate the effect of the proposed method: a propagation agent that mutually estimated local objectives and global purposes using a network-connected two-layer estimation model of emphasizing factors, and an isolated agent that separately estimated the objectives using two isolated estimation models. To evaluate these agents, we measured and analyzed the low-frequency/high-frequency (LF/HF) value (electrocardiograph measurement), which is a physiological index used to measure stress. Additionally, the participant completed a questionnaire about their impressions of the agent.

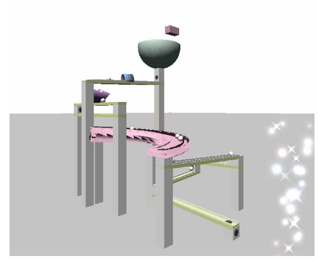

Figure 2: The filming target to record the behavior.

Task

In this study, participants were asked to make a video to record the complex behavior of a target. We call this a filming task. The participants did not know what was appropriate, so a filming agent supported each participant. The agent interactively provided suggestions for how to shoot the target, and collaboratively recorded a movie, considering the participant's preferences and emphasizing points.

Figure 2 shows the filming target. The target has some balls that move interestingly within the object. There are three impressive "gimmicks". Participants mainly considered approaches to filming the gimmicks. For example, dramatically, explanatorily, like a promotion video, a standard recording, and so on. There were ten recording scenes that were divided according to the object parts. The final film was produced using a combination of the ten scenes recorded by the participant and the support agent.

Procedure

At the beginning of the experiment, a participant observed four examples of movies in which a similar, but different, object was filmed by a professional cameramen. Next, the participant freely controlled the virtual video camera and observed the filming target. When the participant had become accustomed to the operation, they began the collaborative filming task, and we recorded the video and physiological indices. During the filming task, the participant repeatedly performed a cycle from planning the collaborative filming to producing a candidate for the final product. One filming cycle was composed of three phases: the planning, recording, and production phases. The planning and recording phases are local phases. The production phase was a global phase. The participant could communicate with the agent. The participant performed the cycle four times. At the conclusion of

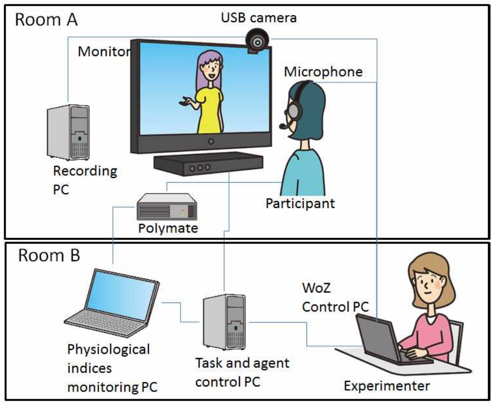

Figure 3: Experimental settings.

the experiment, the participant completed a questionnaire.

Experimental Settings

The experimental setting is shown in Figure 3. The participant sat in front of a 60-inch monitor that displayed the filming target and an interaction agent. The participant could freely control a virtual video camera and confirm the recorded video. Two video cameras recorded the participant's behavior: one was placed on the monitor to record the participant's behavior and another was placed behind the participant to record the task. Polymate was used to measure SCR and the electrocardiogram.

The experimenter was in another room and inputted commands into the system via a WOZ interface. The commands were "change the current task phase, confirm the product of the movie, or communicate with the agent", "input keywords extracted from participant's utterance", "provide a suggestion selected from previously prepared interaction scenarios based on the local objectives and global purposes", and "respond to unexpected questions related to the task". The experimenter provided suggestions and answers using a synthesized voice.

Participant

Twenty Japanese college students (14 men and 6 women) participated in the experiment. They were undergraduate students from 18 to 25 years old (an average of 21.7 years old). The all interacted with one of the agents for approximately 1 hour. Ten participants (7 men and 3 women) interacted with the propagation agent (the "propagation group") and the rest interacted with the isolated agent (the "isolated group"). The experimenter gave the following instructions about the agent: "The agent can recognize your speech. The agent has basic knowledge about the filming task. You can ask how to shoot the target and propose the shooting method".

Results of LF/HF analysis

We used the measured LF/HF value to detect whether or not the participant paid attention to the agent's suggestion in the DEEP system. However, some researchers have reported that

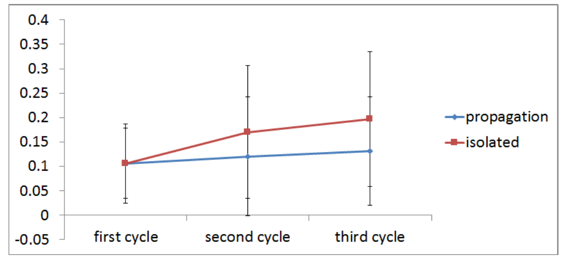

Figure 4: Stressful time rate in each filming cycle.

the LF/HF value was a clue for detecting stress (e.g., (Nater et al., 2006), (Nakazono, Hada, Ataka, Tanaka, & Nagashima, 2008)). Therefore, we analyzed the time series data of the LF/HF value during this experimental task. The LF/HF value exceed five during stressful tasks in a previous study (Nater et al., 2006), so we recorded the times when the LF/HF value was greater than five. We divided the accumulated time by the amount of time spent performing the task, and called this value the "stressful time rate". We calculated the stressful time rate for each filming cycle, each recording phase, each production phase, and each planning phase. The stressful time rate for the fourth cycle was relatively higher than for the other cycles, because the experimenter told participants that the next recording phase was the last at the beginning of the cycle. Therefore, we did not include the fourth cycle data in our analysis. The results are given below. This data does not follow a normal distribution, so we used nonparametric statistics in our analysis.

Figure 4 shows the stressful time rate of each participant in each filming cycle. It is clear from Figure 4 that the stressful time rate increased in the isolated group. We performed a Friedman's X2r-test on the data of each group to compare the stressful time rates in the cycles, which showed that there was a significant difference in the isolated group (p=0.0017) but there was no significant diffrence in the propagation group (p=0.67). We additionally performed Wilcoxon signed-rank tests with Bonferroni correction on the data of the isolated group for post-hoc test, which showed that the stressful time rate in the second cycle was marginally significantly more than in the first cycle (p=0.086) and that in the third cycle was significantly more than in the first cycle (p=0.025). This means that the participants in the isolated group became more stressed as the filming task progressed, but the participants in the propagation group did not.

Next, we analyzed the stressful time rate in each phase of the filming task.

We first considered the stressful time rate in the recording phase. Figure 5 shows our results. We performed a Friedman's X2r-test on the data of each group, but there was no significant diffrence. Since the averages of the data of the isolated group appear to be relatively higher than that of the propagetion group, we performed a Mann-Whitney U test on the data. But there is no significant diffrence (p=0.13).

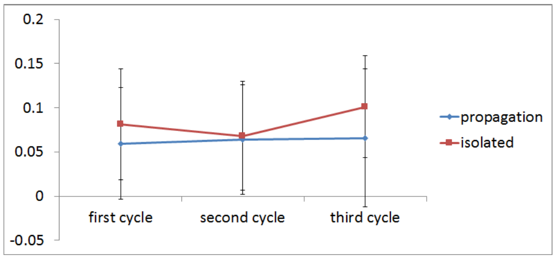

Figure 5: Stressful time rate in the recording phase of each filming cycle.

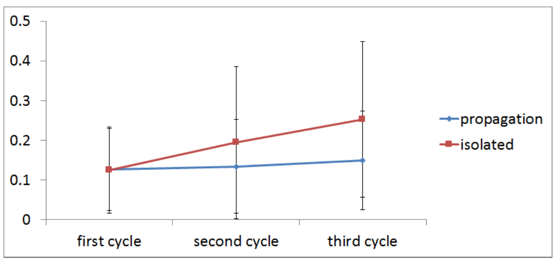

Figure 6: Stressful time rate in the production phase of each filming cycle.

We then analyzed the stressful time rate in the production phase. Figure 6 shows our results, which indicate that the stressful time rate increased in the isolated group. We performed a Friedman's X2r-test on the data of each group, which showed that there was a significant difference in the isolated group (p=0.00050) but there was no significant diffrence in the propagation group (p=0.74). We additionally performed Wilcoxon signed-rank tests with Bonferroni correction on the data of the isolated group, which showed that the stressful time rate in the first and second cycle was significantly less than in the third cycle (p=0.025, p=0.025).

Finally, we analyzed the stressful time rate in the planning phase. Figure 7 shows our results. We performed a Friedman's X2r-test on the data of each group, but there was no significant diffrence. However, the average of the second cycle appears to be relatively higher than that of the propagetion group. Then, we performed a Mann-Whitney U test on the whole data of the isolated group and the propagation group. As a result, the stressful time rate in the isolated group significantly more than in the propagation group (p=0.049).

Questionnaire Results

The purpose of this analysis was to investigate how the determining method of providing suggestions and action plans influenced the participants' subjective impressions. The participants answered five questions regarding the ECA's behavior using a seven-point scale. The scale was presented as seven ticks on a black line without numbers, which we scored from 1 to 7. The results are shown in Figure 8. We performed a

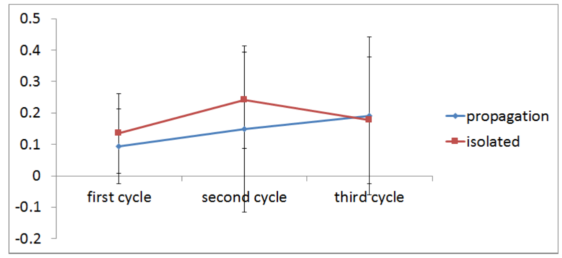

Figure 7: Stressful time rate in the planning phase of each filming cycle.

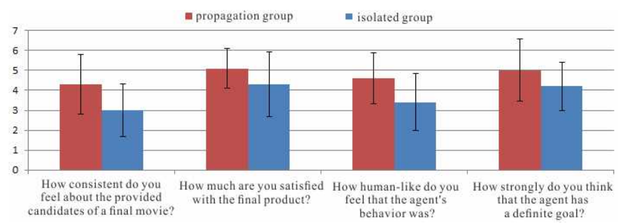

Figure 8: Questionnaire results.

Mann-Whitney U test on the data in the questionnaire.

- **How consistent do you feel about the provided candidates of a final movie?**
  The participants felt that the propagation agent was marginally significantly more consistent than the isolated agent (p=0.074). Additionally, some participants in the isolated group said that "the candidates were often very different from my target image" in interviews. This was not said by any participant in the propagation group. This suggests that the propagation agent provided relatively suitable suggestions and action plans, which considered both the local objectives and global purposes.

- **How strongly do you think that the agent has a definite goal?**
  The participants in the propagation group thought that the agent had a relatively more definite goal for the final product than the participants in the isolated group (5.0 vs. 4.2), but there was no significant difference (p=0.20).

These two results indicate that the participants in the propagation group thought that the final candidates provided by the agent were consistent, but the consistency was not induced by a definite goal for the final movie.

- **How satisfied are you with the final product?**
  The participants in the propagation group were relatively more satisfied with the final product than the participants in the isolated group (5.1 vs. 4.3), but there was no significant difference (p=0.12). The average scores were higher than the midpoint of the scale, so the participants accepted the final product provided by the agents. This means that the algorithms that controlled the human-agent interaction and provided candidates were accepted by the participants.

- **How human-like do you feel that the agent's behavior was?**
  The participants in the propagation group felt that the agent was marginally more human-like than the participants in the isolated group (p=0.074). We believe that one reason for this is that the propagation agent was consistent in terms of its suggestions and candidates. This means that the propagation agent provided candidates that considered the suggestions during the planning and recording phases. Consistency affects the participants' impressions of a human-like agent.

Discussion

From the results of our LF/HF analysis, the participants in the propagation group had a low level of stress throughout the task, but the participants in the isolated group became more stressed as the task progressed. In the isolated grou, the stress level increased gradually in the production phase, but, in other phases, the stress level did not increase. This means that the participants in the isolated group felt the stress because the product probably did not reflect their global purpose. Moreover, in the planning phase, the average stress level of the isolated group was more than that of the propagation group. This means that the participants in the isolated group could not plan based on their local objectives. Additionally, the questionnaires suggest that the participants felt that the propagation agent was consistent. This implies that the propagation agent provided consistent suggestions and candidates, and that this consistency reduced the participants' stress levels during the filming task.

We believe that the proposed method is useful for reducing stress levels during a project style task. The implemented agent was based on the DEEP system. The system requires active human-agent interaction to estimate the emphasizing points, so that it can improve the final product or the satisfaction of the decision-maker. However, the human-agent interaction in this experiment was not active. This is one reason why there were no significant differences between the propagation group and the isolated group. In the future, we intend to implement some mechanisms that induce active human-agent interaction during the task.

Conclusion

In this study, we proposed a method to alternately propagate estimates of local objectives and global purposes into the estimation of another human-agent interaction, using a network-connected two-layer model of emphasizing factors. We believe that this method can provide consistency and coordination between local objectives and global purposes. We extended DEEP to estimate the emphasizing factors, which was proposed in previous studies.

To evaluate our method, we implemented two types of agents: a propagation agent (that mutually estimates local objectives and global purposes using a network-connected two-layer estimation model of emphasizing factors), and an isolated agent (that separately estimates them using two isolated estimation models). We evaluated the proposed method using an experiment, in which a participant performed a project style collaboration task with an interactive agent. The results of our physiological index analysis suggested that participants who interacted with the propagation agent had a low level of stress throughout the task, but that the participants who interacted with the isolated agent became more stressed as the task progressed. Additionally, a questionnaire confirmed that the propagation agent increased the impression of consistency regarding the final product, and increased the impression of human-like behavior. These results suggest that the proposed method could be useful for reducing stress levels during project style tasks by increasing consistency.

References

Allen, J., Guinn, C. I., & Horvtz, E. (1999). Mixed-initiative interaction. *Intelligent Systems and their Applications, IEEE*, *14*(5), 14–23.

Kitamura, M., Shimohata, S., Sukehiro, T., Ikeno, A., Sakamoto, M., Orihara, I., & Murata, T. (2008). *Design and development of dialogue system for laddering search service* (Vol. 108; Tech. Rep.). IEICE technical report. Natural language understanding and models of communication.

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a" team player" in joint human-agent activity. *IEEE Intelligent Systems*, *19*(6), 91–95.

Meyer, J. P., & Gellatly, I. R. (1988). Perceived performance norm as a mediator in the effect of assigned goal on personal goal and task performance. *Journal of Applied Psychology*, *73*(3), 410.

Nakazono, K., Hada, T., Ataka, E., Tanaka, H., & Nagashima, Y. (2008). *Workload evaluation of gaming task by physiological indices and psychological indices* (Vol. 107 - 553; Tech. Rep.). Technical report of IEICE. HIP.

Nater, U. M., La Marca, R., Florin, L., Moses, A., Langhans, W., Koller, M. M., & Ehlert, U. (2006). Stress-induced changes in human salivary alpha-amylase activity?associations with adrenergic activity. *Psychoneuroendocrinology*, *31*(1), 49–58.

Ohmoto, Y., Kataoka, M., & Nishida, T. (2014). The effect of convergent interaction using subjective opinions in the decision-making process. In *Proc. the 36th annual conference of the cognitive science society* (pp. 2711–2716).

Ohmoto, Y., Miyake, T., & Nishida, T. (2012). Dynamic estimation of emphasizing points for user satisfaction evaluations. In *Proc. the 34th annual conference of the cognitive science society* (pp. 2115–2120).

Shneiderman, B., & Maes, P. (1997). Direct manipulation vs. interface agents. *interactions*, *4*(6), 42–61.

Tohyama, S., Omori, T., Oka, N., & Morikawa, K. (2003). Identification and learning of other's action strategies in cooperative task. In *Proceedings of 8-th international conference on artificial life and robotics* (pp. 40–43).