Analysis of Short-term and Long-term Effects on Mental State of Suggestions Given by an Agent using Impasse Estimation

1st Yoshimasa Ohmoto
Department of Informatics
Graduate School of Integrated Science and Technology
Shizuoka University
Shizuoka, Japan
ohmoto-y@inf.shizuoka.ac.jp

2nd Hanako Sonobe
Department of Intelligence Science and Technology
Graduate School of Informatics
Kyoto University
Kyoto, Japan
sonobe@ii.ist.i.kyoto-u.ac.jp

3rd Toyoaki Nishida
Department of Informatics
Faculty of Informatics
The University of Fukuchiyama
Kyoto, Japan
toyoakinishida@gmail.com

***Abstract*—For an agent to teach a person a problem-solving attitude by giving him advice that does not directly contribute to solving the problem, a strategy that considers changes in the person's long-term attitude must be designed. This study aimed to investigate how the mental state of participants performing a task is affected during short-term and relatively long-term periods when they are advised either based on their conditions or mechanically at regular intervals. We focused on metacognitive suggestions during insight problem-solving as an example of advice that would be effective even if given by the agent. By these means, the effect on the human mental state over a relatively long period of time when the agent gives advice is examined. We conducted an experiment using two types of suggestion agents and observed that participants were likely to accept metacognitive suggestions provided by an agent when the suggestions were given based on an inner-state estimation of the participant. An analysis of mental state changes based on physiological indices suggested that the use of metacognitive suggestions by agents based on participants' conditions affected the mental state in problem-solving activities in the short and long term. It is also suggested that if the advice is not given depending on the situation, the effect of the advice in mitigating the impasse reduces as the task progresses. These findings will contribute towards the implementation of a tutoring agent.**

***Index Terms*—Human-agent interaction; metacognitive suggestion; insight problem solving.**

I. Introduction

This article is an extended version of the authors' paper "Difference in Attitudes toward Suggestions Given by an Agent using Impasse Estimation" [1], presented at the Twelfth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC2019). In this paper, based on the temporal changes in the participants' physiological indices measured during the experiment, we analyzed in detail the changes in mental state throughout the experiment, the changes in mental state for each task with different properties, and the changes in mental state due to the intervention by the agent. In this way, we investigated not only the effect on the mental state immediately after the agent's intervention, but also the long-term effect on it due to the repetition of the interventions of the agant. In addition, we described details of the specific interventions of the agents and the tasks performed in the experiment, which were omitted in the proceedings of CENTRIC2019 due to space limitations.

In learning and teaching situations, when learners are working on problem-solving, those who are knowledgeable about the task are often encouraged to develop an attitude toward learning. This is a way of thinking about problem-solving itself by encouraging them to broaden their horizons and learn via trial and error rather than by giving them direct advice to help solve the problem. This teaching strategy does not contribute directly to the solution of the problem. Therefore, if you do not consider the condition of the person you are communicating with, you may not be able to convey your intentions correctly or you may not be able to be considered your opinions. For example, if you repeatedly give advice about something that the listener does not perceive to be a problem, they may ignore the advice. Such problems become more pronounced when systems such as agents provide advice. One of the reasons for this is that agents' ability to grasp the situation seems to be relatively low from a human standpoint.

In order to avoid such issues, the agent needs to understand the human state and give advice. However, even a human often fails to estimate the mental state of the person they are communicating with. When you have a trustful relationship with the communication partner, this is not necessarily a fatal problem. However, in a human-agent interaction, the human often needs to infer the agent's behavior model based on a small number of interactions. Therefore, it is expected that a small number of failures in interaction will cause errors in the behavioral model of the agent constructed by humans. For example, one approach to getting people to accept an agent's advice is to show that the agent has expertise by having the agent consistently provide the appropriate advice [2]. The human often accepts the advice of the agent when the agent provides appropriate advice depending on the task situation.

However, if the agent fails early in the interaction because of misunderstandings, the human may stop accepting further advice. In addition, it is often difficult to determine whether or not the advice is appropriate when the advice does not directly lead to the correct answer for the task being performed. In such instances, effective advice cannot be given without considering the situation and intention of the person performing the task. In order for an agent to teach a person something like a problem-solving attitude by giving him advice that does not directly contribute to solving the problem, it is necessary to consider an advice-dispensing strategy that takes changes in the long-term attitude of the person into account.

Metacognitive suggestions are useful for problem-solving, although they do not lead to the direct outcome of the task being performed. Several previous studies had attempted to improve task performance using pre-training to induce metacognition. Patrick et al. [3] reported the impact of general metacognitive training on performance. Metacognitive suggestions may convey knowledge of how participants solve problems and can facilitate changes in their way of thinking during insight problem-solving (e.g., [4]). In this study, we consider metacognitive suggestions during insight problem-solving as an example of advice that would be effective even if the agent gave it. The effect of an agent's metacognitive suggestions on the human mental state over a relatively long period of time is then examined.

We tried to encourage the acceptance of metacognitive suggestions from agents by controlling the contents and timing of suggestion presentation as per the state of the participant. In many previous studies (e.g., [5] [6]), the agent advises a participant when there is a pause in the conversation. In this study, in order to provide convincing advice to humans, we focused not on the content of the problem, but on the state of thinking about the problem and the awareness of problems. The advantage of this method is that it does not recognize the content of the task and can provide advice based on how difficult the person is feeling the task at an appropriate time.

On the other hand, it takes a relatively long period of time for the learners to develop an attitude toward learning and toward problem-solving. In this study, participants were asked to repeatedly perform a similar task that became progressively more difficult. The purpose of this study was to investigate how the mental state of the participants in performing a task is affected when the participants are given advice either depending on their conditions or mechanically at regular intervals. If it is important to be able to customize the timing of providing advice as per the condition of the participant in order for the agent's advice to be understood with long-term effect, the method proposed in this paper may be useful when developing a tutoring agent.

The present paper is organized as follows. The Suggestion system using impasse estimation section contains an explanation of a system developed to give metacognitive suggestions based on the estimated state of the person performing the insight problem-solving task. The Experiment section describes the results of an experiment to evaluate the system implemented on the agent. In the Discussion section, the achievements of this research and some future works are described. The conclusions are presented in the Conclusion section.

II. Suggestion system using impasse estimation

Insight problem-solving contains four steps: impasse, incubation, illumination, and validation [7]. We focus on the impasse step in which people repeatedly searches inappropriate problem space that does not include a solution. In the impasse step, advice from other perspectives is useful for constraint relaxation and a switch of problem space. Metacognitive suggestion is one method of providing acceptable advice for the constraint relaxation [3] [8] [9] [10]. The metacognitive suggestion is confirmed to be effective even if it is presented at random timing. This is because the insight problem-solving task is prone to fall into the impasse state, and therefore there is a certain probability of being in the impasse state when presented at random.

In order to confirm the appropriate timing of advice in an insight problem-solving task, we conducted a preliminary experiment in which an experimenter determined the content and timing of the agent's metacognitive suggestions using the Wizard of Oz (WoZ) and presented them to the person performing the task. The task in the preliminary experiment was an "escape room game" in which players were often at a stalemate because they were required to think from a different perspective to win this game. In this game task, the participant must escape from a virtual room using various game objects. After this preliminary experiment, some participants reported that they were "given proper advice", so we thought that the advice of the agent by WoZ operation was accepted. When we observed the behavior of the participants and advice of the experimenter in the preliminary experiment, the experimenter provided suggestions when the participant seemed to be at a stalemate. We regarded this state of stalemate as an impasse. We consider that the state of stalemate is one of the appropriate clues to provide metacognitive suggestions. In accordance with this concept, we developed a system to provide advice by estimating whether the interaction partner is in an impasse state while working on an insight problem-solving task.

A. Estimation of strategies to perform the insight problem solving task

In order to find typical strategies to perform the insight problem-solving task, we observed the behavior of the participants in the preliminary experiment. As a result, it was expected that the participants switched two strategies: depth-first search and breadth-first search. In the state of depth-first search, participants focused on a particular object and looked for ways to use it successfully. In the state of breadth-first search, participants saw the overall situation of the task to search whether there were any missing or untried methods. Since it is conceivable that a stalemate may occur while executing each strategy, human inner states in insight problem-solving can be classified into 4 states (table in Figure 1). In

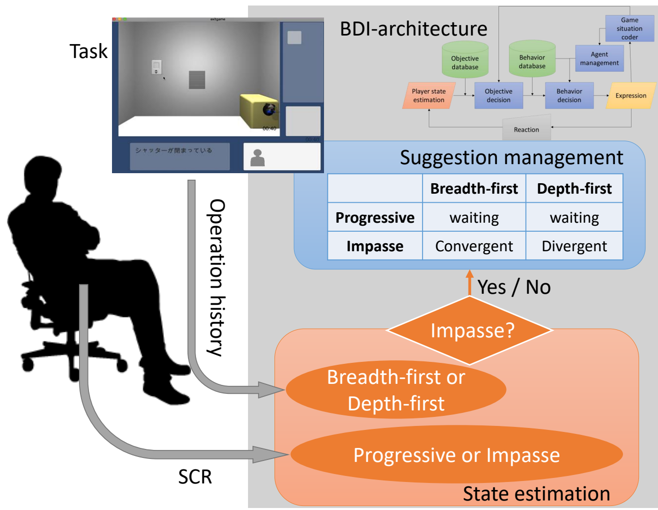

Fig. 1. Outline of the system architecture.

the advice by the experimenter, there were many suggestions that urged the participant to look for other ways to solve the task in the depth-first search, and there were many suggestions that encouraged the participant to look back on his/her own behavior and to focus on the specific object in the breadth-first search.

It is difficult to infer the inner state of thinking from the participant's behavior, specifically the inner state of thinking whether the participant is at a stalemate. To estimate this inner state, we analyzed physiological indices obtained during the preliminary experiment. In our previous work, we reported to estimate the feeling of difficulty of the task by using physiological indices [11]. As a result, it was frequently observed that Skin Conductance Response (SCR) was often activated, when the unfamiliar object was discovered during the task and when the situation in the task was changed. In addition, even when the situation did not change, the SCR was often responsive when with repeated trial and error such as looking for hints or checking previous information. Therefore, we regard the state as a non-impasse state (the participant is not at a stalemate) when the responses of SCR are frequently observed, and we regard the state to have shifted to the impasse state (the participant is at a stalemate) when the response of SCR is not observed for a certain time.

We also measured the electrocardiogram. However, we have not been able to obtain a useful feature for estimating task impasse from the electrocardiogram. Therefore, no electrocardiogram data was input to the system. We used electrocardiographic data to assess participant's mental states to tasks.

B. The outline of the suggestion system using impasse estimation

Figure 1 shows an outline of the agent design. This agent basically decides own behavior based on the Belief-Desire-Intention (BDI) architecture. This agent estimates two kinds

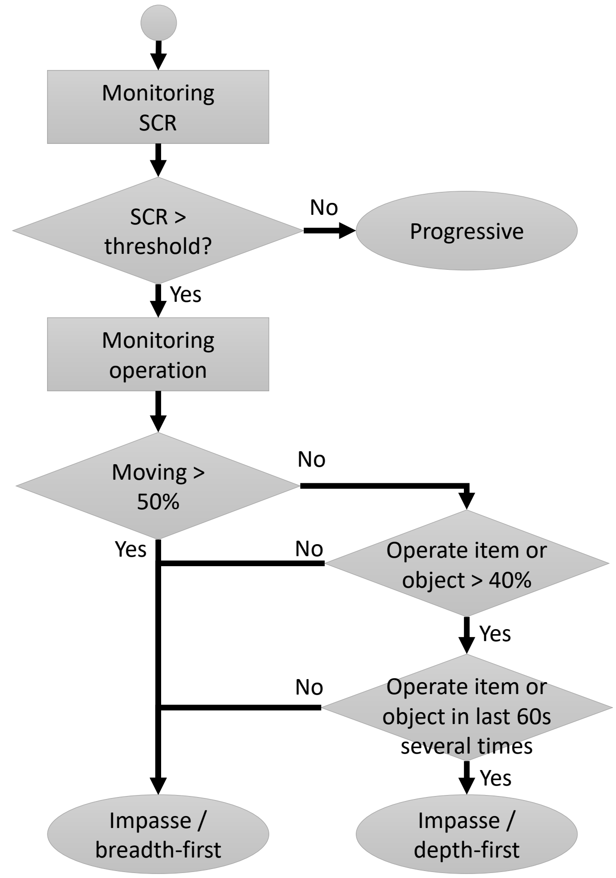

Fig. 2. Outline flowchart of the thinking-mode estimation.

of user states: a thinking mode (depth-first or breadth-first) and a state of the stalemate (impasse or progressive). The user's overall states are categorized into one of four combinations: depth-first/progressive, breadth-first/progressive, depth-first/impasse, and breadth-first/impasse. The agent provides a metacognitive suggestion when the estimated user's state includes "impasse." A convergent suggestion is provided in the state of breadth-first/impasse. A divergent suggestion is provided in the state of depth-first/impasse.

The user's physiological index and behavior are measured to estimate the user's state. The state of the stalemate is estimated using the measured SCR. The agent estimates the user's state as impasse when the SCR does not respond during a defined time window. The time window and the threshold to estimate the state of the stalemate are decided based on the measured data for two minutes from the start of the task. To estimate the thinking mode, the operation history log of the user is used. When the user repeatedly operates a game object (such as, a key, a scissors, a piece of paper, a door, a dial plate, a drawer of a desk, a closet, a button and a safe) in high frequency, the agent estimates the thinking mode to be depth-first search. The outline flowchart of the thinking-mode estimation is shown in Figure 2.

Ten convergent and ten divergent suggestions were prepared. The suggestions were not dependent on a particular task because they were metacognitive suggestions. One of the suggestions is selected randomly when the agent provides an advice. In general situations, it is necessary to give advice considering the context of the task though it is a metacognitive suggestion. In this study, we focused on the effect of controlling the timing of the metacognitive suggestions provided based on the state of the user. Therefore, the agent advised only considering whether the context was divergent or convergent in our experiment. The following is a list of metacognitive suggestions we prepared.

Divergent suggestions

- Why don't you look at something a little different?
- Don't stick to the way you've been doing things, think about different ways.
- Why don't you consider some other ways to proceed?
- Think of a way that's different from the way you failed.
- Is there anything else you can do?
- Why don't you think about something different from what you have been looking at?
- Please try to look at the situation from a different point of view.
- Let's think about what else you can do.
- Is there anything else you haven't done yet?
- Try to get rid of your assumptions.

Convergent suggestions

- Let's think about what's important in what you have seen.
- Let's try what comes to mind.
- Try to sort out what you have been doing.
- Think about what you need to do to escape.
- Have you missed anything so far?
- Let's think about what's been the inspiration so far.
- What element do you think is involved in the escape?
- Let's think back to what you have done so far.
- If you think you can do something, try it.
- Why don't you narrow down what you're focusing on?

III. Experiment

When we try to intervene in the behavior or decision-making of the other person by providing advice, especially during interaction with a less socially related interaction partner, it is important to provide appropriate advice based on the estimation of the partner's inner state. We considered the metacognitive suggestion in the insight problem-solving task as an example of the useful advice that the agent can provide. Then we proposed the suggestion system based on the impasse state estimation of the partner in order to accept the suggestion by the agent. We used two types of suggestion agents in this experiment. One was a state-considering agent that estimated the user's state before providing a metacognitive suggestion (sc-group). Another was a fixed-interval agent that provided a metacognitive suggestion in three-minute intervals (fi-group).

A. Task

Participants played an "escape from the room" game. The objective of this game is to escape from a closed space such as a room by utilizing game objects and items that are placed in that space. In most escape games, the player cannot escape from a room in a simple way, such as unlocking the door. The player escapes by searching for keys that are hard to find, by manipulating game objects in specific steps, and/or by using items in ways that are different from their common uses. In this game, players were often at a stalemate because they are required to think from a different angle to escape from a virtual room using various game objects.

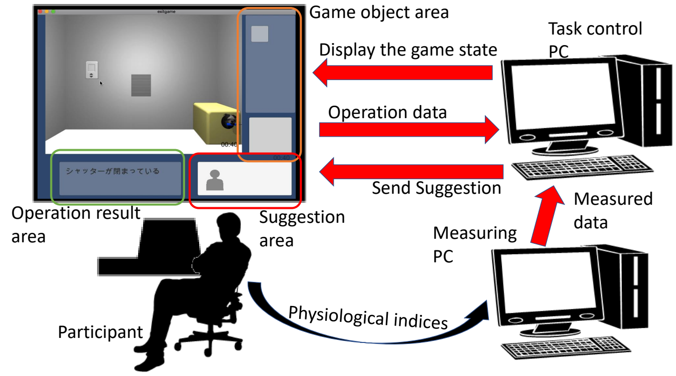

Fig. 3. The experimental setting.

The player can see images representing the four directions of the room, as well as partial enlargements of the images. If the player can explore the room and is able to move to the other side of the wall by opening a door or making a hole in the wall, the escape is successful. There are several non-movable game objects in the room (desk, chair, window, safe, etc.) and movable items (hammer, key, notepad, etc.). Descriptions of the game objects or items are displayed in the description display area in the game screen. The player can use the keyboard and mouse to change his/her viewing direction, zoom in on objects of interest, and use items.

The participants are asked to escape from three rooms. As the number of game objects and items and the steps to escape increase, the difficulty of escaping gradually increases. The order of the rooms that the participant escaped from was fixed. There was a 15-minute time limit to escape from each room. The suggestion agent explained the procedure for escaping from the room when the participant exceeded this time limit. After escaping from a room, the participant was allowed to rest. Participants were able to continue the game by pressing the start button when they wanted to resume.

B. Exprimental setting

Figure 3 shows the experimental setting. Each participant sat in front of a 27-inch monitor that displayed the game. A video camera was placed behind the participant to record his/her behavior and the game playing screen. The participant's voice was recorded using microphones. Polymate was used to measure SCR and the electrocardiogram (heart rate variability). SCR was measured with electrodes attached to the first and third fingers of the participant's non-dominant hand.

The electrocardiogram was measured by connecting electrodes with paste to the participant's left side, the center of the chest, and both ears for ground and reference. The experimenter sat out of view of the participant and observed the participant's behavior. The suggestions by the agent were provided using audio and text. The participants performed the task using a mouse.

C. Procedure

First, each participant was briefly instructed on the experimental procedure. Electrodes for measuring SCR and the electrocardiogram values were then attached to the participant's left hand and chest. After the installation, each participant played a practice game to confirm the operating method and basic flow of the game. The experimenter instructed the participant on the basic operation method. In addition, the participant was given an overview of the agent providing metacognitive suggestions. After receiving questions from the participant and confirming his/her understanding, the participant started the "escape from the room" experiment. After the experiment, the participant answered NASA Task Load Index (NASA-TLX) to measure the mental workload.

Forty-two undergraduate students, 27 males and 15 females, participated in the experiment. The average age was 20.8 years with a standard deviation of 1.9 years. We eliminated 13 participants because they did not need suggestions to escape from one of the rooms. Therefore, we used data of 29 participants (sc-group: 14 participants, fi-group: 15 participants).

D. Results

We analyzed the frequency of metacognitive suggestions, operation history log, mental workload, and physiological indices. For the frequency of metacognitive suggestions, we analyzed how many times each of the participants in each group provided metacognitive suggestions to encourage divergence and convergence in each room. From the operation history log, it was analyzed whether the state transition was carried out within a fixed time. The analysis range was 10 seconds after the suggestion. In the analyses of the physiological indices, we used heart rate variability (this is converted to cardiac sympathetic index (CSI) and cardiac vagal index (CVI)), and SCR. In the analysis of mental workload of the task, we used the Japanese version of the NASA-TLX, which represents the physical and psychological load of the task.

1) The frequency of metacognitive suggestions : We analyzed whether there was a difference in the frequency of metacognitive suggestions provided in the sc-group compared with the fi-group. There were two types of the metacognitive suggestions (divergent and convergent), so we performed a 2 (group: state-considering or fixed-interval) x 3 (room: first, second or third) analysis of variance (ANOVA) separately. Since each participant spent different amounts of time in each room, we compared the number of suggestions per minute. Logit transformed values were used in ANOVA to test for differences. The results are shown in Figure 4, Figure 5, Table I, Table II, Table III, and Table IV. In the tables, "SS" means the sum-of-squares, "df" means the degrees of freedom, "MS" means the mean squares, "F" means the F ratio, and "p" means the p-values.

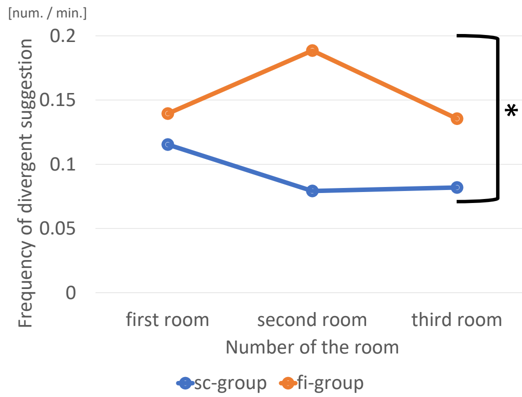

Fig. 4. The frequency of divergent metacognitive suggestions per minute.

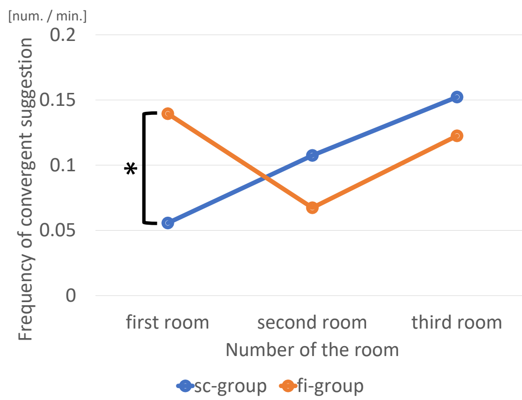

Fig. 5. The frequency of convergent metacognitive suggestions per minute.

In the divergent suggestions, there were significant differences between groups (sc-group $<$ fi-group) and between rooms (first and second $>$ third). The interaction was also significant. When tested for simple main effects, there were significant differences between groups in second room and third room (sc-group $<$ fi-group). There was also a significant difference between the rooms in the sc-group (first $>$ second and third). This result indicates that, in the sc-group, a relatively large amount of divergent suggestions was provided in the first room where the task execution method is unclear for the participants, and that trial and error is encouraged. In addition, in second room and third room where people seem to be used to the task, suggestions were reduced.

In the convergence suggestions, there was no significant difference between groups, but there were significant differences between rooms (first and second $<$ third). The interaction was also significant. A simple main effect test showed a significant difference between groups in first room (sc-group $<$ fi-group). It was also found that there were significant differences between the rooms in the sc-group (first $<$ second and third). This result shows that the convergent suggestions in

TABLE I
RESULT OF THE ANOVA ON THE FREQUENCY OF DIVERGENT METACOGNITIVE SUGGESTIONS.

source	SS	df	MS	F	p
A: group	6.50	1	6.50	14.83	<0.001 ****
error[S(A)]	11.84	27	0.44		
B: room	1.81	2	0.90	5.13	0.0091 **
AB	1.17	2	0.58	3.32	0.044 *
error[BS(A)]	9.49	54	0.18		

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$

TABLE II
THE SIMPLE MAIN EFFECT OF THE ANOVA ON THE FREQUENCY OF DIVERGENT METACOGNITIVE SUGGESTIONS.

effect	SS	df	MS	F	p
A(first)	0.38	1	0.38	1.43	0.235
A(second)	4.32	1	4.32	16.39	<0.001 ****
A(third)	2.98	1	2.98	11.30	0.0012 ***
error		81	0.26		
B(state-considering)	1.90	2	0.95	5.42	0.0072 **
B(fixed-interval)	1.07	2	0.53	0.04	0.056 +
error		54	0.18		

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$

TABLE III
RESULT OF THE ANOVA ON THE FREQUENCY OF CONVERGENT METACOGNITIVE SUGGESTIONS.

source	SS	df	MS	F	p
A: group	0.77	1	0.77	2.03	0.165
error[S(A)]	10.21	27	0.38		
B: room	1.15	2	0.57	4.10	0.022 *
AB	3.17	2	1.58	11.30	<0.001 ****
error[BS(A)]	7.57	54	0.14		

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$

TABLE IV
THE SIMPLE MAIN EFFECT OF THE ANOVA ON THE FREQUENCY OF CONVERGENT METACOGNITIVE SUGGESTIONS.

effect	SS	df	MS	F	p
A(first)	3.83	1	3.83	17.42	<0.001 ****
A(second)	0.015	1	0.015	0.067	0.796
A(third)	0.099	1	0.099	0.45	0.50
error		81	0.26		
B(state-considering)	3.19	2	1.60	11.39	<0.001 ****
B(fixed-interval)	1.12	2	0.56	4.01	0.024 *
error		54	0.14		

+ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .005$; **** $p < .001$

the sc-group was reduced in first room, which was a relatively simple.

Overall, the control of metacognitive suggestions was reasonable to some extent.

2) Operation history log: After the metacognitive suggestion was provided, we analyzed from participants' operation history log to determine whether they were acting in line with the suggestion. From the operation history log, we checked whether the transition to another state occurred within 10 second after the suggestion was given. In divergent suggestions, if a state transition was made, it was considered that the suggestion was accepted. In convergent suggestions, if no state transition was made, the suggestion was accepted. The result is shown in Figure 6 and Figure 7. The chi-squared test was applied to determine whether there was a difference between the groups in the acceptance rate of divergent suggestions and the acceptance rate of convergent suggestions.

We compared the acceptance rate of all suggestions between groups. As a result, the acceptance rate of all suggestions in the sc-group was significantly higher than that in the fi-group (p = 0.0013). We compared the acceptance rates of divergent suggestions and convergent suggestions between groups. Although there was no significant difference in divergent suggestions (p = 0.01), the acceptance rate of convergent suggestions in the sc-group was significantly higher than that in the fi-group (p = 0.0061). We compared the acceptance rates of divergent suggestions and convergent suggestions in each group between rooms. In third room, the acceptance rates of both divergent suggestions and convergent suggestions in the sc-group were significantly higher than those in the fi-group (divergent: first room p = 0.72, second room p = 0.16, third room p = 0.005, convergent: first room p = 0.44, second room p = 0.36, third room p = 0.014). In addition, in the sc-group, the acceptance rates in third room were higher than those in first room, and the acceptance rates seems to be gradually increasing. It is not clear whether this is because the difficulty of the room is increasing or because the reliability of the agent's suggestions is increasing. In any case, the results showed that the participants were likely to accept the metacognitive suggestions provided by the agent when the suggestions were given based on the inner state estimation of the participant.

3) Mental workload: We measured mental workload using NASA-TLX. This is major method to measure the mental workload. Figure 8 shows the results. With the exception of "performance," the sc-group reported an overall lower mental workload than the fi-group. We performed Welch's t-test on the total score between the two groups and there is no significant difference (t(27)=-1.42, p=0.17). We also performed Welch's t-test on each individual score between the sc-group and the fi-group. There was a significant difference regarding the data of "temporal demand" (sc-group < fi-group, t(27)=-2.18, p=0.038). The results suggest that advice based on human internal state estimation reduces some of the human mental workload. At the same time, it shows that the overall effect is not significant.

4) Analysis of changes in mental state based on physiological indices: To investigate whether a change that was not apparent from the human's behavior occurred in their mental state, we analyzed physiological indices. The physiological indices used in this study are heart rate variability (this is converted to CSI and CVI), and SCR.

CSI is one of the indices of sympathetic nerve activity. The sympathetic nervous system's primary function is to stimulate the body's fight-or-flight response, in terms of perceptible reactions such as tension and excitement. The CVI is one of

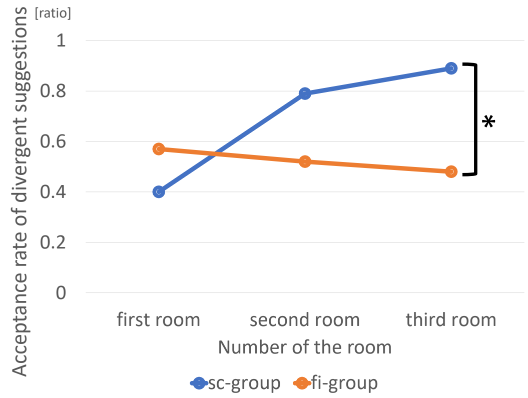

Fig. 6. Acceptance rates of divergent metacognitive suggestions.

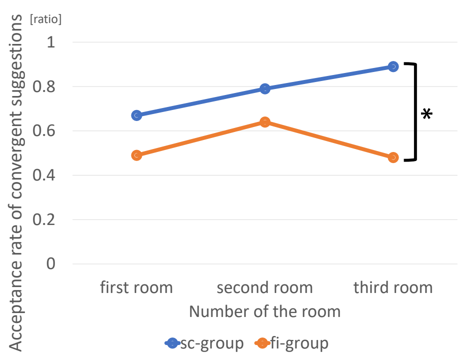

Fig. 7. Acceptance rates of convergent metacognitive suggestions.

the indices of parasympathetic nerve activity. The parasympathetic system is responsible for stimulating the "rest-and-digest" activities that occur when the body is at rest and relaxed. We used the geometric Lorenz plot method [12] to calculate the CVI and CSI.

SCR is a kind of electro-dermal activity that includes skin potential activity and skin conductance activity. People sweat during exercise but their palms have only a few sweat glands for body temperature adjustment. Therefore, by measuring the electrical resistance on the palms, it is possible to check for the presence or absence of emotional perspiration [13]. Given that the underlying mechanisms of SCR and electrocardiograms are different, we assumed that they could be used to distinguish between different responses from different sources of stress. Sweating is controlled by the sympathetic nervous system [14] and can be induced by emotional stimuli, intellectual strain, or painful cutaneous stimulation. The underlying mechanisms of SCR are related more to anticipation, expectation, and attention concentration [15]. We thus expected that the SCR could be used to tell when someone is dealing with an unexpected or thrilling situation.

a) Analysis of changes in mental state in each room: In order to analyze the changes in the mental state of participants who tackled a difficult task while receiving advice from an agent, we analyzed the state in which they started working on the task and the state before they reached a solution in each

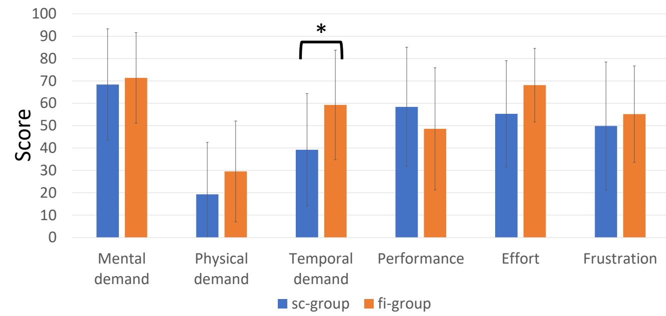

Fig. 8. Results of mental workload measurements.

room. For this purpose, the time required for escaping from each room was divided into three for each participant and the early and last data were extracted. For this data, CSI, CVI, and SCR were calculated. The results are presented in Table V. Paired t-test was performed on these data for each group.

In first room, the sc-group showed a significant increase in CSI as problem-solving progressed (CSI: $t(13)=-2.56$, $p=0.024$), but there was no significant difference in CVI and SCR. In the fi-group, there were no significant differences in CSI and CVI, but there was a marginally significant increase in SCR (SCR: $t(14)=-2.10$, $p=0.054$).

In second room, the sc-group showed a marginally significant increase in CSI as problem-solving progressed (CSI: $t(13)=-2.03$, $p=0.063$) but there were no significant differences in CVI and SCR. In the fi-group, there were no significant differences in CSI and CVI but there was a significant increase in SCR (SCR: $t(14)=-2.38$, $p=0.032$).

In third room, the sc-group showed a marginally significant increase in CSI as problem-solving progressed (CSI: $t(13)=-1.94$, $p=0.074$), but there were no significant differences in CVI and SCR. In the fi-group, there were no significant differences in CSI, CVI, and SCR.

Overall, the sc-group was adaptively advised via metacognitive suggestions and the sympathetic nervous system was more active in the late phase of problem-solving than in the early phase of problem-solving, suggesting that they had been working diligently on the problem-solving until escape. In other words, it can be suggested that by providing advice based on the condition of the participants, they were able to increase their involvement in the problem-solving efforts. In the fi-group, effects on SCR are observed in first room and second room. SCR tends to be lower when a person is absorbed in repeating tasks and higher when they are engaged in various trials and errors [16]. As the difference seems to be caused by the low SCR when they started working on the problem-solving, it is presumed that the fi-group might have fallen into an impasse and repeated the same action in that time. We think that the reason here were no differences in any of the physiological indices when the participants were in third room is that the problem-solving became more complicated, so the problem space became wider, and impasse in the form of repeating the same action was less likely to occur.

TABLE V
THE AVERAGES OF CSI, CVI AND SCR IN EACH ROOM.

		CSI		CVI		SCR	
room	group	early part	last part	early part	last part	early part	last part
first	sc	1.18	1.29	3.22	3.26	15.03	15.20
	fi	1.18	1.26	3.27	3.32	14.79	15.30
second	sc	1.29	1.38	3.29	3.30	14.78	14.97
	fi	1.42	1.36	3.38	3.36	14.00	14.96
third	sc	1.43	1.54	3.33	3.37	14.78	14.97
	fi	1.40	1.52	3.39	3.40	14.99	15.21

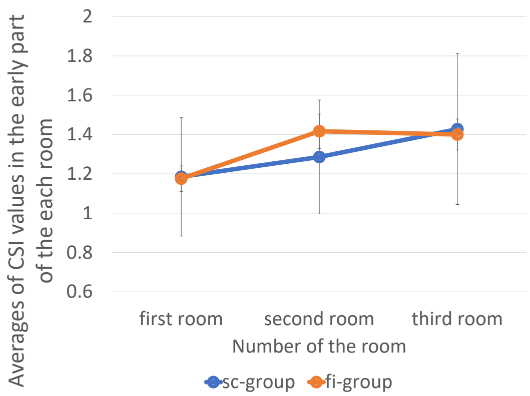

Fig. 9. The averages of CSI in the early part of the problem-solving in each room.

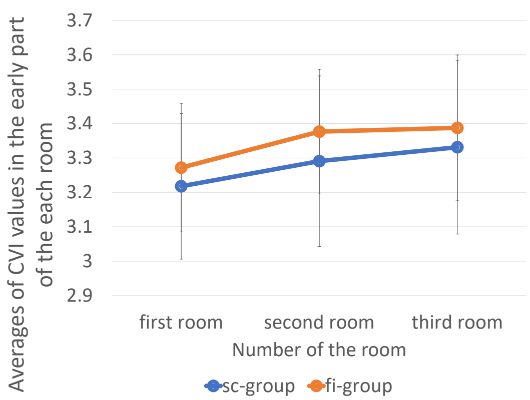

Fig. 10. The averages of CVI in the early part of the problem-solving in each room.

b) Analysis of changes in mental state in three consecutive rooms: In the early part of the problem-solving, differences between sc-group and fi-group seem to appear. Therefore, we analyzed the effect of the three rooms. A one-way analysis of variance was performed on the CSI, CVI, and SCR data in each group. The averages of each data in the early part of the problem-solving in each room are listed in Figure 9, Figure 10, and Figure 11.

The results indicate that there were no significant differences in CSI, CVI, and SCR in the sc-group, but there was a significant difference in CSI in the fi-group ($F(2, 14)=3.40$, $p=0.048$). There was a significant difference between first

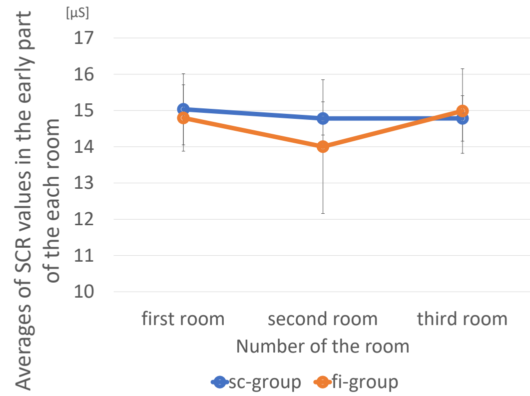

Fig. 11. The averages of SCR in the early part of the problem-solving in each room.

TABLE VI
MULTIPLE COMPARISONS IN THE MAIN EFFECT OF CSI BY RYAN'S METHOD.

pair	r	nominal level	t	p	sig.
second - first	3	0.33	2.33	0.027	s.
third - first	2	0.66	2.17	0.038	n.s.
second - third	2	0.66	0.16	0.874	s.

room and second room as well as between first room and third room (Table VI). The participants in the fi-group were more engaged in the problem-solving from the beginning in second room compared to first room, while participants from the sc-group demonstrated a gradual increase as they moved through the rooms. These differences may reflect the different attitudes toward problem-solving that were learned during the problem-solving trials in the previous room(s).

c) Analysis of changes in mental state after advice was given: The CSI, CVI, and SCR were calculated and compared for 10 seconds before and after giving the advice in each room to investigate the effect on the mental state of the participants immediately after the advice was given. Averages of values before and after advice are presented in Table VII. Paired t-tests were performed on these data for each group.

In first room, the sc-group showed a significant increase in SCR immediately after advice was given (SCR: $t(34)=-4.51$, $p<0.0001$) but there were no significant differences in CSI and CVI. In the fi-group too, there was a significant increase in SCR immediately after advice was given (SCR: $t(45)=-4.41$,

TABLE VII
THE AVERAGES OF CSI, CVI AND SCR BEFORE AND AFTER SUGGESTIONS IN EACH ROOM.

		CSI		CVI		SCR	
room	group	before	after	before	after	before	after
first	sc	1.21	1.20	3.25	3.26	13.26	16.56
	fi	1.23	1.25	3.31	3.29	14.11	19.14
second	sc	1.44	1.42	3.32	3.32	13.59	16.88
	fi	1.41	1.33	3.36	3.36	15.00	18.43
third	sc	1.48	1.44	3.34	3.34	13.44	16.41
	fi	1.35	1.38	3.36	3.37	15.02	16.73

$p<0.0001$) but again there were no significant differences in CSI and CVI.

In second room, the sc-group demonstrated a significant increase in SCR immediately after advice was given (SCR: $t(39)=-3.64$, $p=0.00039$) but there were no significant differences in CSI and CVI. In the fi-group, there was a significant increase in SCR (SCR: $t(37)=-3.10$, $p=0.0019$) but no significant differences in CSI and CVI.

In third room, the sc-group demonstrated a significant increase in SCR immediately after advice was given (SCR: $t(57)=-4.30$, $p<0.0001$), but there were no significant differences in CSI and CVI. In the fi-group, there was a marginally significant increase in SCR (SCR: $t(55)=-1.88$, $p=0.065$) but no significant differences in CSI and CVI.

Observing the overall trend, the sc-group appears to demonstrate that there is a relatively consistent and strong influence of advice on the participants, whereas the fi-group appears to be less influenced by advice as the task progresses.

IV. DISCUSSION

We hypothesized that participants would be likely to accept the advice based on the estimation of the inner state of the human even when the agent provided advice. In this research, we focused on metacognitive suggestions in an insight problem-solving task, which is one of the examples of useful advice that the agent can provide. We investigated the effects of metacognitive suggestions that controlled the timing of presentation based on human inner state. We implemented an agent that estimated two kinds of user states: a thinking mode (depth-first or breadth-first) and a state of stalemate (impasse or progressive). The agent categorized the participant's overall state as one of four combinations: depth-first/progressive, breadth-first/progressive, depth-first/impasse, and breadth-first/impasse. The agent provided a metacognitive suggestion with the goal of getting humans out of the impasse state.

We conducted an experiment using two suggestion agents. One was a state-considering agent that estimated the user's state before providing a metacognitive suggestion. The other was a fixed-interval agent that provided a metacognitive suggestion at three-minute intervals. Based on results from the analysis of the operation history log, the acceptance rate of suggestions in the sc-group was significantly higher than that in the fi-group. In other words, the attitude to the metacognitive suggestions given by the agent was different between the participants in fi-group and those in sc-group. The participants in sc-group believed that the content of the suggestions given by the agent should always be considered. By contrast, participants in the fi-group typically thought that the agent's suggestions presented general knowledge and tended to accept useful ones regardless of the task status. The results of the mental workload suggest that participants in the fi-group might interpret the agent's suggestions as a kind of facilitation of the task execution rather than human assistance.

An analysis of mental-state changes based on physiological indices suggested that the use of metacognitive suggestions by agents based on participants' conditions affected the mental state in problem-solving activities in the short as well as long term. Overall, changes in mental state were mainly reflected in CSI and SCR. While working on a single problem-solving task, it was suggested that when the agent's advice was given based on the participant's condition, the participants worked more diligently on the problem-solving task. In multiple tasks in succession, especially in situations where participants were beginning to tackle a new problem, they tended to be less likely to fall into the impasse of trying the same solution over and over again when they received advice that was provided based on their condition. As a short-term effect, a comparison of physiological indices before and after advice was provided suggested that both groups responded to such advice by trying new trial-and-error activities. However, it was also suggested that if the advice was not given depending on the situation, the effect of the advice in mitigating the impasse would be reduced as the task progressed.

During one room trial, the effect of advice based on the participant's condition appeared in the CSI. By contrast, the direct effect of the advice appeared in the SCR, while there was no change in the CSI. In addition, the direct effect of the advice was basically unaffected by the timing of the advice. The direct effect of advice on participants is providing a specific way of problem-solving. Therefore, it is suggested that the change in the current way of problem-solving by receiving the advice itself affects the mental state of the participants. It also suggests that mitigating the susceptibility to falling into impasses and changing attitudes toward the problem-solving requires appropriate advice based on the participant's condition. This is consistent with the different interpretations of the sc-group and fi-group advice described above. In other words, advice on actual task performance is considered to be acceptable to the person by providing appropriate advice on the status of the task, regardless of the state of the person. However, if we expect to improve the attitude toward the task through advice, such as to make the person have a broader perspective and be less likely to fall into an impasse, it is important to understand the person's condition at the time of giving the advice. The advice provided by the agent in this study was not directly helpful in solving the problem, but was a metacognitive suggestion that indirectly suggested how to solve the problem. Nonetheless, it is interesting to note that the immediate effect of the advice differed from the overall effect of the task execution. This is a finding that will contribute to the implementation of a tutoring agent. It suggests that an agent supporting active learning, which needs to maintain a positive attitude toward learning tasks, should provide advice and intervention based on an understanding of the human's current condition.

V. Conclusions

The aim of this study is to investigate how the mental state of participants performing a task is affected in the short term and for relatively long-term periods when the participants are given advice based on their conditions as opposed to mechanically at regular intervals. We implemented an agent that estimated two kinds of user states: a thinking mode (depth-first or breadth-first) and a state of stalemate (impasse or progressive). Based on experiments using two types of suggestion agents, we suggest that participants are more likely to accept metacognitive suggestions provided by agents when the suggestions are provided based on an inner-state estimation of the participant. With respect to the participant's mental state, an analysis of mental-state changes based on physiological indices suggests that the use of metacognitive suggestions by agents according to participants' conditions affects the mental state in problem-solving activities in the short and long term. As a short-term effect, a comparison of physiological indices before and after advice suggests that both groups responded to advice by trying new trial-and-error activities. It is also suggested that if the advice is not given depending on the situation, the effect of the advice in mitigating the impasse reduces as the task progresses. These findings will contribute towards the implementation of a tutoring agent.

Acknowledgment

This research is supported by Grant-in-Aid for Young Scientists (B) (KAKENHI No. 16K21113), and Grant-in-Aid for Scientific Research on Innovative Areas (KAKENHI No. 26118002) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Y. Ohmoto, H. Sonove, and T. Nishida, "Difference in attitudes toward suggestions given by an agent using impasse estimation," in *The Twelfth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services*. IARIA, 2019, pp. 11–16.

[2] R. G. Hass, "Effects of source characteristics on cognitive responses in persuasion," *Cognitive Responses in Persuasion*, pp. 141–172, 1981.

[3] J. Patrick, A. Ahmed, V. Smy, H. Seeby, and K. Sambrooks, "A cognitive procedure for representation change in verbal insight problems," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 41, no. 3, p. 746, 2015.

[4] Y. Hayashi, "Social facilitation effects by pedagogical conversational agent: Lexical network analysis in an online explanation task," *International Educational Data Mining Society*, 2015.

[5] S. T. Iqbal and B. P. Bailey, "Leveraging characteristics of task structure to predict the cost of interruption," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2006, pp. 741–750.

[6] K. Isbister, H. Nakanishi, T. Ishida, and C. Nass, "Helper agent: Designing an assistant for human-human interaction in a virtual meeting space," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2000, pp. 57–64.

[7] S. Ohlsson, "Information-processing explanations of insight and related phenomena," *Advances in the Psychology of Thinking*, vol. 1, pp. 1–44, 1992.

[8] T. Okada and H. A. Simon, "Collaborative discovery in a scientific domain," *Cognitive Science*, vol. 21, no. 2, pp. 109–146, 1997.

[9] H. Shirouzu, N. Miyake, and H. Masukawa, "Cognitively active externalization for situated reflection," *Cognitive Science*, vol. 26, no. 4, pp. 469–501, 2002.

[10] Y. Hayashi, "Togetherness: Multiple pedagogical conversational agents as companions in collaborative learning," in *International Conference on Intelligent Tutoring Systems*. Springer, 2014, pp. 114–123.

[11] Y. Ohmoto, T. Matsuda, and T. Nishida, "Experimentally analyzing relationships between larner's status in the skill acquisition process and physiological indices," *International Journal on Advances in Life Sciences*, vol. 9, no. 3 and 4, pp. 127–136, 2017.

[12] 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, pp. 79–84, 1997.

[13] R. Edelberg, "Electrical activity of the skin: Its measurement and uses in psychophysiology," *Handbook of Psychophysiology*, vol. 12, p. 1011, 1972.

[14] E. F. Bartholomew, F. Martini, and W. B. Ober, *Essentials of Anatomy & Physiology*. Benjamin Cummings, 2007.

[15] K. Hugdahl, *Psychophysiology: The Mind-Body Perspective*. Harvard University Press, 1995.

[16] Y. Ohmoto, S. Takeda, and T. Nishida, "Effect of visual feedback caused by changing mental states of the avatar based on the operator's mental states using physiological indices," in *International Conference on Intelligent Virtual Agents*. Springer, 2017, pp. 315–324.