

Having Avatar Nestle to User through Dialogues to Develop Exercise Habits with Intention Maintained

Tomoya Yuasa

Graduate School of Information Science and Engineering
Ritsumeikan University, Japan
Email: tomoya @ de.is. ritsumei.ac.jp

Fumiko Harada, Hiromitsu Shimakawa

Connect Dot Ltd., Japan
Graduate School of Information Science and Engineering
Ritsumeikan University, Japan
Email: {harada, simakawa}@ cs. ritsumei.ac.jp

Abstract—Continuation and habituation of exercises are in preventing lifestyle-related diseases. However, existing habituation applications fail to address mental factors of individual users. This study proposes a method to support habituation of exercise. It makes the best use of an accompanying avatars assigned to users according to their motivation. The avatar dialogues with a user every day. It proposes a goal based on the intention level and the current goal achievement of the user. It also shares the results of the analysis on collected data with the user at regular intervals. This method enables the user to continue the daily exercise easily. From the group of pre-survey subjects, we obtained 2 kinds of groups. We created avatars based on each group to verify their effectiveness. We found that the avatars have improved the goal achievements of the subjects.

I. INTRODUCTION

LIFESTYLE-RELATED diseases are a major problem worldwide. Diabetes, hypertension and coronary artery disease are on the rise worldwide [1]. The risk of lifestyle-related diseases can be reduced by physical activities. For example, a study by Manson et al. shows that physical activity may be a promising approach to primary prevention of NIDDM [2]. A study by Pfaffenberger Jr. et al. shows that initiating moderately vigorous sports activity, quitting smoking, maintaining normal blood pressure, and avoiding obesity are individually associated with reduced mortality from all causes and coronary heart disease [3].

There are IT-based ways to help people exercise on a daily basis. One example is the use of habit-forming applications that run on smartphones. Regardless of the type of exercise, there are a number of apps that run on smartphones to support habit formation [4]. These apps support the implementation of continuous exercise in a variety of ways. For example, one app allows users to record their actions. It let them reflect on themselves in order to make exercise a habit. Others set an alarm time to remind the user to exercise at a certain time. Despite the fact that many studies have shown that exercise is effective against lifestyle-related diseases, one-fourth of the world's adults do very little physical activity [5]. The main reason is that conventional apps promote exercise habit uniformly to all users. Users have individual differences, such as those who can maintain their motivation by being praised by others and who can act with a strong will. The applications on devises should reflect these differences to promote exercise

in a way appropriate for each user. In other words, there is a need for user-centered applications.

In this paper, we propose a method to support the habituation of exercise by using use of an accompanying avatars. In this paper, the avatar dialogues with the user to understand the user's state. It sets appropriate short-term goals to achieve long-term exercise goals. In addition, the avatar analyzes the collected data to share the analysis results with the user. The paper refers to avatars taking these behavior features of the users, as avatars accompanying them individually.

In the proposed method, users are classified into personas by motivation analysis based on (MSLQ), see[6]. Reflecting the analysis result, an avatar based on the persona dialogues with the user. The avatar calculates how positive the user's comments are, comparing them with a word-emotion polarity correspondence table [7]. The avatar calculates the user's motivation to exercise based on the difference of current states from the past ones. The avatar also asks the user to indicate the degree of achievement of the goal on a 5-point scale, indicating how much exercise the user take on that day. The avatar shows the optimal short-term goal to the user based on the user's motivation to exercise and the degree of goal achievement.

In this study, we first conducted a survey of what kind of avatars are effective for making exercise a habit. As a result, it is found that users need either avatars that sympathize with the user's utterances or ones that suggest specific things for the habituation of exercise. We developed two types of avatars based on the results of this survey. The study tries to address the following questions.

- 1) Does dialoguing with avatars improve users' motivation to exercise and achieve their exercise goals?
- 2) Is it effective for avatars to make concrete suggestions or to sympathize with the user's utterances, rather than to simply engage in dialogue?
- 3) Is there a relationship of the help seeking ability with long-term reflection which is derived from collected user utterances and exercise goal achievement?

We have conducted three experiments. The results of the experiments have revealed the followings. There is a significant difference in the increase in physical activity when the avatar is used. There is a significant difference in the increase in physical activity when a user is accompanied by an avatar that empathized with the user's utterances or one

that makes specific suggestions. To support habit formation, it is effective to provide long-term reflection with users of high help-seeking ability. The remaining parts of the paper is organized as follows. Chapter 2 discusses existing habit-forming applications and the need for avatars. In Chapter 3, a method is proposed to support the habituation of daily exercise using an avatar. Chapter 4 describes the survey method and the results of the survey on what kind of avatars users need in order to construct the avatar. In Chapter 5, the experiment and its results are explained to demonstrate the usefulness of the proposed avatar. Chapter 6 discusses the causes and future issues inferred from the results. In Chapter 7, the remarks of the paper is summarized as the importance of the research, suggests for applications, and future extensions.

II. MAKE DAILY EXERCISE A HABIT

A. Current status of habit-forming apps

Currently, various applications on smartphones have been published for habituation. Katarzyna et al. the features of 115 habituation applications [4]. As a result, they found that many habituation applications have task tracking and graph display functions. On the other hand, we found that there are very few functions that motivate users through sending messages showing supports and comments from other users. This suggests that many habituation applications are designed for self-monitoring.

It is also important to set goals frequently. Locke et al. have suggested that goal setting should be an iterative process in which users evaluate their performance either to modify their goals or to set completely new goals [8]. However, the applications have no function that frequently presents the most appropriate goal to the users. The users must change their goals towards the most appropriate ones for themselves all the time. However, this places a large burden on the users.

Phillippa et al. found that it takes 60 days in average for specific behavior to get habitualized [9]. In addition to that, Navin et al. have reported that continuous participation of users in the gym activities is highly correlated with "their frequency of the participation" [10]. Therefore, it is necessary to work on specific behavior continuously to make it a habit. However, motivation can not always be kept high. Engagement in specific behavior with high frequency needs abilities to regulate oneself toward it. It is difficult to benefit from the existing habituation applications unless their users have strong will.

B. The need for avatars

In order to maintain human motivation for a long time, dialogue is an important means because of persuading. For example, it has been found that persuaders who express positive emotions to persuade others are more likely to succeed in persuading others than those who do not express emotions [11][12]. On the other hand, a research has been conducted to increase the success rate of attempts to persuade users with dialogues considering their emotions [13].

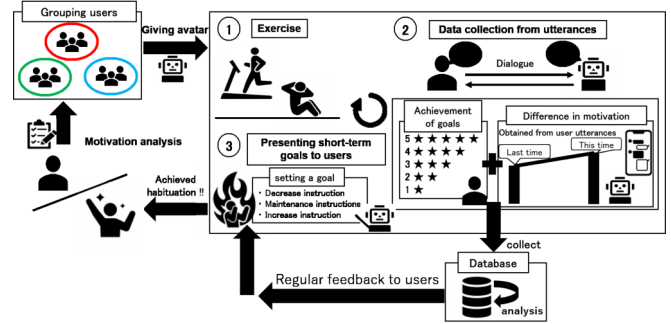


Fig. 1. a schematic diagram of the proposed method

If we focus only on dialogue for a small number of users, it is sufficient to prepare a human-to-human dialogue environment. However, in case of habituation applications to be used by tens of thousands of users or more, it is difficult to realize a function to respond immediately to the user's utterances. To address the problem, it is a promising way to implement avatars with chatbots. The chatbots make it easy to realize avatar-based applications respond immediately to many users 24 hours a day. According to Reeves et al., we would esteem exchanges of words. Humans have a tendency to treat any interaction as an interaction, between humans, even if the other party is a computer [14]. Chatbots that are available 24 hours a day are considered to be superior to human partners in that they can interact with users at any time.

In fact, chatbots that automatically interact are put into practical use in various fields. Fitzpatrick et al. developed Woebot, an automatic interactive chatbot. They found that it significantly reduced the severity of depression [15]. This suggests that the use of dialogue can be a great support for habituation.

III. MAKING BEHAVIORAL HABITS WITH AVATARS THAT ARE CLOSE TO USERS

A. Accompanying avatars

The paper proposes a method to support habituation of exercises. It makes the best use of an accompanying avatars assigned to users according to their motivation.

This study, realizes an "accompanying" avatar which understands the user's state through dialogues with the user. Its role is to set appropriate short-term exercise goals to achieve long-term goals toward a healthy life, promoting the user to continues exercises, while it analyzes the collected data from the user to sharing the analysis results with the user. Figure 1 shows a schematic diagram of the proposed method. This method allows users to continuously perform daily exercises that they determine themselves without difficulties.

In this method, we assume that users have a long-term exercise goal. First, users answer a questionnaire about their motivation. To the user, the method assigns both an appropriate persona based on the questionnaire result, and an avatar corresponding to the persona.

The persona is a classification of a user's personality in terms of motivation to take exercises. The avatar provides the user with motivation for exercise through an appropriate dialogue according to the persona. For example, an avatar can make specific suggestions or sympathize with the user's utterances.

Suppose users trying daily exercise. At the end of the day, a user dialogues with the avatar. The paper assumes the avatar presents the analysis results of user data. The dialogue is initiated by the avatar's utterance on the analysis results. It is followed by dialogues consisting of the user comments and succeeding responses from the avatar. The contents of responses depend on the avatar's persona. The user's comment and the avatar's response may be repeated several times.

From the content of the dialogue, the avatar obtains two values; the degree of achievement of the day's goal and the motivation of the user. The avatar sets a short-term goal based on the two values. It proposes the goal at the end of the dialogue.

B. Goal setting by avatar

The short-term goal values are calculated with a logistic regression, which takes the sum of the values of goal achievement and motivation as explanatory variables, while it. A logistic regression model is prepared for each persona. The short-term goal setting for exercise is one of the followings: to "reduce", "maintain" or "increase" the exercise load. It determines short-term goals that users can achieve the next day so that they may continue to exercise towards achievement of their long-term goals.

The degree of goal achievement shows how well the daily exercise is achieved. It is evaluated by the user in five levels at the end of the dialogue.

The degree of motivation of the user indicates the variation of the user feeling toward taking exercises. It is calculated from a numerical the user's utterances. It tells us whether the user's positive feeling toward exercises is increasing or not. If the avatar requests the user to increase the exercise load while the motivation is low, the user's motivation decreases, which makes it difficult for the user to continue exercising. On the other hand, an instruction to decrease the exercise load would make highly motivated users feel insufficient, which lessens their motivation to exercise. The degree of motivation of the user makes it possible to check whether exercise is painful or not.

Equation (1) is used to measure the degree of motivation of user P .

$$P = \frac{1}{k} \sum_{i=1}^k p_i - \frac{1}{k} \sum_{i=k+1}^{2k} p_i \quad (1)$$

Where i is the positive degree of the user's utterances i times ago in the dialogue. To know the motivation of the user, it is better to examine the tendency rather than instantaneous values. The proposed method uses the moving average.

The avatar performs morphological analysis on each comment C obtained from the user. The word prototypes extracted through the morphological analysis are compared with the word sentiment polarity correspondence table [7] to obtain the positivity of each word. Summing up the positivity of each word presents, the degree of positive of each user's comment. The avatar takes the difference of the average of the positivity of the user's last k comments from the average of the positivity of their preceding k comments. If P is greater than 0, the user's motivation to exercise tends to increase, while P less than 0, means the user get demotivated. The user's comments are analyzed sequentially. The calculation of every utterance enables us to obtain the time series of the user's positivity.

C. Create personas based on user motivation

Avatar are different in terms of ways of interventions with their messages. Some avatars suggest numerical goals, while others present sympathy with the user's utterances. Since users also have different personalities, how they are motivated to exercise also varies with each use. It is necessary to provide users with avatars suitable for their personality in advance. This study has developed a questionnaire based on the [6] to measure the motivation of users. MSLQ is a questionnaire to examine the motivation of learning in education. The study adopt 5 kinds of motivational factors in MSLQ: Control of Learning Beliefs, Extrinsic, Intrinsic, Self-Efficacy, and Task Value. Modifying examples in MSLQ as they go well with engagement in healthy exercises, the study prepared 4 questions for each of the motivational factors; 20 questions in total are presented to users to know their motivation. Every user answers each question on a 7-points scale. This study uses the 20 answers of the questionnaire as explanatory variables of a logistic regression model. Separately prepared users who answer the same questionnaire will be interviewed in advance about which avatar is suitable. The avatars belong to either of one that specifically suggest to achieve numerical goals, or one that empathize with the user's utterances. The study trains the model using their questionnaire answers labeled with their suitable avatars. The model is used to determine the avatar suitable for each of new users.

D. Suggestions based on users' reply comments

Users do not always have time to exercise. When they are too busy to exercise, they are expected to engage in another exercise to get the same effect in a shorter time. Repeating the same thing every day is important, but doing different exercises with the same effect is one way to continue the daily exercise.

Avatars showing concrete numerical goals are suitable for the way. For example, avatars should avoid presenting stereotype messages, such as "Keep it up!", which encourage the same exercise over and over again. Instead of it avatars had better suggest another exercise, say, "Doing 10 sit-ups has the same effect as 1000 step walking exercise". In this way, users can engage in exercise in various ways, which enable them to continue their daily exercise according to their own schedule.

E. Sympathy based on user response comments

In order to encourage users, it is important to offer words of encouragement. However, depending on the user's personality, simply saying "Good work" may seem like a formulaic sentence, which may lessen the effect of encouragement. It is important to offer words of encouragement sympathy with the content of the user's utterances.

An avatar that sympathizes with the user's utterances will respond to a user statement such as "It was cold today, but I did my best to go outside to walk a lot", might say "Walking in the cold until your body gets warm burns a lot of calories. It is very effective for dieting. If you can lose weight, it will also have a preferable effect on your health". In this way, the user feels that the avatar understands him or her, which raises the user's will to continue daily exercise.

F. Long-Term Reflection

The avatar stores the user's utterances and the degree of achievement of the day's goal in a database. Using data collected for one week, avatars analyze the states of users and their exercise status, to share the analysis results with the users. The system provides the users with the exercise trends of how much exercise they have done in a week, goals for the next reviews based on the exercise trends, to concrete suggestions for achieving the long-term goals. This allows the users to understand their own conditions. It also contributes to their sustainable engagement in daily exercise.

IV. IMPLEMENTING AVATARS FROM REAL-WORLD EXAMPLES

A. Searching for accompanying avatars

In order to create avatars that accompany to users, the study investigates what kind of avatars they need. The survey targets were 15 working adults (7 males and 8 females).

In this study, we investigated the avatars needed to improve the amount of walking. Since walking does not require any special skills or equipment, it is an exercise that is more accessible and can be incorporated into daily life compared to many sports[16][17].

The survey targets answered to a motivation questionnaire based on the MSLQ. We collected examples of dialogues which took place between the survey targets and the avatars. The survey provides as many kinds of messages as possible, to encourage engagement in walking exercise to the survey targets. After the dialogue, the survey targets accepted interviews. In the dialogues, the survey caused the survey targets to set a goal of walking 10000 steps every day. Mean while, the actual amount of walking of each survey target was classified in to the stages shown in Table I. The five stages in Table I are settled based on the results of the National Health and Nutrition Survey of the Ministry of Health, Labour and Welfare of Japan [18].

The survey used the "Wizard of Oz" testing, where the survey, instead of an avatar, sent expected various messages prepared in advance to the utterance of the survey targets. The dialogues are taken place, using the open chat function

TABLE I
5 LEVELS OF 10000 STEPS SETTING

5 levels of 10000 steps setting	
1	Less than 5000 steps
2	Less than 7000 steps~More than 5000 steps
3	Less than 8000 steps~More than 7000 steps
4	Less than 10000 steps~More than 8000 steps
5	More than 10000 steps

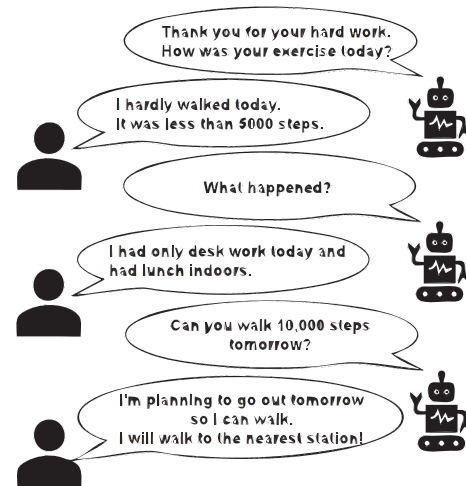


Fig. 2. Example of a dialogue with a survey avatar

of LINE [19]. An example of the actual dialogue is shown in Figure 2.

The interview was conducted on Zoom [20]. In the interview, the survey targets looked back at the actual dialogues in LINE. They answered whether the message from the avatar was appropriate, and if not, what would have been the ideal response. The survey collects appropriate messages in dialogues as well as answers or short-term goals. At the same time, they specify avatars needed to make habituation of exercises.

The K-means method is applied to the results of the questionnaire responses to classify the survey targets into personas. However, the classification of the questionnaire responses fails to find distinctive personas, such as a cluster with a high extrinsic motivation factor and a one with a low intrinsic motivation factor. The survey targets are also classified into personas based on the interview results. This time, the results showed that there were two types of personas: one who prefers to suggest specific things during dialogues, and one who prefers to empathize with the user's utterances. However, there are several survey targets those who could not be classified into either persona. The last group is referred to as the "miscellaneous persona". Due to the number of survey targets, common characteristics might not be found, because a result

TABLE II
PARTIAL REGRESSION COEFFICIENTS FOR PERSONAS WHO PREFER SPECIFIC SUGGESTIONS

Personas who prefer to suggest specific things	
Control of Learning Beliefs	-0.435
Extrinsic	-0.914
Intrinsic	0.261
Self-Efficacy	-0.660
Task Value	0.349

TABLE III
PARTIAL REGRESSION COEFFICIENTS FOR SYMPATHY-FAVORING PERSONAS

Personas who prefer to sympathize with the user	
Control of Learning Beliefs	0.0267
Extrinsic	0.0717
Intrinsic	-0.498
Self-Efficacy	0.1694
Task Value	-0.878

of the classification is only one person. The person might be assigned to miscellaneous persona. For miscellaneous persona, it is not possible to determine appropriate avatars because it is a mixture of different personas. Therefore, to check the effectiveness of avatars, the proposed method prepares avatars corresponding to two personas, one that prefers concrete suggestion of numeric goals, and the other that prefers sympathy of the avatar with the user's utterances.

To investigate the two personas, multiple regression analysis is applied to the results of the questionnaire. Let us check which of the five items is the questionnaire is important for each persona, looking at the partial regression coefficients of the multiple regression. The partial regression coefficients for both of the personas are shown in Table II and Table III.

Examination of the partial regression coefficients reveals, the persona preferring concrete suggestion consider the task value factor important. The task value favor corresponds to the evaluation on how assigned task is interesting, important, and useful. In other words, people who often think about how they feel the exercise to be performed or proposed tend to have preference on concrete suggestion. In the same way, the persona who prefers to avatar sympathy with the user's utterances values the self-efficacy factor. The factor is based on self-evaluation of one's ability to accomplish a task. It implies people who can make an effort without giving up easily even in difficult situations tend to prefer sympathy from avatars with their utterances.

B. Construct an accompanying avatar

For constructing a accompanying avatar, this study used Telegram [21], a tool for building messenger applications. In order to automatically respond to the user's dialogues by

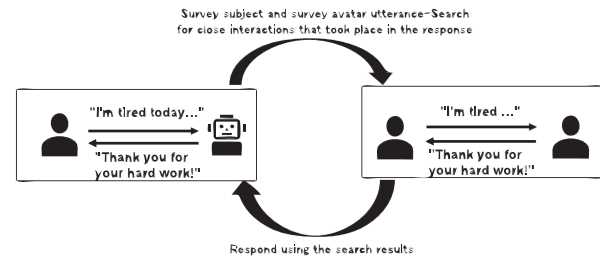


Fig. 3. How to dialogue with avatars

avatar, we created data based on the dialogues collected in the survey in Section 4.A, listing the pairs of the survey target's utterances and the survey avatar's responses or the ideal responses answered by the survey targets as examples. These data was stored in Elasticsearch [22]. Elasticsearch is a full-text search engine that can find appropriate examples from a large number of examples. The utterances-responses pairs of each avatar were mapped to each user persona, which was classified based on the interview results, to create a collection of examples for each avatar. The avatars dialogues each other using the example data stored in Elasticsearch. This is shown in Figure 3.

The avatar's response is generated as follows. To find sentences that are similar to the user's utterance, the similarity between the user's utterance and the survey target's utterance in each example is calculated using cos similarity. We calculate the cos similarity by assigning a vector with the frequency of words in the sentence as a component. By doing so, it finds the example query that is most similar to the user's utterance and responds to the user with a response to that example query.

C. Long-term reflection by an accompanying avatar

The user's utterances are stored in a database for long-term reflection. A user's long-term review is based on the subject's motor tendencies and the exercise tendencies based on the user's speech and motor status collected by Avatar over a period of time. This time, we couldn't implement the functionality shared by Avatar, so the overseer used her text message to share with the user the goals to reflect and specific suggestions for achieving them.

V. USEFULNESS OF EXERCISE WITH ACCOMPANYING AVATARS

A. Pilot field study Summary

In this study, we conducted a pilot field study to verify the following three points:

- 1) Does dialoguing with avatars improve users' motivation to exercise and achieve their exercise goals?
- 2) Is it effective for avatars to make concrete suggestions or to sympathize with the user's utterances, rather than to simply engage in dialogue?
- 3) Is there a relationship of the help seeking ability with long-term reflection which is derived from collected user utterances and exercise goal achievement?

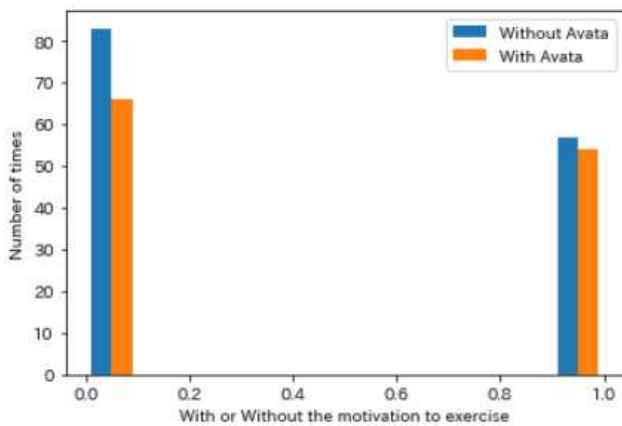


Fig. 4. Frequency distribution of subjects' motivation to exercise

The subjects were ten working adults (four males and six females). Each subject was answered a motivation questionnaire based on the MSLQ and a Help-Seeking questionnaire. Help-seeking is the act of asking for help from others [6]. Each subject submits daily walking and motivation to exercise for 14 days before the pre-pilot field study started. Motivation to exercise is represented with two options: motivation to exercise or not. The degree of goal achievement is calculated based on the submitted walking data, assuming a daily walk of 10000 steps. This period is referred to as the pre-pilot field study period. During the pre-pilot field study period, avatars were not used. For pilot field study period, the subjects were asked to take their daily activities for 12 days with the goal of walking 10000 steps every day. During the 12 days, each of the two kinds of avatars, as described in Section 4.B, is used for six days each. At the end of each day of the pilot field study, the subject reported the achievement of the exercise goal to the avatar and dialogued with it. The long-term reflection are realized using the Wizard of Oz test, that is, the supervisor of the pilot field study, instead of the avatar, reflected on the user's utterances and the degree of achievement of the exercise goals for 6 days periods. The results were shared with the subjects. We conducted questionnaires after the use of each avatar and after the completion of the entire pilot field study. The questionnaire included questions mainly about the avatars and whether or not the subject was motivated to exercise.

B. Usefulness of using leaning avatars

Figure 4 shows the frequency distributions of the data of "the degree of achievement of the goal of exercise during the 14 days during the pre-pilot field study". Figure 5 shows the frequency distributions of the data of "the degree of achievement of the goal during the 12 days during the pre-pilot field study" for all subjects.

In Figure 4, the horizontal axis indicates motivation. It means no motivation to exercise when the value is 0, while motivation to exercise when the value is 1. The vertical axis indicates the number of times the motivation to exercise. In

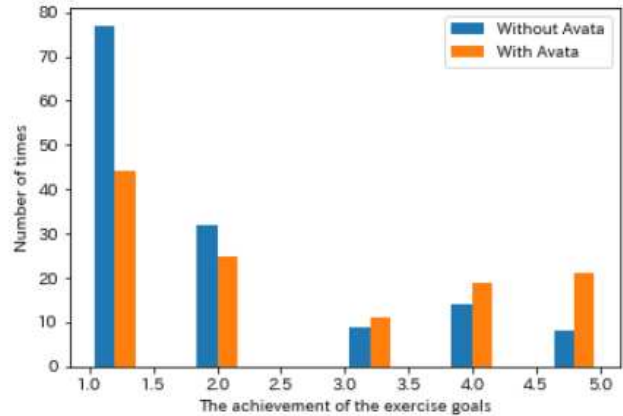


Fig. 5. Frequency distribution of subjects' goal achievement

Figure 5, the horizontal axis indicates the degree of achievement of the exercise goal. It indicates that the person is walking more steps when the value is higher. The vertical axis indicates the degree of achievement of the exercise goal.

The results of Figure 4 shows that the state of lack of motivation to exercise is lower than that of people who do not use avatars. The results of Figure 5 shows the number of times they achieved the highest value of the exercise goal increased compared to those who did not use the avatar. Therefore, we believe that users can improve their motivation to exercise and their achievement of exercise goals by dialoguing with avatars.

In order to verify whether the improvement in exercise motivation and goal achievement is significant, we conducted a statistical test. In the statistical test, we compared the motivation to exercise and goal achievement data of all subjects for each day of the 14-day pre-pilot field study period and the motivation to exercise and goal achievement data of all subjects for each day of the 12 days pilot field study period. As the distribution of the data was not normalized, the Mann-Whitney U test was used with a two-tailed significance level of 5%. As a result, a statistical value of 0.69 for motivation to exercise and 3.69 for goal achievement were calculated. We can say that there is a significant difference when the statistic value is 1.96 or higher. Therefore, it was found that the use of avatars did not significantly improve the motivation to exercise, but it significantly improve the achievement of exercise goals.

VI. USEFULNESS OF USING SYMPATHIZING AVATARS

As it is in 5.B, in order to test whether using an avatar that sympathizes with the user's utterances, rather than simply using an avatar, leads to an increase in the subject's motivation to exercise and achievement of exercise goals, we conducted a Mann-Whitney U test at a 5% two-tailed significance level. In the statistical test, we compared the motivation to exercise and goal achievement data of all subjects for 14 days pre-pilot field study period and the motivation to exercise and goal achievement data of all subjects for 6 days pilot field study periods using avatars that sympathize with the user's

TABLE IV
STATISTICAL TEST RESULTS OF EACH SUBJECT'S MOTIVATION TO EXERCISE IN THE SYMPATHIZING AVATARS

Subject	Motivation to exercise
A	1.89
B	2.21*
C	-
D	2.79*
E	0.20
F	2.09*
G	2.09*
H	1.24
I	4.35**
J	3.15**

statements. As a result, the statistical values of 0.09 for motivation to exercise and 3.68 for goal achievement were calculated. We can say that there is a significant difference when the statistical value is 1.96 or higher. Therefore, it was found that the use of avatars that sympathize with the user's utterances did not significantly improve the motivation to exercise, but it significantly improve the achievement of the exercise goals.

Next, we conducted a Mann-Whitney U test at the 5 two-tailed significance level for each subject in order to verify whether the use of avatars that sympathize with the user's utterances would not lead to an improvement in motivation to exercise. In the statistical test, we used the data of motivation to exercise of each subject for 14 days pre-pilot field study periods and the data of motivation to exercise each subject for 6 days pilot field study periods using avatars that sympathize with the user's utterances. The statistical values obtained from the statistical test results are shown in Table IV.

We can say that there is a significant difference if the value is 1.96 or higher. The values marked with * indicate that there is a significant difference this time. The values marked with ** indicate that there is a significant difference this time when the avatar was not used. The "-" indicates that there was no change before and after the use of avatar.

The results in Table IV shows that for some subjects, the use of avatars that sympathize with the user's utterances significantly improve the motivation to exercise.

A. Usefulness of using avatars as specifically proposed

As it is in the same way in 5.B, in order to test whether using an avatar that makes concrete suggestions, rather than simply using an avatar, leads to an increase in the subject's motivation to exercise and achievement of exercise goals, we conducted a Mann-Whitney U test at a 5% two-tailed

TABLE V
STATISTICAL TEST RESULTS FOR EACH SUBJECT'S MOTIVATION TO EXERCISE IN THE CONCRETE PROPOSAL AVATAR

Subject	Motivation to exercise
A	1.89
B	-
C	-
D	1.52
E	1.55
F	2.09*
G	2.09*
H	2.52*
I	4.35**
J	1.19

significance level. In the statistical test, we compared the motivation to exercise and goal achievement data of all subjects for 14 days pre-pilot field study periods and the motivation to exercise and goal achievement data of all subjects for 6 days pilot field study periods using avatars that makes concrete suggestions. As a result, a statistical value of 1.21 for motivation to exercise and 2.17 for goal achievement were calculated. We can say that there is a significant difference when the statistical value is 1.96 or higher. Therefore, it was found that the use of avatars that suggest specific things did not significantly improve the motivation to exercise, but it significantly improve the achievement of the exercise goals. Next, we conducted a Mann-Whitney U test at the 5 two-tailed significance level for each subject in order to verify whether the use of avatars to make concrete suggestions would not lead to an improvement in motivation to exercise. In the statistical test, we used the data of motivation to exercise of each subject for 14 days pre-pilot field study periods and the data of motivation to exercise each subject for 6 days pilot field study periods using avatars to make concrete suggestions. The statistical values obtained from the statistical test results are shown in Table V.

We can say that there is a significant difference if the value is 1.96 or higher. The values marked with * indicate that there is a significant difference this time. The values marked with ** indicate that there is a significant difference this time when the avatar was not used. The "-" indicates that there was no change before and after the use of avatar. The results in Table V shows that for some subjects, the use of avatars that provided concrete suggestion significantly improved the motivation to exercise.

TABLE VI
EACH SUBJECT'S HELP-SEEKING VALUE AND EVALUATION OF
LONG-TERM REFLECTION

Subject	Help-seeking	Long-term reflection
A	5.33	4
B	2.33	1
C	4.66	2
D	4.33	4
E	3.33	5
F	4.66	5
G	4.00	4
H	4.33	3
I	2.33	2
J	5.00	3

B. Sharing long-term reflections with users

As help-seeking is the act of asking for help from others, subjects with a strong help-seeking should feel that avatars are helping them with their own daily exercise habits. We researched the relationship between each subject's Help-Seeking value from the pre-experiment questionnaire and the 5-point evaluation of long-term reflection on the user obtained from the post pilot field study periods questionnaire.

In the post pilot field study periods questionnaire, the subjects answered whether they felt that the long-term reflection from avatars was effective and why. Subjects' opinions included "I feel that they understand me" and "I did not find the comments from avatars very appealing". It was found that not all the subjects had positive opinions. Table VI shows the values of Help-seeking and the 5 points evaluation of the reflections.

The higher the value of Help-seeking, the stronger the Help-seeking.

Subject A has a higher value of Help-seeking than the other subjects. Therefore, it is thought that Subject A is likely to feel that the long-term reflection is effective. In fact, subject A gave a high rating of 4 for the long-term reflection. On the other hand, because Subject B has a lower Help-seeking value than the other subjects, it is thought that Subject B is less likely to feel that the long-term reflection is effective. In fact, Subject B gave the lowest evaluation of long-term reflection among all subjects.

In order to research the relationship between Help-seeking values and reflection, the correlation coefficient between Help-seeking values and the 5 points rating for long-term reflection was researched. As a result, we obtained a positive correlation of 0.47. This indicates that sharing long-term reflections for subjects with high Help-seeking values is effective in supporting habit formation. The long-term reflection is based on the user's utterances and the achievement of the exercise goal. This indicates that dialogue is important for people with a strong help-seeking. On the other hand, people with low Help-Seeking values work to solve problems on their own, so they

do not need to receive advice from collected data.

VII. DISCUSSION

A. Non-significant increase in motivation to exercise

In the present study, the use of avatars did not significantly improve the motivation to exercise. The subjects answered "yes" or "no" whether they wanted to exercise their motivation to exercise. However, each person's consciousness of movement is different. Even in the same situation, one subject may be "yes" while another may be "no". Therefore, it was necessary to clarify the situation setting.

In addition to clarifying the situation, it was also necessary to set the evaluation in more detail, such as five levels, instead of two levels. For example, a questionnaire has "Do you want to exercise today?". It is possible that a person who answers "yes" if there is enough time, but "no" if he does not want to go outside because it is raining, cannot answer clearly if there is enough time and it is raining. It would be nice if the situation setting could be subdivided, but it is difficult to subdivide everything appropriately. Therefore, we think it is necessary to use a rank out of 5 or a rank out of 7 in order to have a certain range of responses.

B. Appropriateness of avatars to be attached

Based on the preliminary questionnaire, we investigated which avatars matched the persona of each subject: avatars that sympathize with the user's utterances, avatars that suggest concrete things, or other avatars. The five items of Control of Learning Beliefs, Extrinsic, Intrinsic, Self-Efficacy, and Task Value in the questionnaire results obtained in section 4.A were used as explanatory variables, and the persona classification of preference concrete, empathy, other as the objective variable.

A logistic regression model was created from the explanatory and objective variables. The regression equation was used to predict the optimal avatar by applying it to the subjects' questionnaire results. As training data, the subjects answered a questionnaire after the pilot field study : "What kind of avatars do you feel is most appropriate for you in supporting your exercise?". They answer from three types of avatars.

Figure 6 shows the prediction results of the logistic regression and the results of the questionnaire.

In Figure 6, the horizontal axis is the number of the most suitable avatar for each subject predicted by the logistic regression, and the vertical axis is the number of avatars that each subject answered as suitable for him/herself obtained from the results of the subject's questionnaire. 0 represents avatars that sympathize with the user's utterances, 1 represents other avatars, and 2 represents avatars that suggest specific things. Figure 6 shows that the prediction by logistic regression is not very adequate, 3/10.

In the questionnaire based on the MSLQ described in section 3.B, five motivational factors were employed: Control of Learning Beliefs, Extrinsic, Intrinsic, Self-Efficacy, and Task Value. However, there are other motivational factors that were not used in this study. Therefore, in order to improve the

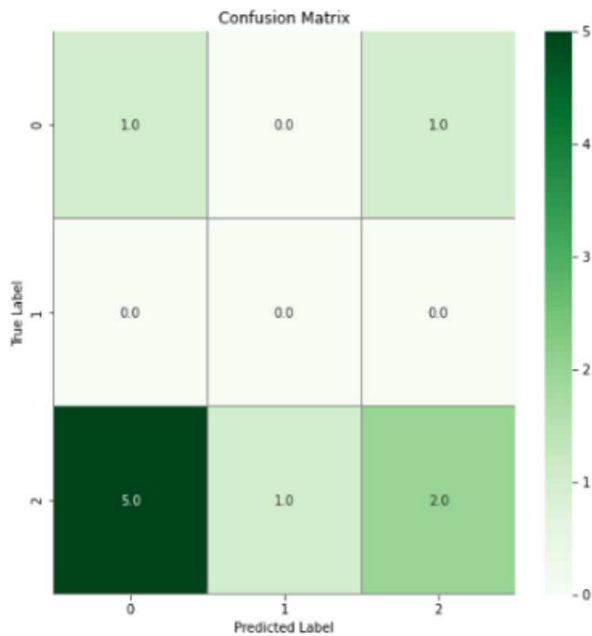


Fig. 6. Confusion matrix of persona prediction results for all subjects

accuracy, it is necessary to prepare explanatory variables that correspond to more motivational factors.

C. Improving avatars with specific suggestions

Looking at the True Label in Figure 6, most of the subjects preferred the number 2. This indicates that the subject needs an avatar to suggest specific things. However, looking at the results in Table V, few subjects showed a significant improvement in motivation and achievement in exercise by avatars who made specific suggestions. We think the reason is that the proposal was not feasible for them.

In fact, in the free descriptions of the subjects after the experiment, some of the subjects for whom no significant difference wrote that "the proposed concrete plan was unrealistic and difficult". This means that the concrete suggestions made by avatars to users need to be appropriate for the users, not just suggestions. For example, a user who likes to go outside can be suggested to take a walk, and a user who likes housework can be suggested to do housework. By incorporating information about the fields that are not difficult for each user in advance, significant differences in motivation to exercise and achievement by avatars who suggest specific things are expected to appear in more people.

Next, the means of the values for each item of the Control of Learning Beliefs, Extrinsic, Intrinsic, Self Efficacy, and Task Value of the questionnaire using the MSLQ of all subjects were calculated. The mean values were used to classify the participants into "high" and "low" groups. Based on the these groups, the relationship between the groups of subjects with significant differences in exercise motivation and achievement was examined. The results showed that there was no commonality between the groups of subjects with significant

differences in any of the items. This may also be due to the fact that there were only five explanatory variables, so it is necessary to prepare more explanatory variables.

D. Improving the interaction between the user and avatar

When we checked the content of the dialogue between the subject and avatar described in Chapter 5 from the perspective of linguistics, we found that the avatar responded to the user in a way that was out of line and that the same phrases were responded to over and over again.

In this case, the interview was conducted by one person based on the interview in 4.A. Therefore, since there were cases where different survey targets had similar dialogues, the phrases were similar and the answers were commonplace. Therefore, when adding a new response content, rather than creating a response to the content of the user's utterances by one person, multiple people come up with the response content individually. By doing this, it is possible to give responses from various people to one of her remarks by the user, which not only increases the types of responses, but also eliminates mundane responses.

One of the shortcomings of our method is that the dialogue is not conducted in a time-series. For example, if you were sick the day before, in human-to-human situations, you would worry about the other person. On the next day, taking into account the previous day's condition, we will ask, "Are you feeling better?". However, our method does not have such a function. Consider the other party's condition in time series, it is the embodiment of a more "accompanying avatars". We believe that it is necessary to consider the time series in dialogue processing. In this paper, we will discuss how to use the time series in the dialogue.

E. Study Limitation

In this experiment, we focused on the number of steps and examined the effectiveness of the goal achievement of exercise. However, in reality, exercise varies. For example, some of the subjects did exercise other than walking, such as riding a road bike on a regular basis. Such people get enough exercise because they ride road bikes, even if the number of steps is small. By taking into account items other than the number of steps, such as the amount of exercise and the time of exercise, rather than just the number of steps, we were able to verify the usefulness of the avatars. For this purpose, it is more appropriate to measure the number of steps using a smartwatch instead of using the acceleration sensor of a smartphone. Since few people wear smartwatches on a daily basis, this study mainly used a pedometer on a smartphone. But then, it is highly likely that people do not wear their smartphones during the time they are doing housework or exercising at the gym, for example. Those exercises are a very effective way to acquire the number of steps. Therefore, in order to measure the number of steps more accurately, the smartwatch should have been adapted in the experiment.

Since there were 10 subjects, the sample size was small. Therefore, there are limitations in generalization. We received

volunteers to help us with the experiment. The subjects were also asked to participate in the study as volunteers. Therefore, they varied from those who exercise on a daily because we are more interested in increasing the awareness of exercise and the amount of exercise for those who do not have exercise habits than in increasing the awareness of exercise and the amount of exercise for those who already have exercise habits.

VIII. CONCLUSION

In this paper, we proposed a method of using an avatar to accompany users to help them make daily exercise a habit, by having users interact with the avatar. The strength of this research is not to adapt the application in a uniform manner, but to assign an appropriate avatar to each individual and to present short-term goals on a daily basis.

We investigated what kind of avatar is needed through interaction with the research avatar. As a result of the survey, we found that some personas in the research group preferred to receive specific suggestions to help them make daily exercise a habit, while others preferred to be sympathetic to the speech.

We also conducted an experiment to test the effectiveness of avatars. The results showed that the use of the avatar did not significantly improve the subjects' motivation to exercise, but it significantly improve their achievement of their daily exercise goals.

The use of avatars that empathize with the user's speech and suggest specific things to do was found to significantly improve the achievement of daily exercise goals. On the other hand, motivation to exercise was not significantly improved. However, some of the subjects significantly increased their motivation to exercise.

In the future, the avatars need to be improved by increasing the number of explanatory variables in the regression equation for predicting appropriate avatars, and by having multiple people think about and add the response content of what the user is expected to say, rather than having one person create it. In addition, it is necessary to collect information on feasible motions from the user in advance and consider specific proposals based on this information. Furthermore, since the dialogue was not conducted in time-series data, it is necessary to process the avatar's responses in time-series.

REFERENCES

- [1] Arora, Charu, et al. Development and validation of health education tools and evaluation questionnaires for improving patient care in lifestyle related diseases. *Journal of clinical and diagnostic research: JCDR*, 2017, 11.5: JE06.
- [2] Manson, Joann E., et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *The Lancet*, 1991, 338.8770: 774-778.
- [3] Paffenbarger Jr, Ralph S., et al. The association of changes in physical-activity level and other lifestyle characteristics with mortality among men. *New England journal of medicine*, 1993, 328.8: 538-545.
- [4] Stawarz, Katarzyna; COX, Anna L.; BLANDFORD, Ann. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015. p. 2653-2662.
- [5] World Health Organization. *Global action plan on physical activity 2018-2030: more active people for a healthier world*. World Health Organization, 2019.
- [6] Pintrich, Paul R., et al. *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. 1991.
- [7] Takamura, Hiroya; Inui, Takashi; Okumura, Manabu. Extracting semantic orientations of words using spin model. In: *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*. 2005. p. 133-140.
- [8] Locke, Edwin A.; Latham, Gary P. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 2002, 57.9: 705.
- [9] Lally, Phillippa, et al. How are habits formed: Modelling habit formation in the real world. *European journal of social psychology*, 2010, 40.6: 998-1009.
- [10] Kaushal, Navin; Rhodes, Ryan E. Exercise habit formation in new gym members: a longitudinal study. *Journal of Behavioral Medicine*, 2015, 38.4: 652-663.
- [11] Mazzotta, Irene; De Rosis, Fiorella; Carofiglio, Valeria. Portia: A user-adapted persuasion system in the healthy-eating domain. *IEEE Intelligent systems*, 2007, 22.6: 42-51.
- [12] Carnevale, Peter JD; Isen, Alice M. The influence of positive affect and visual access on the discovery of integrative solutions in bilateral negotiation. *Organizational behavior and human decision Processes*, 1986, 37.1: 1-13.
- [13] Forgas, Joseph P. On feeling good and getting your way: Mood effects on negotiator cognition and bargaining strategies. *Journal of personality and social psychology*, 1998, 74.3: 565.
- [14] Reeves, Byron; Nass, Clifford. *The media equation: How people treat computers, television, and new media like real people*. Cambridge, UK: Cambridge university press, 1996.
- [15] Fitzpatrick, Kathleen Kara; Darcy, Alison; Vierhile, Molly. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health*, 2017, 4.2: e19.
- [16] Morris, Jeremy N.; Hardman, Adrienne E. Walking to health. *Sports medicine*, 1997, 23.5: 306-332.
- [17] Ogilvie, David, et al. Interventions to promote walking: systematic review. *bmj*, 2007, 334.7605: 1204.
- [18] <https://www.mhlw.go.jp/content/10900000/000687163.pdf>
- [19] LINE, LINE | always at your side. <https://line.me/en/>.
- [20] Zoom, Zoom Meetings - Zoom. <https://zoom.us/meetings>.
- [21] Telegram, Telegram Messenger. <https://telegram.org/>.
- [22] Elasticsearch, Elasticsearch: The Official Distributed Search & Analytics Engine | Elastic. <https://www.elastic.co/elasticsearch/>.

[1] Arora, Charu, et al. Development and validation of health education tools and evaluation questionnaires for improving patient care in lifestyle