

Construction and Evaluation of a System to Identify People's Personality and Mental State Based on Dialogues with Chatbots

Ryohei Nakatsu Ten Suzuki Naoko Tosa Takashi Kusumi
Kyoto University Kyoto University Kyoto University Kyoto University
Nakatsu.ryohei@gmail.com stuzeukni0901@icloud.com tosa.naoko.5c@kyoto-u.ac.jp kusumi.takashi.7u@kyoto-u.ac.jp

Abstract

One method of identifying a person's personality and mental state is the TAT (Thematic Appreciation Test). The TAT is a method of having a person create stories based on presented images, and it can express a person's unconscious. However, TAT is not easy to use because it is based on a conversation between a person and an expert, and therefore, it has not been used much. We have developed an AI-TAT system that uses AI to replace experts and can identify a person's personality and mental state through dialogue with the AI. We also evaluated the system using an online experiment with 30 subjects and obtained promising results. In addition, we compared the use of still images and videos as TAT diagrams and verified the effect of repeated use.

Keywords: personality, mental state, TAT, psychological evaluation

1 Introduction

For people to live healthy lives, they need to keep their bodies and minds healthy. Regarding physical health, many measures have been developed and prepared to maintain a healthy body. While various systems are in place to quickly check physical health daily and provide advice on maintaining and improving it, they are not yet entirely in place for mental health. Therefore, even if a person has a mental illness, it is not taken care of until it becomes a serious illness, such as depression or criminal behavior.

When we were children, we all dreamed about who we want to be and what we want to accomplish. However, as we grow up, we must compromise our ideas and arguments with those around us. For example, if one works for a company, it becomes necessary to prioritize its rules and profits, and asserting and implementing one's opinions and ideas becomes difficult. As a result, the dreams and hopes one had gradually drove from the conscious to the unconscious. However, such unconsciously relegated dreams and hopes of one's own sometimes surface. Many people, therefore, have a sense of "what am I doing?" or "what I am today is not my true self," which is thought to lead to problems.

There are many ways to address such concerns. One such method is to take time to reflect on oneself, which leads to mindfulness [1]. This is why mindfulness has been adopted by many people in developed countries and cities. A purer form of mindfulness would be Zazen [2], which is used in Zen. In Zen, the ideal self and the present self are contrasted, and it is said that the ultimate form of this contrast is "enlightenment." For this reason, Zen uses the Ten Ox-Herding Pictures, which represent the ten states that lead to enlightenment [3].

Another way to deal with problems is to know one's personality [4][5] and think and act appropriately based on that personality. However, personality judgments and guidelines are often used to heal the minds of people who are

already unbalanced. Just as it is easy to check the health of the body, is it possible to construct a system to check the health of the mind quickly?

This research aims to develop such a mental health management system. We describe a system in which an AI easily checks a person's personality and mental status by interacting with an AI chatbot [6] and giving advice. We begin with a general description of personality testing methods. In particular, the personality testing method called TAT (Thematic Appreciation Test) [7] will be described. Next, we will describe the three chatbots (Consulting AI, Absurd AI, and Zen AI), and a personality judgment system (AI-TAT) configuration that combines the TAT and these chatbots. We then describe the online evaluation experiment in which participants evaluated these systems.

2 Personality Assessment

One method of measuring mental health is to determine an individual's personality. In this Chapter, we will briefly explain the Big Five, the Rorschach method, and the Baum test, which are well-known methods for personality assessment. The TAT used in this research will be explained in Chapter 3.

2.1 Big Five

The Big Five is a theory on the compartmentalization of personality. It states that a person's personality comprises five dimensions: openness, honesty, diplomacy, cooperativeness, and neurotic tendencies, which psychologist Lewis R. Goldberg proposed in the 1990s [8]. It has been widely known and used to successfully explain human traits and behavior. However, it is difficult to extract the mental problems of an individual, especially unconscious problems, by simply trying to explain each person's personality using the five factors.

2.2 Rorschach method

The Rorschach method was invented in 1921 by the Swiss psychiatrist Hermann Rorschach [9]. By dropping ink on a piece of paper and folding it in two, a nearly symmetrical image is created and shown to the subject. The answers to what the image looks like are analyzed to discover the subject's unconscious state of mind and any problems he or she may have. A set of ten Rorschach illustrations is used. The Rorschach method is a personality evaluation method called the projective method, which takes advantage of the tendency of people to project their unconscious thoughts and desires onto figures.

2.3 Baum test

The Baum test is another projective method, a simple test in which subjects are asked to draw a single fruiting tree [10]. It was proposed by Swiss psychologist K. Koch in 1945. The test is intended to learn about the subject's thinking habits, inner feelings, and depth of mind from the overall impression, the form of the tree, and its position. It is widely used because it is easy to perform, but it requires an expert with experience to decipher the painted picture.

3 Personality Assessment Method: TAT

3.1 What is TAT?

The TAT (Thematic Appreciation Test) is a personality assessment method proposed by Henry Alexander Murray in 1943 and is one of the projective methods [7]. Twenty sets of illustrations (some illustrations are prepared for men, women, children, and adults, making a total of 31 illustrations) are shown to the subjects in turn, and the subjects are asked to write a narrative of what they perceive from the illustrations. The stories created in response to the illustrations are used to learn about the subject's inner life and even unconsciousness. The TAT illustrations are figurative and suggest a crisis. In addition, they depict a person (the protagonist) with whom the subject can easily empathize and identify. Therefore, by looking at TAT illustrations, subjects are likely to create a life story, as if they were reliving what has happened from childhood to the present. Moreover, in that story, the subject's unconscious feelings are expressed.

However, because each TAT illustration has such power, it gives a highly dark sense. It is difficult for a healthy person to draw a story using TAT illustrations. Another problem is that many TAT illustrations require considerable time to determine personality. Therefore, the TAT is used less frequently than the Rorschach method, and the analytical method for the TAT has not been established.

However, TAT has excellent potential as a personality assessment method. Therefore, we attempt to construct a system that can easily perform personality assessments using TAT and AI. The following new aspects can be added to TAT by combining TAT with AI.

- (1) Generate new TAT illustrations using a generative AI.

The standard illustrations for the TAT are still those in the Harvard Edition manual, published by Murray in 1943 by Harvard University Press [7]. The original illustrations can compel one to become emotionally involved with the characters in the illustrations. However, this makes it difficult for the general public to use them comfortably.

Using the recently developed generative AI [11][12], we tried to create new TAT images based on the original images by improving the extremely dark impression of the original images. In addition, since the AI can generate moving images, it is possible to compare the use of moving images with the

use of still images.

- (2) AI replaces the role of inspectors

The TAT is a one-on-one, face-to-face session between a psychologist and a subject. However, this requires much time and a large workforce and is not something the general public can comfortably use. We aim to develop a tool to check mental health daily with the same ease as checking physical health daily. To this end, it is desirable to check mental health through dialogue with an AI, rather than with a specialist.

3.2 Legitimacy of developing AI-TAT

This study was conducted with the active participation of Professor Takashi Kusumi, a specialist in educational psychology at Kyoto University, who provided guidance throughout the planning and implementation phases. He also reviewed the TAT image generation process and confirmed the plausibility of the evaluation results. In addition, he has been added as a co-author.

4 Chatbot Used for AI-TAT

Since AI-TAT has the following requirements, we constructed suitable chatbots using Chat GPT API.

- (1) Present TAT illustrations to facilitate conversations with subjects smoothly and create an atmosphere in which they are comfortable creating stories.
- (2) The stories created are analyzed, and subjects are asked more in-depth questions. The goal is to make the subjects aware of their mind state by bringing to light unconscious problems they may have.

4.1 Consulting AI

4.1.1 Concept

In today's busy world, many people are so busy with their work that they do not have time to reflect on themselves. Their dreams and hopes when they were young, such as whom they wanted to become and what they wanted to accomplish, are gradually forgotten and pushed into the unconscious due to busyness. But they occasionally come to the forefront of their minds. This is thought to lead to anxiety and other feelings. It is desirable to have an AI that can clarify such vague concerns through conversation with it. We have named such a conversational AI "Consulting AI" and are developing it.

4.1.2 Construction of Consulting AI

The basic structure is to use the Chat GPT API to build the Consulting AI as an interface system between Chat GPT and the user (Fig. 1).

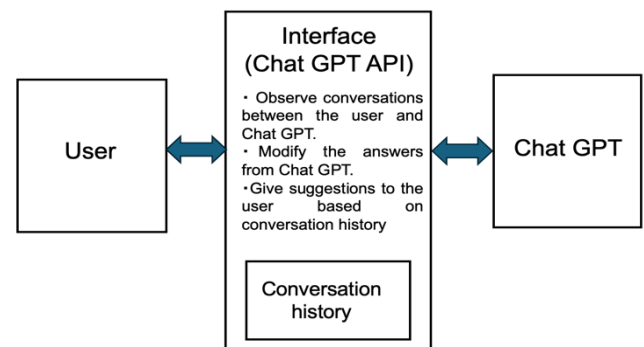


Fig. 1 Basic structure of the Consulting AI

To realize the Consulting AI, we have repeatedly conversed with Chat GPT and studied the structure of the Chat GPT API.

We found that adding the following functions is appropriate.

- (1) Using the Chat GPT API, we aimed to coordinate the conversation between the user and Chat GPT and elicit the user's current concerns and goals through the conversation.
- (2) We developed a function that stores the conversation history, creates a summary based on several past conversations, combines it with the user's input, and inputs it to the Chat GPT to obtain an appropriate response.
- (3) Instead of passing Chat GPT responses directly to the user, we developed a function to make adjustments (e.g., adjust length, tone, etc.) before outputting them.
- (4) A function was developed to calculate the distance from the previous conversation content and have Chat GPT generate a response again if it is not within the appropriate range.

The specific process flow is shown in Fig. 2.

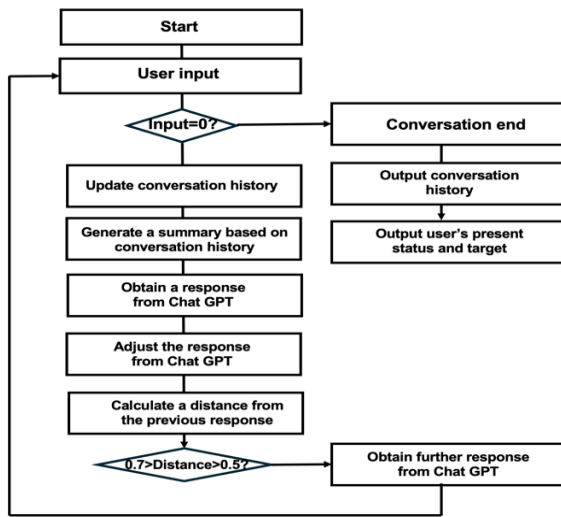


Fig. 2 Processing flow of the Consulting AI

4.2 Absurd AI

4.2.1 Concept

Conversations with the Consulting AI are, in a sense, like conversations with a counselor and may cause the user stress. Therefore, we prepared a chatbot that chats lightly between conversations with the Consulting AI. However, users are likely to get bored with mere chats. We came up with the idea of an "Absurd AI" that engages in absurd conversations.

4.2.2 Absurd conversation

Absurdity refers to being irrational or contrary to common sense. Because absurdity brings about the effect of nonsense, absurd developments are often used in specific literary works. Representative absurd authors include Eugène Ionesco [13] and Samuel Beckett [14]. Conversations with the absurd AI may ignore logic and leaps in conversational flow, making the AI's responses unpredictable from the user's point of view.

4.2.3 Construction of Absurd AI

As basic data for the absurd conversation, we have taken Samuel Beckett's "Waiting for Godot" [15], which is considered one of the most famous works of absurdist theater. The entire basis is a conversation between two characters,

Vladimir and Estragon, who are waiting for Godot. The conversation takes place in a truly absurd setting where the characters themselves do not know why they are waiting for Godot, a person they do not know. The conversation is not logical, has no story, and is absurd.

We constructed Absurd AI using the following procedure.

- (1) Digitization of "Waiting for Godot": The entire "Waiting for Godot" was digitized.
- (2) Training data: The entire "Waiting for Godot" is a play in two acts, of which the first act is used as data for the AI to learn.
- (3) Creation of conversation data for training: The entire "Waiting for Godot" is a series of short lines. To treat this as conversation data, the sequence of lines "ABCDE..." was made into the form (A, B) (B, C) (C, D) (D, E)...., and the data was created as a set of lines and the lines of answers to them.
- (4) Fine-tuning of Chat GPT: Open AI offers a service to fine-tune [16] Chat GPTs. This is useful as a mechanism to quickly tune to new training data while retaining the functionality of the Chat GPT as a whole (Fig.3).

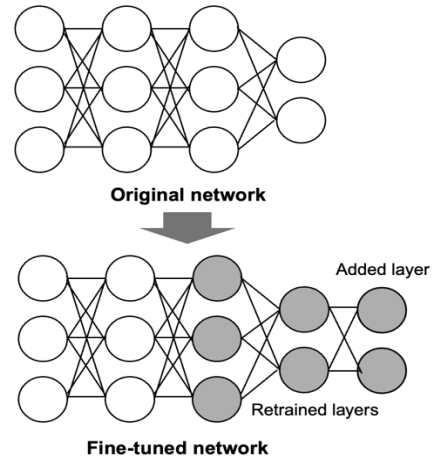


Fig. 3 Conceptual diagram of fine-tuning

4.3 Zen AI

4.3.1 Concept of Zen AI

Zen question-and-answer is a method for subjects to know their state of mind and solve their problems through awareness. We have developed Zen AI as a chatbot that can perform Zen question-and-answer sessions.

Zen monasticism is an exchange of questions and answers between a Zen master and a Zen practitioner [17]. The practitioner aims to reach enlightenment through illogical questions and leaps of logic, called "Koans" [18], by the Zen master. For example, a well-known example of a Koan is the question, "What is a one-handed clap?" Because such questions are illogical, they have been regarded as an area in which conversational AI, such as Chat GPT, does not excel. However, Zen questions are possible for AI for the following reasons.

- (1) Though illogical, Zen questions have logic at the bottom of them.
- (2) "What is a one-handed clap" is a Zen question that breaks with the logic of two hands clapping at the bottom of the question.
- (3) If it is possible to change a logical question to an illogical question, then Zen questions are considered possible for

AI.

- (4) Considering that the practitioner reaches enlightenment and AI's role is to provide impetus, Zen questions and answers are possible for AI.

4.3.2 Building Zen AI

Zen AI was achieved by using the Chat GPT API and prompts to make the output of Chat GPT look like the following.

- (1) Sentence simplification: By making Chat GPT's long logical sentences extremely short, the connections between words fade and appear illogical.
- (2) Tone change: By setting the tone of the words to that of a Zen master, the question and answer become more like a Zen question and answer.

5 AI-TAT, a Personality Judging System Combining TAT and Chatbots

By combining the newly generated TAT illustrations with the chatbots described above, we constructed the AI-TAT. This system allows subjects to perform the TAT test without an examiner and analyze their personalities. Its processing flow is shown in Fig. 4.

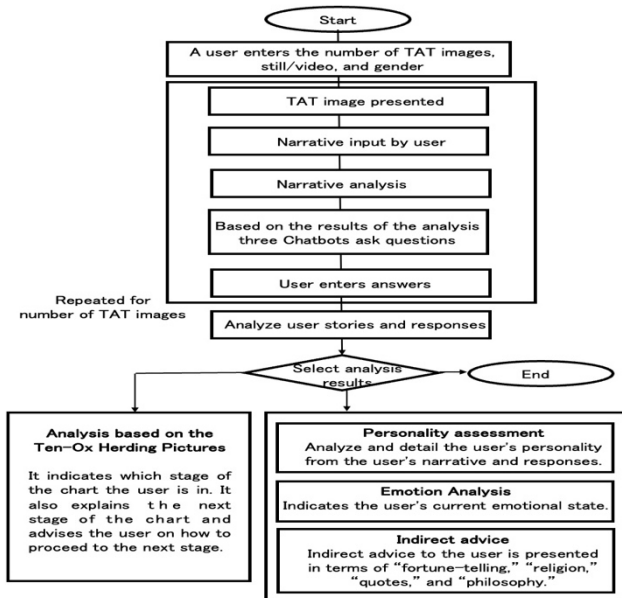


Fig. 4 Schematic processing flow of AI-TAT

5.1 TAT images to be used

The subjects are asked to select the number of TAT illustrations to be used for story generation, are presented with the illustrations in sequence, and are asked to generate a story for each illustration. The system to be used by the general public daily should have fewer cards than the original number. Hartmann proposed a basic set of 8 cards [19], which many clinicians use. Therefore, we used this basic set. The appropriateness of a smaller basic set has not yet been proven, but we did so because even eight illustrations seem too many for today's busy people. The number of cards will become clearer through repeated evaluation experiments of the AI-TAT.

The TAT illustrations to be used were generated using a generative AI, but using the original as a reference. Two types of TAT illustrations were generated: still and moving images. Examples of the still image version are shown in Fig. 5.

In addition to still images, a video TAT was generated.

Although the video is only about 5 seconds long, the significance of using video for TAT is as follows.

- (1) Since the younger generation is familiar with video, the use of video will make AI-TAT more familiar to the younger generation and easier to use.
- (2) While still images tend to force the viewer to become emotionally involved, moving images can be viewed more easily than still images because the people in them appear to be moving independently and with life, allowing the viewer to keep a little distance from them.



Fig. 5 Examples of TAT images produced using generative AI

5.2 Role of chatbots

The three types of chatbots described in Chapter 4 ask questions based on each character regarding the story input by the subject. In other words, they ask the subject about the deeper meaning of the subject's story, thereby confirming the subject's innermost or unconscious thoughts that the subject has expressed without realizing it. While the Consulting AI tends to ask direct questions, the Absurd AI and Zen AI ask indirect and abstract questions. It is hoped that subjects will become aware of the unconscious thoughts expressed in the stories they create.

5.3 Visualization of the subject's emotion

The stories created by the subjects are entered into Chat GPT for analysis, and the subjects' emotions are expressed in terms of four emotions: "joy," "sadness," "anger," and "anxiety." By expressing this as a time-series graph, subjects can understand changes in their emotions, or in other words, changes in their mental state. An example of an emotion change graph is shown in Fig. 6.

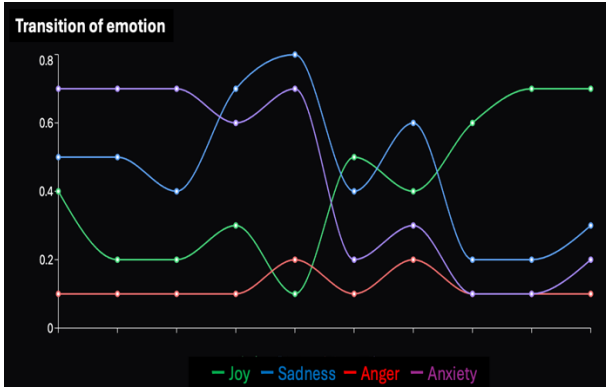


Fig. 6 An example of a time series graph of emotions

5.4 Personality analysis

Since conventional TAT is not often used as a personality analysis method, the analytical method has not yet been established, and the analyst's experience and intuition were used in many cases. In our AI-TAT, after the presentation of all TAT illustrations, the subject's narrative input, the input of questions from the chatbot, and the subject's answers were input together into Chat GPT to request a personality analysis of the subject and the output of Chat GPT is presented to the subject as the result of the personality analysis.

5.5 Presentation of indirect advice

Chat GPT personality analysis tends to be lengthy, analytical, and logical. The participants are not likely to accept them as concrete advice. Therefore, instead of giving direct advice, we decided to give indirect advice, such as the following, which the participants would enjoy.

- (1) Fortune-telling: Providing advice in the form of fortune-telling.
- (2) Proverbs: Providing advice selected from well-known proverbs.
- (3) Philosophers' words: Providing advice selected from well-known philosophical quotes.

5.6 Personality analysis using the Ten Ox-Herding Pictures

As a straightforward indication of the subject's current mind and the path to be taken in the future, we adopted the Ten Ox-Herding Pictures [3]. The pictures are based on the concept that there are two selves: the self that one should aspire to be and the self that is compromised with others in the real world. Moreover, the final destination of the mind is to recover the true self, in other words, enlightenment. A cow represents the true self, and the process of finding the lost cow, taming it, and uniting it with oneself is the process of attaining enlightenment, which is represented by the ten pictures.

Based on the story the subject has created, the subject's position on these pictures may be more acceptable to people than the analytical and logical personality analysis results that Chat GPT tends to produce. Also, since there are many books on the Ten Ox-Herding Pictures, the system does not need to advise the subject on what it means to be at a particular stage and how to go to the next stage, leaving it to the subject's initiative.

Although it is inherently difficult to determine which stage of the ten pictures a subject is in, we adopted a simple method: input the subject's story into the Chat GPS, let it make a judgment, and output it.

5.7 Examples of dialogue with AI-TAT

Several examples of dialogues with AI-TAT are shown below.

Dialogue 1 (Subject viewed the upper-left image in Fig. 5):

Subject: A boy is looking at a baseball glove and seems worried. "Do I have no talent for baseball? Even if I practice, I don't think I'll ever become a regular player."

Consulting AI: Please tell me about the boy's concerns and your relationship to him.

Subject: This boy might be me when I was young.

Zen AI: Where do your concerns come from?

Subject: Perhaps they come from my heart.

Absurd AI: Concerns? Such things don't exist.

Subject: I see, that might be true.

Dialogue 2 (Subject viewed the center image in Fig. 5):

Subject: People are busy with rice planting. A woman who came from the city seems confused.

Consulting AI: What is she thinking?

Subject: She seems to feel that this isn't her place. I sometimes feel that way too.

Absurd AI: Rice planting is fun.

Subject: I don't have the leisure to think that way yet.

Zen AI: Where is your place?

Subject: Where is the place I should be?

6 AI-TAT Evaluation Experiment

We have evaluated the system. As this system utilizes the Chat GPT API in various places, evaluating it from the following points is necessary.

- (1) Can we assume that this system can check the state of mental health?
- (2) Can the system be used easily by the general public?
- (3) Which is more appropriate TAT illustrations, still or moving images?

6.1 Evaluation method

The experiment was conducted online, with two TAT illustrations, one using still images (AI-TAT1) and the other using moving images (AI-TAT2), accessible from different URLs. The number of used TAT illustrations was set to 5 in both cases. Each subject was asked to access AI-TAT1 and AI-TAT2 two times for each system on different days to reset his or her mind state.

6.2 Test subjects

The target population of this system is assumed to be healthy individuals who can use online tools, particularly university students or young adults with basic digital literacy. The subjects were Japanese undergraduate students at Kyoto University, recruited through the Kyoto University Co-operative Association website. To ensure no bias in the sample, the recruitment notice simply stated, "We are seeking participants for an experiment involving online personality and psychological state tests," without providing further details about the experiment. Regarding digital literacy, the only requirement was that participants had experience using Google Forms; no further skills were required. A total of 37 students applied, but some withdrew during the experiment or had incomplete responses in Google Forms, resulting in a final sample of 30 participants (18 males and 12 females), aged approximately 18 to 24. The participants covered a wide range of departments at Kyoto University, including both science and liberal arts, with no significant biases observed.

6.3 Ethical and Privacy Issues

The purpose and procedures of the experiment were explained to all participants in advance, and written informed consent was obtained. To protect personal information, participants were informed that their data would be stored and analyzed only after anonymization.

To minimize emotional impact, participants were told that the AI-generated responses were not clinical diagnoses, but conversational simulations designed for research. They were also allowed to discontinue participation at any time during the experiment.

To prevent misuse of the system as a mental health assessment tool, we clarified in the study documentation and experiment interface that the system is not intended for clinical use. Its purpose is to explore narrative interactions and self-reflection, and any clinical interpretation must be left to professionals.

6.4 Evaluation items

6.4.1 Impression of TAT images (or videos)

- (1) Did the TAT images (or videos) stimulate your emotions?
- (2) Did the TAT images (or videos) stimulate your imagination?

6.4.2 Narrative generation experience

- (3) Was story generation fun and easy?
- (4) Did you feel burdened or stressed about the story?

6.4.3 Unconscious representation of narrative

- (5) Do you think the story you created reflects your unconscious thoughts and feelings?
- (6) Have you noticed new aspects of yourself through story generation?

6.4.4 Effectiveness of chatbot questions

- (7) Did the chatbot's (Consulting AI, Absurd AI, Zen AI) questions cause you to think more deeply about your story?
- (8) Did each chatbot's approach offer a different perspective?

6.4.5 Evaluation of the characterization process

Evaluation of Emotional Analysis

- (9) Was the visualization of emotional changes easy to understand?
- (10) Did the AI's emotion estimates match actual feelings?

Personality Analysis Assessment

- (11) Was the personality analysis convincing?
- (12) Did the analysis provide new perspectives on your ideas and values?

Evaluation of advice

- (13) Did you find the advice (fortune-telling, philosophy, proverb, etc.) helpful?
- (14) Did you feel the advice was appropriate for your situation?

Evaluation of user experience

- (15) Was the flow of feedback (emotion -> personality -> advice) smooth?
- (16) Did it help you to know your state of mind?

6.4.6 Evaluating the process of the Ten Ox-Herding Pictures

Recognition and understanding of the Ten Ox-Herding Pictures

- (17) Did you know about the Ten Ox-Herding Pictures?
- (18) Did this experience help you understand the pictures?

The benefit of presenting the stages of the Ten Ox-Herding Pictures

- (19) Have you benefited from being shown which stage of the pictures you are in?
- (20) Are you satisfied with the judgment of which stage of the pictures you are in?

User self-growth and continuity of experience

- (21) Would you like to move on to the next stage of the pictures?
- (22) Will it be enjoyable as a pure game or entertainment?

7 Evaluation Results and Discussion

7.1 Results per evaluation item

Figure 7 shows the results for each evaluation item described in 6.3.

All evaluation items except Q4 and Q17 received a higher value than the midpoint value 3. The low value for Q4 indicates that the subjects did not feel stressed.

Q2, Q5, Q11, and Q16 have ratings of 3.5 or higher, indicating that subjects felt that the TAT image (or video) stimulated them to create a story that reflects their unconscious mind. As a result, their personality and current state of mind are correctly judged.

Q15 received a rating value of nearly 4, indicating that the presentation of information to the user, such as emotional changes, personality judgments, and advice, was rated as appropriate.

Even though most subjects were unfamiliar with the Ten Ox-Herding Pictures (Q17 was less than 2), they did not feel uncomfortable with its display of the state of mind (Q18-Q21 received a rating value of nearly 3.5).

Q22 received a rating of 3.5 or higher, indicating that AI-TAT could be expanded into games and other applications.

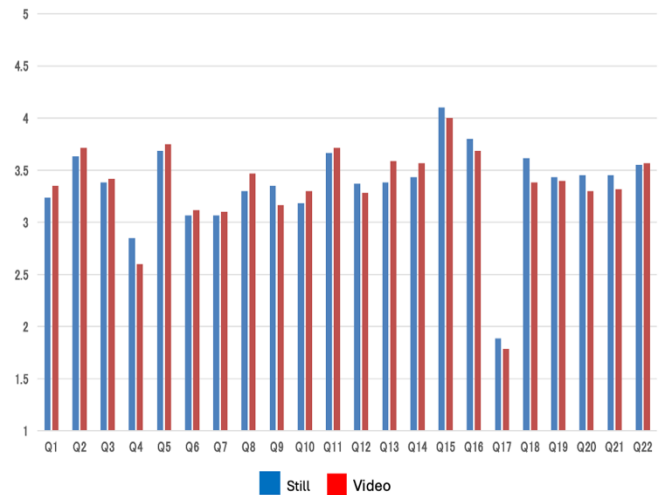


Fig. 7 Evaluation results by evaluation item

7.2 Analysis of questionnaire response trends

To address the concern of central tendency bias often observed in Likert-scale evaluations, we calculated the standard deviation (SD) of the scores. The results were SD = 1.137 for the still image condition and SD = 1.170 for the video condition, suggesting a relatively broad distribution of responses and indicating that participants did not consistently choose neutral values.

In addition, we created Fig. 8, which shows the top five participants' average scores. These scores are significantly higher than the overall averages shown in Fig. 7, further supporting the interpretation that the evaluations were not concentrated around the median but were broadly distributed.

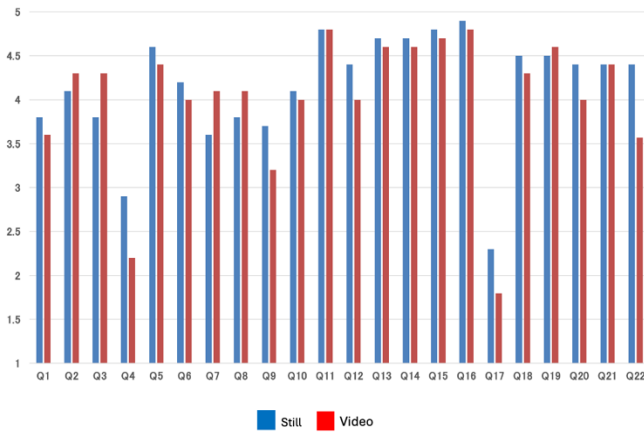


Fig. 8 Evaluation results by evaluation item for the top five subjects.

7.3 Verification of significant differences between still and moving images

This experiment generated a new TAT diagram using the generative AI based on the original TAT diagram. One is a still image similar to the original TAT diagram. The other is a 5-second movie based on the still image. We wanted to know how moving images would affect the results.

Figure 7 likely indicates little difference between still and moving images. The next step was to verify whether the difference was significant. The verification was conducted using the analysis of variance (ANOVA). Since there were 22 evaluation items, a factor analysis was first conducted to reduce the number of items.

The factor analysis results show that the 22 evaluation items can be grouped into three groups. Then, the original evaluation results for still and moving images were averaged for these three groups, and ANOVA was performed (Fig. 9).

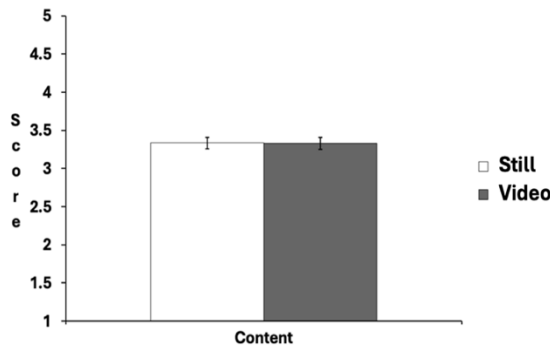


Fig. 9 ANOVA results for comparison of still and video images

As Fig. 8 shows, ANOVA using the three grouped evaluation items revealed no significant difference between still and moving images ($p=.866$). This suggests that video for TAT can be used and opens up new possibilities.

7.4 Verification of the effect of repetition

Next, we tested the effect of repeating the experiment using the same TAT diagram. For the AI-TAT to repeatedly examine

the state of mind daily, repetition should have no effect. Therefore, we asked participants to repeat the experiment using still and moving images twice, leaving at least one day between each, and compared the results of the first and second tests. As there is no significant difference between still and moving images, we only show the results for still images.

Figure 10 compares the first and second results using still images.

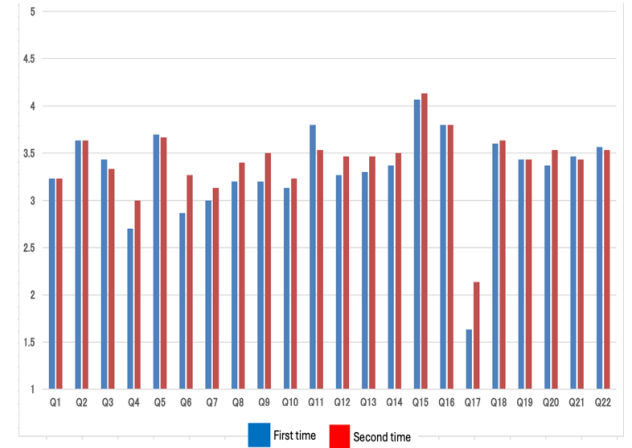


Fig. 10 Comparison between the first and the second experiment (still image)

Then the factor analysis was conducted showing that 22 evaluation items can be grouped into three groups. Then, the original evaluation results for first and second trials were averaged for these three groups, and ANOVA. was performed. Figure 11 shows the results of the ANOVA.

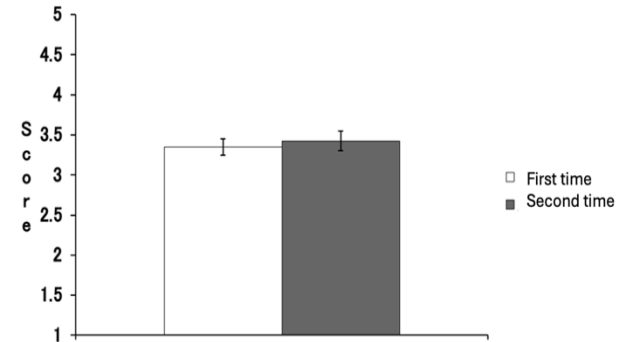


Fig. 11 ANOVA results for the comparison between the first and the second experiments (still image)

The ANOVA result shows no significant difference between the first and the second for still images ($p=.199$). From these results, the following can be said.

7.5 Discussion

7.5.1 Comparison of still and moving images

No significant difference existed between the experimental results using a still image TAT and those using a moving image TAT. This indicates that either can be used. However, since the video is only about 5 seconds long, the effect of a longer video should be verified in the future.

7.5.2 Effect of repeated use

The results of the first and second experiments, separated by at

least one day, were similar for still images. This suggests the possibility of using repeated TAT experiments with this AI-TAT. However, this was a two-repetition study, and it will be necessary to confirm the effect of more repetitions in the future.

7.5.3 AI-TAT against existing validated instruments

At this stage, the AI-TAT system has not been directly compared with existing validated psychological instruments such as the Big Five Inventory (BFI) or Minnesota Multiphasic Personality Inventory (MMPI). However, it is important to note that AI-TAT is designed to serve a different function from these structured self-report questionnaires. While BFI and MMPI aim to quantify personality traits using predefined dimensions, AI-TAT is based on projective techniques that emphasize open-ended expression and narrative generation.

Given this difference, the AI-TAT system is expected to provide complementary insights, especially in capturing subtle, context-dependent aspects of personality or psychological state that may not be easily detected through fixed-choice instruments. In future studies, we plan to administer BFI or similar validated instruments alongside AI-TAT and investigate correlations and divergences between the results. This will help clarify the unique contributions and limitations of AI-TAT and establish its credibility more firmly through triangulated evaluation.

8 Conclusion

People must maintain their mental and physical health to spend their days in good health. While there are many ways to check physical health daily, no method exists to check the mental state daily.

This paper describes the concept and content of AI-TAT, which introduces AI into the TAT method of checking a person's personality and mental state while interacting with the AI daily and allowing the system to advise the person as needed. In addition, we described the results of the evaluation of the prototype system using 30 subjects. The following results were obtained.

- (1) Relatively high evaluation results were obtained for most of the 22 evaluation items. This suggests that this system can be used by the general public to check their mental health daily, as was our initial goal.
- (2) There was no significant difference between the evaluation results of AI-TAT using still and moving images. This suggests that introducing moving images may open up new possibilities for TAT based on conventional still images.
- (3) When the subjects were asked to repeat the experiment twice, there was no significant difference between the results of the first and second experiments. This suggests that this system has the potential to be improved into a system that can be used repeatedly daily.

In future iterations of the system, we plan to incorporate a user feedback mechanism that enables participants to indicate whether the AI's responses were meaningful, helpful, or emotionally engaging. This feedback would serve as a basis for refining the AI's dialogue strategies, allowing the system to adapt more effectively to users' needs and preferences. We believe that integrating such a feedback loop will enhance user experience, promote more personalized interactions, and support the broader applicability of the system in various real-world contexts.

Acknowledgment

This study is based on joint research between Kyoto University and Toyota Motor Corporation. The authors thank Mr. Keisuke Yasuda of Toyota Motor Corporation.

References

- [1] Jon Kabat-Zinn, "Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life," Balance (2023).
- [2] Shunryu Suzuki, "Zen Mind, Beginner's Mind: 50th Anniversary Edition," Shambhara; Anniversary edition (2020).
- [3] Zensho W Kopp, "The ZEN Ox-herding Pictures: Following the Path to Enlightenment," Books on Demand (2021).
- [4] Jerry Burger, "Personality," Cengage Learning (2018).
- [5] Daniel Nettle, "Personality: What Makes You the Way You Are," Oxford University Press; 1st edition (2009).
- [6] Robert Ciesla, "The Book of Chatbots: From ELIZA to Chat GPT," Springer, 1st edition (2009).
- [7] Henry Murray, Christina Morgan, "Thematic Appreciation Test: Manual," Harvard University Press (1943).
- [8] Lewis R. Goldberg, "The Development of Markers for the Big-Five Factor Structure," Psychological Assessment, Vol. 4, No. 1 Vol. 4, No. 1, pp.26-42 (1992).
- [9] Taylor L. Shantz, John Corbyn Cravero, Morgan Newman, "The Legacy of Hermann Rorschach and His Inkblots: Portrait of a Man or Mental Disease?" Cureus (open access), Vol. 16, No. 8 (2024).
- [10] Michelangelo S. Maserati, et. al, "The Tree-Drawing Test (Koch's Baum Test): A Useful Aid to Diagnose Cognitive Impairment," Behavioral Neurology, Vol. 2015, Article ID: 534681 (2015).
- [11] David Foster, "Generative Deep Learning: Teaching Machines To Paint, Write, Compose, and Play," O'Reilly Media; 2nd edition (2023).
- [12] Mark Liu, "Learn Generative AI with PyTorch," Manning (2024).
- [13] Eugène Ionesco, https://en.wikipedia.org/wiki/Eugène_Ionesco
- [14] Samuel Beckett, https://en.wikipedia.org/wiki/Samuel_Beckett
- [15] Samuel Beckett, "Waiting for Godot: A Tragicomedy Two Acts," Grove Press (2011)
- [16] Fine-tuning, [https://en.wikipedia.org/wiki/Fine-tuning_\(deep_learning\)](https://en.wikipedia.org/wiki/Fine-tuning_(deep_learning))
- [17] Thomas Cleary, "Book of Serenity: One Hundred Zen Dialogues," Shambhara; 1st Paperback Edition (2005).
- [18] Koan, <https://en.wikipedia.org/wiki/Koan>
- [19] A. Arthur Hartman, "A Basic TAT Set," Journal of Projective Techniques and Personality Assessment, Vol. 34, No. 34, pp. 391- 396 (1970).