International Journal of Social Robotics (2025) 17:1003–1017
https://doi.org/10.1007/s12369-025-01258-x

Effectiveness of Conversational Robots Capable of Estimating and Modeling User Values

Yuki Sakamoto[1] · Takahisa Uchida[1] · Hiroshi Ishiguro[1]

Accepted: 6 April 2025 / Published online: 15 May 2025

Abstract
Personalizing a dialogue system according to the user has been recognized to have various positive effects. Despite the significance of user values, concepts guiding choices and evaluations being recognized in communication, they have not been considered in personalized dialogue systems. Therefore, this study constructs a dialogue system that understands user values through conversation. Furthermore, the impact of understanding values on the interactions between dialogue systems and users is examined. The method is organized with a user model of preferences and values based on the established means-end chain model. We used a large language model (LLM) to estimate values based on the users' preferences and the reasons they prefer them. Furthermore, an infinite relational model (IRM) estimates the relationships between multiple elements within the user model. The experiments show that the proposed method could estimate user values and enhance animacy and perceived intelligence in users' impressions of an android robot, prompting new insights into users' own values. The perception of the robot contributes to improved-quality interactions, and new insights into values facilitate a deeper self-understanding of users. This achievement, demonstrating the effects of using values for interaction, can provide valuable insights into human-robot interaction.

Keywords Human values · Personalization · Conversational robot · Human-robot interaction (HRI)

1 Introduction

Dialogue robots and systems that interact with humans are attracting attention and are being researched in various fields, such as education [1] and mental health [2]. Dialogue systems are categorized into task- and non-task-oriented dialogue systems. Task-oriented dialogue systems engage in conversations designed to accomplish specific tasks, such as providing transportation information [3] or booking seats [4]. In contrast, non-task-oriented dialogue systems do not have tasks to accomplish and aim to sustain a conversation, like chatting. Regarding non-task-oriented dialogue systems, research has been conducted on generating natural utterances using human dialogue data [5] and on developing dialogue systems that enhance user satisfaction by attempting to understand users [6].

An important function of dialogue systems is to conduct conversations tailored to the user interacting with the system. In the healthcare field, it has been demonstrated that interacting with a system that adjusts its dialogue to match a user's mood is effective at reducing depression symptoms [7]. Furthermore, improvements in the recommendation efficiency of personalized recommendation dialogue systems based on user information [8] and the increase in user familiarity by remembering user information [9] have demonstrated the importance of personalization in both task-oriented and non-task-oriented dialogue systems.

In personalized dialogue systems research, one user characteristic that has not been given much consideration thus far is user values. Values are concepts or beliefs that guide people's selections or evaluations [10] and influence human behavior, such as the perception of morality [11]. Understanding values leads to predicting human behavior [12] and plays a significant role in building relationships [13].

✉ Yuki Sakamoto
sakamoto.yuki@irl.sys.es.osaka-u.ac.jp

Takahisa Uchida
uchida.takahisa@irl.sys.es.osaka-u.ac.jp

Hiroshi Ishiguro
ishiguro@irl.sys.es.osaka-u.ac.jp

[1] Graduate School of Engineering Science, Osaka University, Toyonaka, Japan

Moreover, it has been shown that understanding values is important in communication between individuals from different backgrounds [14]. Thus, values are important in understanding others in communication, and understanding a user's values in a dialogue system can potentially improve the quality of communication.

In this study, we aim to develop a system that understands user values and engages in dialogue. Furthermore, we investigate the effect of dialogues that understand user values on the interaction between the user and the system. The contributions of this study are as follows:

- We propose a method and dialogue strategy for estimating user values solely from information obtained via chat dialogues.
- We demonstrate that our proposed method can enhance the user's impression of the system.
- Furthermore, we demonstrate that our proposed method has the potential to promote a deeper understanding of their own values in users.

The remainder of this paper is organized as follows: Section 2 reviews previous research on personalized dialogue systems and systems that focus on values. Section 3 describes the system through which the proposed method estimates user values and conducts dialogues. Section 4 details the experimental setup, and Sect. 5 presents the experimental results. Section 6 presents a detailed discussion of experimental results. Section 7 concludes the paper.

2 Related Work

This section reviews previous research on personalized dialogue systems and systems focused on values and explains the novelty of this study.

Previous research has been conducted on personalized dialogue systems to adapt to users by estimating and understanding their characteristics. Mo et al. proposed a method for efficiently conducting task-oriented dialogue by personalization according to the user through transfer learning [15]. Inaba et al. proposed a model for estimating users' interests from information in open-domain conversations [16]. Aicher et al. proposed a system to sustain discussions by estimating and utilizing the user's implicit interests in an argumentative dialogue [17]. Uchida et al. presented a method for quickly estimating user preferences using dialogue robots and demonstrated that their proposed method improves user conversation satisfaction [6]. Thus, research is progressing regarding the speed of understanding users and aspects to be understood.

In addition, values influencing user behavior help others understand users. Therefore, they have been studied across various fields. Ihara et al. hypothesized that in communication, people deduce value similarities by repeating agreeing and disagreeing with each other's opinions. Furthermore, they analyzed evaluations related to agreement and disagreement in the starting phrases of statements [18]. Kern et al. demonstrated the potential to extract user personalities, including values, using social media data and to suggest careers suitable for users [19]. In recommendation systems, Hattori et al. showed that modeling personal values based on item attributes that users consider important in decision-making. Furthermore, making recommendations based on these models can improve the recommendation accuracy during cold-start situations [20]. Kosaka et al. proposed a method for estimating personal values using rough sets and applying the estimation results to collaborative filtering [21].

Considering the above, although it has been shown that user values are important for understanding the user, research on the interaction between dialogue systems personalized according to user values and users has not been conducted. Therefore, in this study, we constructed a system that understands user values through dialogue and incorporates them into the dialogue. We then examine, through dialogue experiments, the impact of understanding values on user interaction.

3 Proposed Method

To build a dialogue system that understands users' values, this section first organizes previous research on values and constructs a user model that incorporates values for this study (Sect. 3.1). Based on this user model, a method for estimating users' values is proposed (Sect. 3.2). The dialogue flow and content, including these estimations, are discussed (Sect. 3.3). Finally, the robots and operation management systems used in this study are described (Sect. 3.4).

3.1 Values Model

The means-end chain model [22] is widely known as a theory that links user behavior and values and has been used in many studies [23–25]. This model connects choice behavior, consequences, and values based on two assumptions. The assumptions are as follows: (1) values defined as the preferred end states play a dominant role in guiding choice patterns, and (2) people deal with the complexity of choices by grouping them into sets or classes to simplify the decision-making process. A laddering method is often used to estimate user values based on the means-end chain model.

Laddering is a technique that reveals consumers' values by repeatedly asking questions such as "why is it important"? [26].

Based on this model, we construct a user model that includes the values used in this study (Fig. 1). In this model, preferences are set as the target of the user's choice behavior, and these preferences are linked to the values through the reason for the preferences. During the dialogue, user preferences are collected mainly as information regarding hobbies, and values are inferred based on the reasons for these preferences. Hobbies provide general information regarding personal preferences and are common in everyday conversations. Furthermore, dialogue systems that consider user preferences, such as hobbies, have been shown to enhance user dialogue satisfaction [27]. Therefore, hobbies are considered appropriate topics within user preferences. Values are inferred by targeting preferences as the subject of choice behavior and asking, "Why do you like it?" (a question that probes the reasons for preferences).

In addition, a common theory for classifying values is the theory of values developed by Schwartz et al. [10]. They classified human values into ten values: power, achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, security. Furthermore, they categorized the values into four higher-order values: openness to change, self-enhancement, conservation, self-transcendence. This theory is useful for understanding values and has been employed in various fields [28, 29]. In this study, we used this classification to define the values. In the following text, "value" primarily refers to Schwartz's 10 types of values. When referring specifically to the four higher-order values, the term "higher-order value" is used.

3.2 Estimation of User Models Including Values

In this section, we describe the user model estimation system shown in Fig. 1 through dialogue. Here, we primarily use a large language model (LLM) to estimate values based on the users' preferences and the reasons they prefer them. Furthermore, an infinite relational model (IRM) is used to estimate the relationships between multiple elements within the user model.

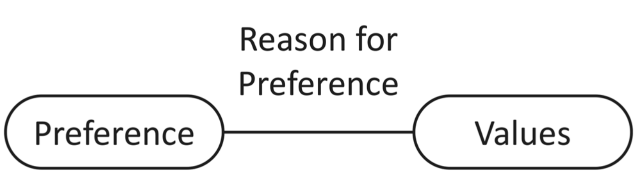

Fig. 1 User model including values in this study

3.2.1 Estimation of Values Using Large Language Model (LLM)

Using a large language model (LLM), we estimate values based on user preferences and the reasons for these preferences. LLMs are language models trained with vast amounts of text data demonstrating various capabilities. For example, the potential of LLMs to simulate human behavior [30] and their inferential abilities, such as estimating a user's personality, have been demonstrated [31]. In addition, studies analyzing LLMs using Schwartz's theory of values have shown that not only can LLMs generate text mapped to the theory of human values, but also that there is no observed bias in values [32]. Based on these considerations, LLMs can be employed to estimate user values.

Neural matching methods, which are technologies capable of considering the meaning of words and are frequently used in natural language processing, can also be used to estimate values from dialogue. In neural matching, domain shifts between the training and test datasets often pose challenges. However, combining these methods with techniques designed to address domain shifts [33, 34] makes it possible to adapt models trained for other tasks for value estimation. Nevertheless, considering the inherent properties of LLM regarding values and their ability to be used in conversation, employing LLM is advantageous for performing estimations and integrating the results within dialogues. Consequently, this study adopts an LLM to estimate values from dialogue. Control through prompting is often employed when LLMs are used for specific tasks. Prompt-based control is achieved by providing the LLM with text that includes instructions for the task to obtain appropriate outputs. Despite the low cost, prompt-based control can exhibit performance comparable to methods that adjust model parameters [35]. In this study, user values are estimated using a prompt-based control. The prompts used within the system for estimating the values are as follows.[1]

> You are an expert in estimating values. Based on the following statement, what kind of values does this person hold? Please estimate their values based on a list of values, and indicate the proportion of each value in a percentage format such as "Value: XX%." Do not include reasons in the output.
>
> ### list of values ###
>
> Power, Achievement, Hedonism, Stimulation, Self-Direction, Universalism, Benevolence, Tradition, Conformity, Security
>
> ### statement ###
>
> {user's statement}
>
> ### Estimated Values ###

Here, the information regarding preferences and reasons for preferences, obtained from conversations with users, is inserted in the form "Because of XX, I like YY" in the

[1] The prompts used in the system have been translated from Japanese into English in this paper.

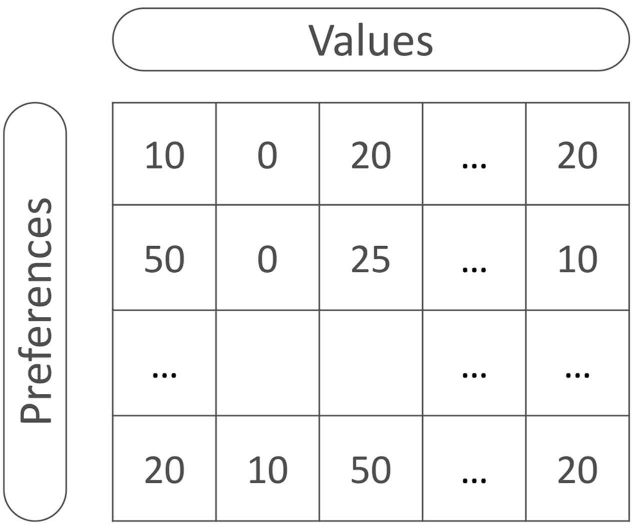

Fig. 2 Relational data between the user's preferences and values

{user's statement}. The list of values was compiled using Schwartz's theory of values [10]. Following the format of previous research on human-robot value alignment [36], the output is structured in a percentage form. Specifically, the output is "Value1: XX%, Value2: YY%." In this study, we employ GPT-4 by OpenAI[2] [37] as the LLM.

3.2.2 Estimation of Relationships within the User Model Using the Infinite Relational Model (IRM)

The values obtained from the large-language model are estimated based on a user's single preference and the reasons for preference. However, the user value is often related to multiple preferences. An infinite relational model (IRM) is used to estimate the relationships between multiple elements of a user model.

IRM is a clustering method for relational data that can be applied to learning systems of concepts [38]. The use of an enhanced IRM with a Poisson process makes it possible to accommodate binary and count data [39]. The IRM does not require a predetermined number of clusters, which is particularly useful for cases where the number of clusters is unknown or may vary among individuals, such as those with personal values, thus allowing clustering. Furthermore, the means-end chain model assumes that people engage in the categorization process [22]. Hence, we can discover user categories of preferences and understand the values associated with them by clustering the preferences. Clustering values also allows us to examine their correspondence with the four higher-order values in Schwartz's theory of values [10]. Therefore, this study employed this method to estimate the relationships between multiple elements in a user model.

[2] https://openai.com/research/gpt-4

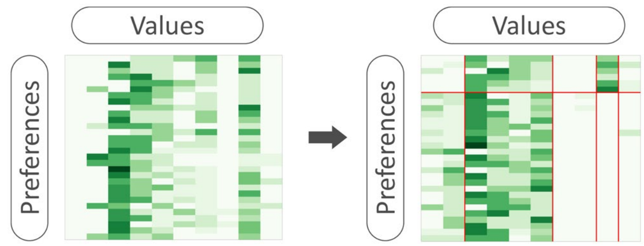

Fig. 3 Example of clustering relational data using IRM

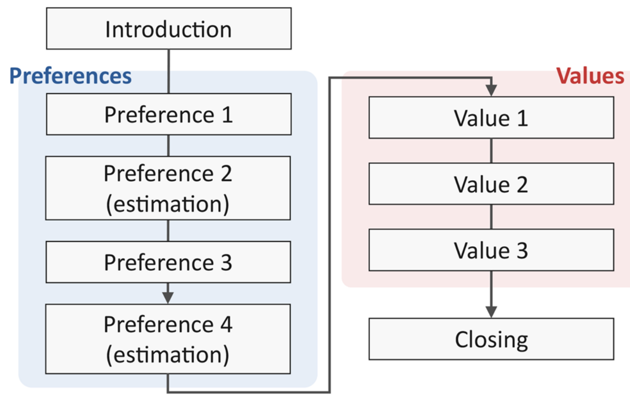

Fig. 4 Dialogue flow of the proposed method

The relational data targeted in this study include the relationship between user preferences and values (Fig. 2). For each preference, the data included estimates of the proportion of each value based on the reasons the users gave, which were obtained using LLM. When clustering such data using IRM, they are organized as shown in Fig. 3. The proportions of user values estimated by the LLM are represented by the dark green color, with each cluster corresponding to the sections divided by red lines. Thus, it is possible to cluster preferences and values based on their relationships using the IRM.

3.3 Dialogue Flow and Dialogue System

The dialogue flow of the system, built based on the above value-estimation method, is shown in Fig. 4. Here, we describe the dialogue regarding the preferences occurring in the first half of the conversation and the dialogue regarding the values conducted in the latter half.

3.3.1 Dialogue Regarding Preferences

In the first part of the conversation, information about the user's preferences is collected to estimate the user's values. We utilized LLM, which has proven useful in preference estimation [40], to comprehend user preferences rapidly.

Specifically, we queried the users' hobbies as preferences and why they prefer them. System utterances were constructed by embedding information regarding user preferences into predefined templates. Using this information, a large language model is employed to estimate the other preferences of the user and to engage in a conversation regarding the predicted preferences by conducting two sets of this process. User preference is estimated by adding the following prompt to the values estimation prompt presented in Sect. 3.2.1: "Estimate one hobby of this person other than {preferences already answered by user} and state it in one word."

While increasing the amount of data is necessary for user modeling, asking the user about many preferences during a conversation can be burdensome. Therefore, the number of preferences discussed with the user during the conversation was set to four. Based on this information, a large language model was used to extend the prediction to 30 preferences by estimating the user's other preferences. The prompts used within the system to estimate preferences are as follows:

> Please list five other "things this person likes" and five "things this person does not like" for the person making the statements below and explain the reasons for each. Format the reasons as "This is because XX."
>
> ### statements ###
>
> {The information regarding the preferences and the reasons the user answered.}
>
> ### things they like ###

In this study, estimating a large number of preferences in a single prompt can be time-consuming; therefore, the prompt is designed to estimate a smaller number of preferences. By conducting this process in parallel, the preferences can be estimated efficiently.

To avoid interruptions in dialogue owing to the time required for the estimation, the system generated simple context-aligned questions using the LLM (GPT-4) in parallel with each estimation. This approach prevents system response delays due to the estimation process. The following prompt was used to generate these questions: "Please generate a question for someone who says they like {user's stated preference} because of {reason user likes the preference}, to seek their opinion about {user's stated preference} further. Maintain the question within 50 characters." This dialogue occurred once during estimation for Preferences 2 and 4 (Fig. 4). After Preference 4, these questions were continually generated until all estimations were complete.

Table 1 Symbols used in the formulation of values utterances

Symbols	Definition
V_l	The set of lower-order values
V_h	The set of higher-order values
$W_i (i \in V_h)$	The set of lower-order values contained within higher-order values
P	The set of preferences
$a_j^i (i \in V_l, j \in P)$	The proportion of value i in preference j
P_{cls}	The set of preference clusters
V_{cls}	The set of value clusters

3.3.2 Dialogue Regarding Values

After modeling based on information from the dialogue regarding preferences, a dialogue regarding values was conducted. Utterances regarding values were created by embedding the estimated values into predefined templates. Three patterns regarding the statements on values can be considered: statements focused solely on user values, preference clusters, and value clusters. Therefore, the following statements were prepared for the system utterances:

1. Utterances that convey the estimation results of the user's most substantial value.
2. Utterances focusing on preference clusters convey the estimation results of values with a high information content.
3. Utterances focusing on value clusters convey the most vital value among the higher-order values.

Each utterance is conducted in a manner that provides feedback to the user in the form of questions or, as the robot's opinion, ensuring that the system's estimated values are not imposed on the user but rather confirm the user's opinion. Below, we outline the formulation of each utterance along with actual examples of the utterances.

1. This involves estimating the user's most substantial value and conveying that value through statements. Specifically, an average was calculated for each of the ten values, and the value with the highest average was pronounced as the user's strongest value. Using the symbols listed in Table 1, the value v_{focus} to be referred can be formalized as follows:

$$v_{focus} = \arg\max_{v \in V_l} \frac{1}{|P|} \sum_{j \in P} a_j^v \tag{1}$$

The actual statement is, "Earlier, you mentioned that you like/dislike {the preference the user mentioned} because of {the reason for preference the user provided}. Hence, I thought you value $\{v_{focus}\}$. Am I right?"

2. This statement focuses on the preference clusters. An average is calculated for each preference cluster within the ten values, and the difference between the highest and lowest average values among these clusters is computed. The value with the greatest difference in the average values is referred to as having a high amount of information. This can be interpreted as a value that

appears strongly in some preferences but not as much in others. Using the symbols listed in Table 1, the value v_{focus} to be referred can be formalized as follows:

$$v_{focus} = \arg\max_{v \in V_l}(maxM^v - minM^v) \tag{2}$$

$$M^v = \left\{ \frac{1}{|C|} \sum_{j \in C} a_j^v | C \in P_{cls} \right\} \tag{3}$$

The actual statement is, "Additionally, I thought that you might also value $\{v_{focus}\}$. After all, you mentioned earlier that you like/dislike {the preference the user mentioned} because of {the reason for preference the user provided}."

3. The utterance focuses on the value clusters. The correspondence was verified between clusters of 10 values and the 4 higher-order values and communicated the strongest value within the higher-order values. Specifically, the average value within each value cluster was calculated, and the cluster with the highest average value was highlighted. The method for determining the cluster V_{focus} using the symbols presented in Table 1 can be formalized as follows:

$$V_f = \arg\max_{D \in V_{cls}} \frac{1}{|D|} \frac{1}{|P|} \sum_{i \in D, j \in P} a_j^i \tag{4}$$

The elements obtained within the clusters were compared with the elements of value clusters from Schwartz's theory of values [10]. The structures of the higher-order clusters used in the system are listed in Table 2. A cluster with two or more matching elements in the preceding value clusters is referred to in the dialogue. The value v_{focus} to be referred can be formalized as follows:

$$v_{focus} = \begin{cases} v & (if \arg\max_{v \in V_h} |V_f \cap W_v| \geq 2) \\ none & (else) \end{cases} \tag{5}$$

The statement is, "Thus, you value $\{v_{focus}\}$, right?" In addition, if there are no clusters with two or more matching elements, statement is, "It seems you hold various values."

Because directly conveying the labels of values from previous research make the dialogue difficult to understand, in conversations, the labels are paraphrased (Tables 3 and 4). This paraphrasing is based on the output obtained by inputting the labels and definition of values [41] into the LLM and instructing it to translate them into more understandable format.

Table 2 Higher-order values clusters structure

Higher-order value types	Lower order value types
Openness to Change	Self-direction, Stimulation, Hedonism
Self-Enhancement	Hedonism, Achievement, Power
Conservation	Security, Conformity, Tradition
Self-Transcendence	Benevolence, Universalism

Table 3 Rephrasing of the ten values

Value	Rephrase
Power	Power such as social status or prestige
Achievement	Achieving success in society through one's own abilities
Hedonism	Enjoyment and satisfaction
Stimulation	Excitement and challenges in daily life
Self-direction	Thinking and acting independently
Universalism	Understanding all people and having an open mind
Benevolence	The happiness of those around
Tradition	Traditional culture and ways of thinking
Conformity	Being careful not to break rules and getting along well with others
Security	Maintaining good relationships with others and ensuring one's own stability

Table 4 Rephrasing of the four higher-order values

Value	Rephrase
Openness to Change	Open to change
Self-Enhancement	Acting to improve oneself
Conservation	Thinking conservatively
Self-Transcendence	Having a broad perspective

3.4 Robot and Motion Management System

The proposed system was implemented on a humanlike android, Geminoid F., which has 12 joints (primarily in the head area). Each joint is pneumatically driven, allowing Geminoid F to create humanlike expressions and movements. Figure 5 shows the Geminoid F android robot.

The android robot is used because people prefer androids over robots that look more robotic when discussing hobbies [42]. Moreover, studies have suggested that individuals can recognize androids as social beings [43, 44], implying that android robots that resemble humans can facilitate natural dialogue for users. In dialogues that address values that are both important and complex [45], it is reasonable to employ robots that resemble humans.

The motion control system was implemented using an existing system [46]. Microsoft Kinect was used as the sensor to detect the conversation partner's position. Unconscious movements, such as eye movement control and blinking, are generated using existing software [47]. Android speech

Fig. 5 Android robot Geminoid F

Fig. 6 Dialogue setup during the experiment

sounds were produced using the text-to-speech software ERICA[3], developed by HOYA Corporation.

The Wizard of Oz (WoZ) method [48] is used for controlling turn-taking in conversations. The WoZ method is widely employed in the field of human-robot interaction (HRI) and is utilized to prevent conversational breakdowns in dialogues regarding values, where users often speak while thinking.

4 Experiment

To verify the usefulness of the dialogue robot integrated with the values-estimation function described above, a conversational experiment was conducted. The experimental setup is shown in Fig. 6. All participants provided informed consent before the start of the study, which was approved by the Ethics Committee of Osaka University, Japan.

In the experiment, the following hypotheses were tested:

- H1: Proposed method is capable of estimating the user's values.
- H2: Proposed method improves the quality of the conversation.
- H3: Proposed method enhances the user's impression of the robot.
- H4: Proposed method deepens the relationship between the user and the robot.
- H5: Proposed method deepens the user's understanding of their own values.

For the verification of the hypotheses, the following three conditions were prepared:

- C1: Condition in which the conversation is based on values estimated using the proposed method.
- C2: Condition where the conversation is based on random values.
- C3: Condition where no values are mentioned, and the conversation focuses solely on preferences.

C1 is a condition where the user's values are estimated using the proposed method, and the conversation is based on the estimated values. C2 is a condition where the conversation is based on randomly selected values without estimating the user's values. This is designed to verify whether the estimation of values by the proposed method is useful compared to C1 and to consider the possibility of phenomena such as the Barnum effect [49] (where generic statements apply to anyone perceived to be specifically applicable to oneself) that may occur as a result of utterances regarding values. This is realized by randomly selecting v_{focus} for utterances related to values, as described in Sect. 3.3.2. C3 is a condition in which no values are stated, and to maintain the length of the conversations comparable with other conditions, the dialogue is extended by discussing preferences.

The flow of dialogue for each condition is shown in Fig. 7. In the experiment, structured dialogues were adopted, starting with conversations about preferences followed by conversations about values. While the experiment is controlled using such scenarios, the proposed method can be applied to other scenarios, such as casual conversations. Preferences, such as hobbies, are commonly mentioned in daily conversation [50], and dialogue systems that consider user preferences have been shown to improve user satisfaction [27]. Thus, our method, which uses preference-related topics, can be applied in various situations. In addition, in the experiment, structured dialogues with predetermined

[3] https://readspeaker.jp/

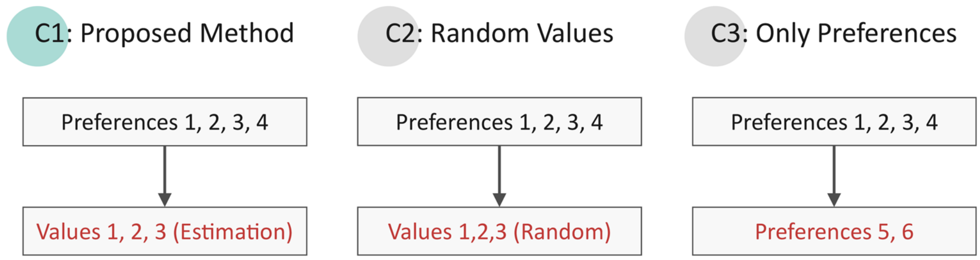

Fig. 7 Flow of dialogue under each condition

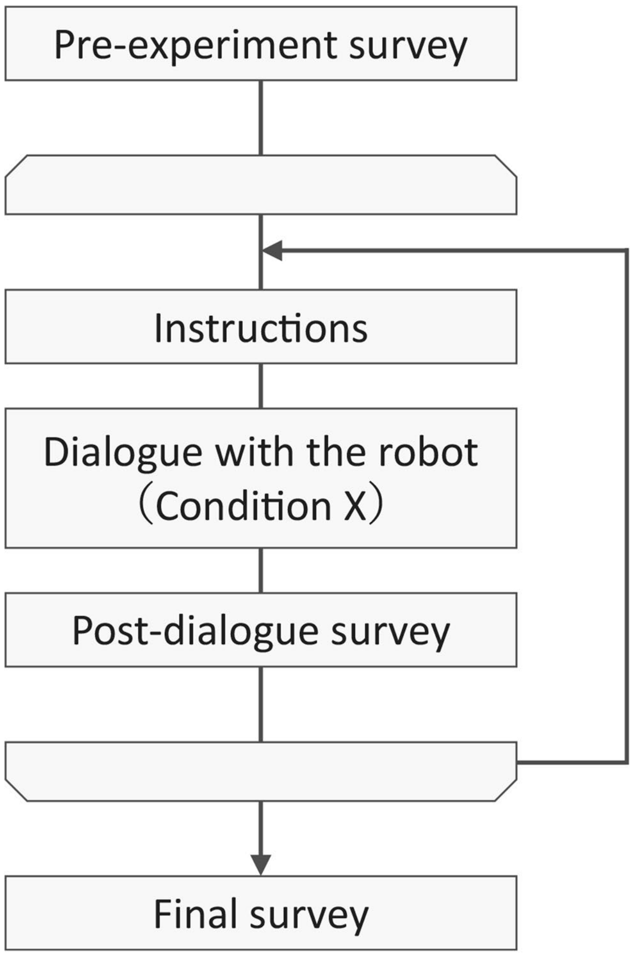

Fig. 8 Overall flow of the experiment

flows were used to clarify the differences between conditions and improve the interpretability of the results. This approach also reduces the occurrence of utterances unrelated to value modeling (such as statements of objective facts) that could potentially affect the impressions of the dialogue. However, the proposed method can also be applied to unstructured dialogues by integrating it with other methods that extract user preferences and reasons from the dialogue. Furthermore, while it is conceivable to validate the proposed method using existing dialogue datasets (e.g., chit-chat format datasets), conducting real dialogue experiments with human participants is necessary. This approach helps to directly evaluate the participants' impressions when the estimated user models are applied in interactions.

The overall flow of the experiment is shown in Fig. 8. The participants conducted dialogues with the android, and the order in which the conditions were presented was counterbalanced. The instructions mainly communicated following points:

- The robot you will be conversing with is intended to be your assistant robot, which attempts to understand and support you.
- (For the 2nd and 3rd dialogues) Please assume it is a different robot from the previous conversation.
- (For the 2nd and 3rd dialogues) It does not matter if the hobbies and the reasons you like them, asked during the dialogue, are the same or different from previous dialogues.

Moreover, to verify the hypotheses, the following evaluation metrics were prepared:

- Accuracy of the system's estimation of participants' preferences.
- Accuracy of the system's estimation of participants' values.
- Participants' agreement rate with the system's estimation of their values.
- Quality of dialogue [6].
- GodSpeed [51].
- IOS scale [52].
- New insights into one's preferences and values.

The accuracy of the system's estimation of participants' preferences was measured by asking them how much they liked each of the 30 preferences used for the values estimation, and responses were collected on a 7-point Likert scale. Preferences that received a score of 4 or less were classified as disliked, and those that received a score of 5 or more were classified as liked. The system's estimation accuracy was determined by the match rate between the system's estimated likes and dislikes and those of the participants. The estimation accuracy was calculated as the proportion of preferences for which the system's estimated likes and dislikes matched the participants' responses.

Regarding the estimation accuracy of the participants' values predicted by the system, participants were asked, "Please tell us how much you value the following values." They responded using a 7-point Likert scale for 10 different values. The estimated user model was then averaged across each value to derive the estimated value. For each dataset, the proportion of each value within the total set of values was calculated, and the estimation accuracy was determined as the Root Mean Squared Error (RMSE) between participants' responses and the system's estimated values.

The participants' rate of agreement with the system's values estimation was measured by the rate of agreement between the participants and the statements regarding their values deduced by the robot during the dialogue. The affirmation rate was calculated as the proportion of times the participants responded positively to the total number of value-related statements made by the robot.

In the questionnaire items regarding new insights into one's preferences and values, the questions asked were: "Through your conversations with this robot, did you gain any new insights into your hobbies?" and "Through your conversations with this robot, did you gain any new insights into your values?" The responses were obtained using a 7-point Likert scale.

5 Results

The details of the attributes of the study participants are as follows: 37 participants were included, of which 35 (15 males, 20 females; average age 21.29 (SD=2.50)) completed the experiment. The experiments for two participants were interrupted halfway owing to system problems. Figure 9 presents the statistics regarding the participants' hobbies, illustrating the distribution of hobbies that the participants mentioned in response to the robot's inquiries during the dialogue. Table 5 presents an example of a dialogue under the proposed conditions.

First, we verified the estimation accuracy of the preferences, which is fundamental for estimating values. The estimation accuracy obtained was 64.38%. As this exceeds the chance level of 50%, it is apparent that the system can estimate user preferences to a certain degree. The RMSE score for the value estimation accuracy was 0.08510.

In addition, the affirmation rate of the questions regarding values was checked as the accuracy of estimating values. Cases in which participants did not respond to the robot's utterances or answered "I don't know", were excluded from the calculation of the affirmation rate. The affirmation rates under conditions C1 and C2 for the utterances related to values are presented in Table 6. It is evident that the affirmation rate for C1 was higher than that for C2. Thus, the hypothesis "H1: The proposed method is capable of estimating the user's values." is supported.

One-way ANOVA was conducted for each questionnaire item prepared for dialogue quality (comprehension ability, willingness to understand, dialogue satisfaction, and dialogue motivation), no statistically significant differences were observed (see Fig. 10). Therefore, the hypothesis "H2: The proposed method improves the quality of dialogue." was not supported.

In conducting a one-way ANOVA on the participants' responses to the GodSpeed questionnaire items (anthropomorphism, animacy, likability, and perceived intelligence) aimed at measuring the participants' impressions toward the robot (Fig. 11), significant differences between conditions were found for animacy $(F(2,68) = 3.95, p < .05$), and multiple comparisons revealed a significant difference in the comparison ($C1 > C3 : p < .05$). Furthermore, in perceived intelligence, significant differences between

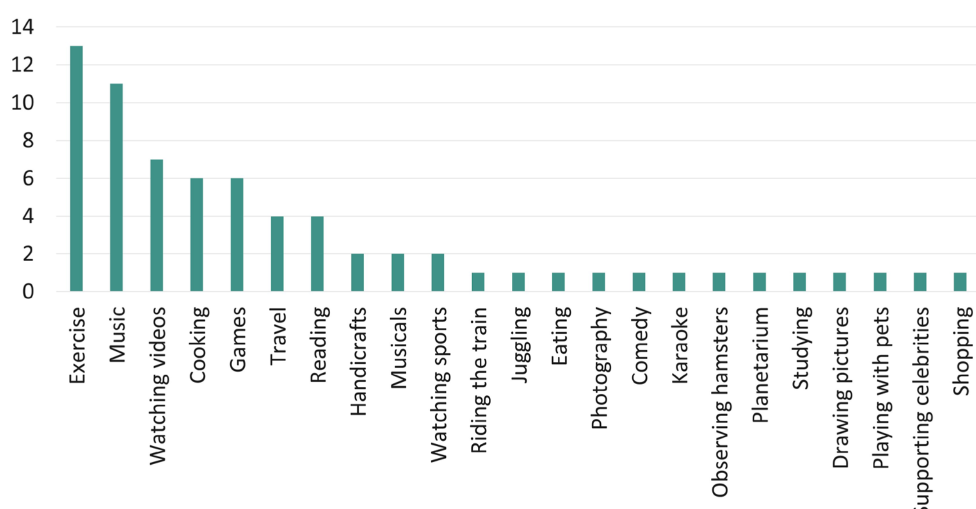

Fig. 9 Distribution of participants' preferences. The vertical axis indicates the number of counts. Preferences with the same meaning have been consolidated

Table 5 Dialogue example in condition 1 (proposed method)

System 1	Hello. I am an android named F. I am excited to talk with you today. Could you tell me one of your hobbies?
User 1	My hobby is listening to music.
System 2	You like listening to music, huh? Why is that?
User 2	Because it excites me.
System 3	I see. Specifically, which genres or artists' music excites you the most?
User 3	I become excited when I listen to classical music.
System 4	I see. Then I thought you might also like dancing in addition to listening to music. Is that correct?
User 4	I am not really into dancing.
System 5	Oh, you are not into dancing. Why is that?
User 5	Because I am not good at physical activity.
System 6	This makes sense. Hmm, I would like to hear about your other hobbies. What other hobbies do you have?
User 6	Another hobby of mine is taking walks.
System 7	You like taking walks, huh? Why is that?
User 7	Because it refreshes my mind.
System 8	I see. Where do you find it particularly refreshing to take walks?
User 8	I walk in various places, but I often walk in parks near my home.
System 9	I see. Then I thought you might also like reading. Is that correct?
User 9	Yes, I like reading too.
System 10	You like reading as well, huh? Why is that?
User 10	Because reading allows me to discover worlds unknown to me.
System 11	I see. Hmm, which genres of books expand your horizons the most?
User 11	I like novels.
System 12	I see. Which author's novels have broadened your perspective the most, and why do you feel that way?
User 12	Proust's novels have incredibly detailed backgrounds that teach me about new worlds.
System 13	I see. What was the new world you felt most from Proust's works?
User 13	In "In Search of Lost Time," I empathized greatly with the protagonist.
System 14	I see. Earlier, you mentioned that you don't like dancing because you're not into physical activities. Hence, I thought you value thinking and acting independently. Am I right?
User 14	Yes, that is correct.
System 15	I suspected as much. Additionally, I thought that you might also value excitement and challenges in daily life. After all, you mentioned earlier that you like listening to music because it excites you.
User 15	Yes, that is right.
System 16	I see. Thus, you value conservation, right?
User 16	More or less, yes.
System 17	I thought so. I enjoyed getting to know you through the conversations today. Thank you very much.

Table 6 Affirmation rate toward statements on values

Condition	Affirmation Rate (%)
C1	78.01
C2	60.95

conditions were revealed ($F(2,68) = 3.18, p < .05$), and multiple comparisons indicated significant differences in the comparisons ($C1 > C2 : p < .05, C1 > C3 : p < .05$). In addition, no statistically significant differences were found in the anthropomorphism and likability. Therefore, the hypothesis "H3: The proposed method enhances the user's impression of the robot." is partially supported.

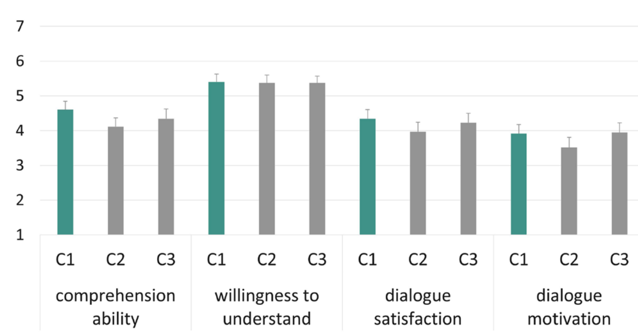

Fig. 10 Quality of dialogue (indicating the mean and standard error)

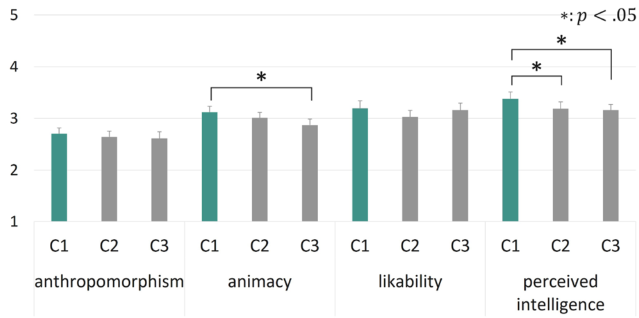

Fig. 11 Impressions toward the robot (indicating the mean and standard error)

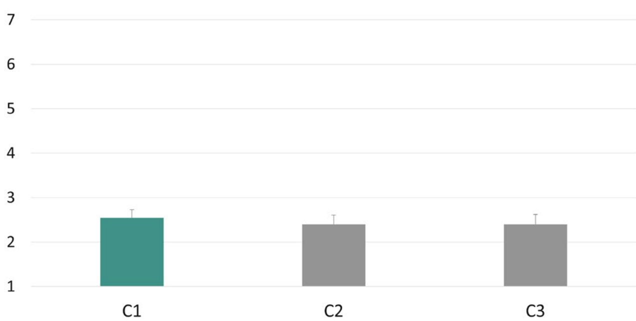

Fig. 12 Relationship with robots (indicating the mean and standard error)

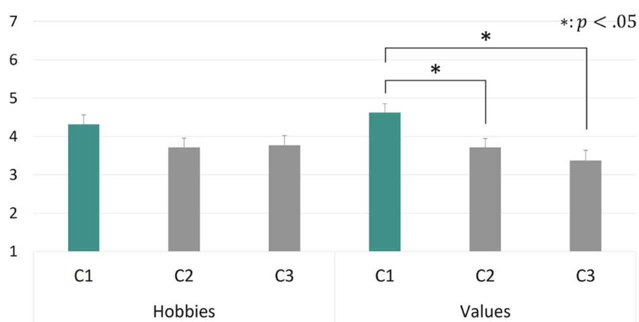

Fig. 13 New insights into one's own preferences and values (indicating the mean and standard error)

When a one-way ANOVA was performed on the participants' responses to the IOS scale items measuring their relationship with the robot, no statistically significant differences were observed (Fig. 12). Therefore, the hypothesis "H4: The proposed method deepens the relationship between the user and the robot." was not supported.

As a result of performing a one-way ANOVA on the participants' responses to the new insights into their preferences and values (Fig. 13), a significant difference between conditions was found in the questionnaire items related to the insights of values $(F(2,68) = 14.27, p < .01)$. Furthermore, following multiple comparisons, a significant difference was found in the combinations $(C1 > C2 : p < .05, C1 > C3 : p < .05)$. It suggests that the effect of users gaining new insights into their values is enhanced when the robot articulates the user's estimated values compared to the conditions where random values are articulated or where no values are articulated. Therefore, the hypothesis "H5: The proposed method deepens the user's understanding of their own values." is supported.

6 Discussion

Based on the experimental results, among the hypotheses tested to verify the effectiveness of the proposed method, H1 and H5 were supported, and H3 was partially supported. From these findings, the proposed method can estimate user values, particularly by enhancing the animacy and perceived intelligence of the robot and further prompting new insights into its values for users. Additionally, the acceptance rate of the dialogue concerning the values based on structures extracted using IRM was 78.01%, which partially supports the effectiveness of IRM in our method.

Regarding the impression of the robot, enhanced animacy and perceived intelligence can be influenced by the robot's utterances regarding the user's internal values. By articulating user values, users may perceive more human-like characteristics in the robot, leading to an improved perception of its animacy and perceived intelligence. Animacy and perceived intelligence are important for enhancing the quality of interaction and are important indices in the field of human-robot interaction (HRI) (e.g [53]). Previous research has shown that imparting a degree of social intelligence to household communication systems can increase the system acceptability and induce social behavior from users toward the system [54]. Because the perceived intelligence of robots is related to trust in robots [55], utterances regarding user values may also lead to increased trust in robots. Furthermore, it is not surprising that no significant differences were observed in the anthropomorphism and likability of the robot's impression. The difference between the experimental conditions was whether the robot expressed the values estimated from the user, which could have minimal impact on the robot's anthropomorphism. Moreover, under all conditions the robot engaged in a dialogue concerning the user's preferences, which likely contributed to its likability across all conditions, leading to no significant differences being observed between them.

New insights into their values for users were achieved through a process in which the robot understood values that the user might not have been aware of through dialogue, and the user gained a sense of agreement with the values presented by the robot. A previous study showed that people's awareness of their values led to choices that aligned with those values and enhanced their confidence in decision-making

[56]. It would also be meaningful to investigate the impact of the proposed method on decision-making.

A deeper understanding of one's values through interaction with the robot can be seen as an effect of constructive interaction [57]. Constructive interaction is a concept that explains the process of deepening one's thoughts through dialogue with others in the context of problem-solving. Although this study was not set in the context of users and robots solving a specific problem together, it is conceivable that the process of engaging in dialogue with a robot that understands users' values based on the question of "understanding the user" and delivers interpretations from the robot's perspective led users to develop their interpretations of their values further.

No significant differences were observed in the quality of the dialogue or the relationship between the user and the robot. One reason the quality of the dialogue did not improve could be that the content differed from that of typical everyday conversations. Considering human-human interaction, even if one understands the other's values, it might be rare to mention them explicitly. It is necessary to reconsider how to apply the estimated values to dialogue, including reviewing the design of utterances related to the values. In addition, the reasons for not deepening the relationship between the user and the robot could include the short duration of the dialogue and the fact that the robot's values were not considered. To deepen the relationship with robots, long-term interaction is important, and various studies have been conducted aiming at long-term interaction between social robots and users [58–60]. Furthermore, considering that similarity in values is critical for relationship building [13], it is necessary to understand user values and construct values for the robot to deepen the relationship. Future studies should aim to vary the interaction and robot design to verify whether these changes deepen the relationship.

The method proposed in this study involves applying zero-shot learning to an LLM to estimate user values, as described in Sect. 3.2.1. The experiment primarily aimed to compare conditions with and without value estimation and did not test multiple variations of the value estimation method. The user acceptance rate for value estimation using the proposed method reached 78.01%, supporting the performance of our approach. However, experiments that include variations in prompts for value estimation using LLMs (few-shot learning) should be conducted in the future, and ways to improve the quality of estimation further should be explored.

As detailed in Sect. 3.3.2, this study explored three approaches to reference values during dialogue based on the estimated user model. However, diverse methods potentially exist for applying the estimated user values within dialogues. By considering more complex methods and nuanced phrasing, the observed effects of this experiment could be further enhanced. This has been identified as a direction for future research.

Our method estimated the user model from structured dialogues about preferences in the experiment. When applied to unstructured datasets (e.g., chitchat format datasets), it should be integrated with techniques that extract information about preferences and reasons for liking them from such datasets; this will enable the estimation of user models.

We designed a dialogue flow that estimated user values based on information regarding the four preferences (Fig. 4). However, the validity of this value has not been verified. If insufficient preferences are mentioned and the dialogue concerning preferences is inadequate in duration, the user's subjective feeling of the reliability of the robot's estimation might decrease. Conversely, if surplus preferences are mentioned, the dialogue can become lengthy. Further verification should be conducted regarding the number of dialogue rounds required for the estimation and quality of dialogue.

For new insights into one's values obtained in this experiment, simple questions were used to measure participants' perceptions. Future research should explore various ways to assess new insights into one's values.

Furthermore, various studies have discussed the relationship between values and culture [10, 61, 62], and cultural differences have been considered to influence values. Thus, considering social values derived from culture could potentially enable faster estimation of a user's values model. Models that further refine Schwartz's values model have also been proposed [63], and such models could potentially allow for more precise estimation of values.

In this experiment, the android robot was used, which presents the possibility that the type of robot may have influenced the experimental results. For example, it has been shown that the robot's appearance can change people's attribution of responsibility to the robot or their preference for it [64, 65]. In addition, the gender of humanoid robots has been shown to influence stereotyping [66]. Given that the appearance of a robot has been demonstrated to affect users' perceptions of it, future research should involve experiments with different types of robots to examine the impact of robot type on interactions concerning values.

7 Conclusion

In this study, we constructed a system to estimate user values from dialogue and apply the estimation results to interactions. We conducted dialogue experiments to verify the usefulness of values estimation. The experimental results revealed that the proposed method could estimate the values of dialogue partners. Furthermore, conducting dialogues based on the estimated values enhanced the user's perception of the robot's animacy and perceived intelligence. It could also prompt users to discover new insights into their values. The results of this study demonstrate the effectiveness of the proposed method and suggest the importance of focusing on values in the interaction between humans and conversational robots. In the future, key challenges will include the faster modeling of user values using social values and the implementation of values models in robots.

Acknowledgements This work was supported by JSPS KAKENHI Grant Numbers 19H05693, 22K17949, and JST PRESTO Grant Number JPMJPR23I2.

Funding Open Access funding provided by The University of Osaka.

Data Availability Data will be made available on reasonable request.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

References

1. Kanda T, Hirano T, Eaton D, Ishiguro H (2004) Interactive robots as social partners and peer tutors for children: a field trial. Hum Comput Interact 19(1–2):61–84
2. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB (2019) Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry 64(7):456–464
3. Komatani K, Adachi F, Ueno S, Kawahara T, Okuno HG (2003) Flexible spoken dialogue system based on user models and dynamic generation of voicexml scripts. In: Proceedings of the Fourth SIGdial Workshop of Discourse and Dialogue, pp 87–96
4. Seneff S, Polifroni J (2000) Dialogue management in the mercury flight reservation system. In: ANLP-NAACL 2000 Workshop: Conversational Systems
5. Bessho F, Harada T, Kuniyoshi Y (2012) Dialog system using real-time crowdsourcing and twitter large-scale corpus. In: Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp 227–231
6. Uchida T, Minato T, Nakamura Y, Yoshikawa Y, Ishiguro H (2021) Female-type android's drive to quickly understand a user's concept of preferences stimulates dialogue satisfaction: dialogue strategies for modeling user's concept of preferences. Int J Social Rob 1–18
7. Fitzpatrick KK, Darcy A, Vierhile M (2017) 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 4(2):7785
8. Thompson CA, Goker MH, Langley P (2004) A personalized system for conversational recommendations. J Artif Intell Res 21:393–428
9. Tsunomori Y, Higashinaka R, Yoshimura T, Isoda Y (2020) An evaluation of a chat-oriented dialogue system that remembers and uses user information over multiple days (in japanese). Tran Jpn Soc Artif Intell AI 35(1):1
10. Schwartz SH (1992) Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In: Advances in experimental social psychology, vol 25. Elsevier, New York, pp 1–65
11. Finegan J (1994) The impact of personal values on judgments of ethical behaviour in the workplace. J Bus Ethics 13:747–755
12. Pozzebon JA, Ashton MC (2009) Personality and values as predictors of self-and peer-reported behavior. J Individual Differ 30(3):122–129
13. Murstein BI (1970) Stimulus. value. role: a theory of marital choice. J Marriage Family 465–481
14. Nordby H (2008) Values, cultural identity and communication: a perspective from philosophy of language. J Intercultural Commun 8(2):1–10
15. Mo K, Zhang Y, Li S, Li J, Yang Q (2018) Personalizing a dialogue system with transfer reinforcement learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol 32
16. Inaba M, Takahashi K (2018) Estimating user interest from open-domain dialogue. In: Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue, pp 32–40
17. Aicher A, Gerstenlauer N, Minker W, Ultes S (2022) User interest modelling in argumentative dialogue systems. In: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp 127–136
18. Ihara M, Kobayashi M (2005) Opening statements in turn taking to extract similarity of human sense of values for mutual understanding (in Japanese). IPSJ J 46(1):138–146
19. Kern ML, McCarthy PX, Chakrabarty D, Rizoiu M-A (2019) Social media-predicted personality traits and values can help match people to their ideal jobs. Proc Natl Acad Sci 116(52):26459–26464
20. Hattori S, Takama Y (2014) Recommender system employing personal-value-based user model. J Adv Comput Intell Intell Inf 18(2):157–165
21. Kosaka Y, Okamoto K (2021) Personal values modeling with rough sets for collaborative filtering. In: Proceedings of the 7th International Workshop on Advanced Computational Intelligence and Intelligent Informatics, pp 1–5
22. Gutman J (1982) A means-end chain model based on consumer categorization processes. J Marketing 46(2):60–72
23. Chiu C-M (2005) Applying means-end chain theory to eliciting system requirements and understanding users perceptual orientations. Inf Manage 42(3):455–468
24. Shimonishi K, Kawashima H (2020) A two-step approach for interest estimation from gaze behavior in digital catalog browsing. J Eye Mov Res 13 (1)
25. Le GHK, Le XTT (2022) Employing the means-end chain theory and the laddering technique to identify obstacles to practice green consumption from consumers' perspectives. Sci Technol Dev J Econ Law Manage 6(4):3458–3470

26. Reynolds TJ, Gutman J (2001) Laddering theory, method, analysis, and interpretation. In: Understanding consumer decision making. Psychology Press, UK, pp 40–79
27. Kobyashi S, Hagiwara M (2016) Non-task-oriented dialogue system considering user's preference and human relations (in Japanese). Tran Jpn Soc Artif Intell AI 31(1):1
28. Krystallis A, Vassallo M, Chryssohoidis G (2012) The usefulness of schwartz's 'values theory'in understanding consumer behaviour towards differentiated products. J Marketing Manage 28(11-12):1438–1463
29. Boer D, Fischer R (2013) How and when do personal values guide our attitudes and sociality? explaining cross-cultural variability in attitude–value linkages. Psychol Bull 139(5):1113
30. Aher GV, Arriaga RI, Kalai AT (2023) Using large language models to simulate multiple humans and replicate human subject studies. In: International Conference on Machine Learning, PMLR, pp 337–371
31. Yang T, Shi T, Wan F, Quan X, Wang Q, Wu B, Wu J (2023) Psycot: psychological questionnaire as powerful chain-of-thought for personality detection. arXiv preprint arXiv:2310.20256
32. Fischer R, Luczak-Roesch M, Karl JA (2023) What does chatgpt return about human values? exploring value bias in chatgpt using a descriptive value theory. arXiv preprint arXiv:2304.03612
33. Li J, Han P, Ren X, Hu J, Chen L, Shang S (2021) Sequence labeling with meta-learning. IEEE Trans Knowledge Data Eng 35(3):3072–3086
34. Li J, Feng S, Chiu B (2023) Few-shot relation extraction with dual graph neural network interaction. In: IEEE Transactions on Neural Networks and Learning Systems
35. Tu L, Xiong C, Zhou Y (2022) Prompt-tuning can be much better than fine-tuning on cross-lingual understanding with multilingual language models. arXiv preprint arXiv:2210.12360
36. Yuan L, Gao X, Zheng Z, Edmonds M, Wu YN, Rossano F, Lu H, Zhu Y, Zhu S-C (2022) In situ bidirectional human-robot value alignment. Sci Rob 7(68):4183
37. Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, Aleman FL, Almeida D, Altenschmidt J, Altman S, Anadkat S et al (2023) Gpt-4 technical report. arXiv preprint arXiv:2303.08774
38. Kemp C, Tenenbaum JB, Griffiths TL, Yamada T, Ueda N (2006) Learning systems of concepts with an infinite relational model. In: AAAI, vol 3, p 5
39. Blundell C, Beck J, Heller KA (2012) Modelling reciprocating relationships with hawkes processes. Adv Neural Inf Process Syst 25
40. Uchida T, Sakamoto Y, Koshino A, Ban M, Yoshikawa Y, Ishiguro H (2023) A pilot study on preference estimation function of large-scale language model for hospitality avatars. In: 2nd Symbiotic Society with Avatars (SSA) workshop at the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI'23)
41. Sagiv L, Schwartz SH (1995) Value priorities and readiness for out-group social contact. J Personality Social Psychol. 69(3):437
42. Uchida T, Minato T, Ishiguro H (2024) Opinion attribution improves motivation to exchange subjective opinions with humanoid robots. Front Rob AI 11:1175879
43. Pütten AM, Krämer NC, Becker-Asano C, Ogawa K, Nishio S, Ishiguro H (2014) The uncanny in the wild. analysis of unscripted human–android interaction in the field. Int J Social Rob 6:67–83
44. Watanabe M, Ogawa K, Ishiguro H (2015) Can androids be salespeople in the real world? In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp 781–788
45. Boyd R, Wilson S, Pennebaker J, Kosinski M, Stillwell D, Mihalcea R (2015) Values in words: using language to evaluate and understand personal values. In: Proceedings of the International AAAI Conference on Web and Social Media, vol 9, pp 31–40
46. Glas DF, Minato T, Ishi CT, Kawahara T, Ishiguro H (2016) Erica: the erato intelligent conversational android. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, pp 22–29
47. Higashinaka R, Minato T, Sakai K, Funayama T, Nishizaki H, Nagai T (2021) Spoken dialogue system development at the dialogue robot competition (in japanese). J Acoust Society Jpn 77(8):512–520
48. Fraser NM, Gilbert GN (1991) Simulating speech systems. Comput Speech Lang 5(1):81–99
49. Meehl PE (1956) Wanted—a good cook-book. Am Psychologist 11(6):263
50. Sehulster JR (2006) Things we talk about, how frequently, and to whom: frequency of topics in everyday conversation as a function of gender, age, and marital status. Am J Psychol 119(3):407–432
51. Bartneck C, Kulić D, Croft E, Zoghbi S (2009) Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. Int J Social Rob 1:71–81
52. Aron A, Aron EN, Smollan D (1992) Inclusion of other in the self scale and the structure of interpersonal closeness. J Personality Social Psychol. 63(4):596
53. Bartneck C, Kanda T, Mubin O, Al Mahmud A (2009) Does the design of a robot influence its animacy and perceived intelligence? Int J Social Rob 1:195–204
54. De Ruyter B, Saini P, Markopoulos P, Van Breemen A (2005) Assessing the effects of building social intelligence in a robotic interface for the home. Interact Comput 17(5):522–541
55. Becker D, Rueda D, Beese F, Torres BSG, Lafdili M, Ahrens K, Fu D, Strahl E, Weber T, Wermter S (2023) The emotional dilemma: influence of a human-like robot on trust and cooperation. In: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), IEEE, pp 1689–1696
56. Jonkmans A, Wurl J, Snelders D, Onselen L (2016) Junior designers' awareness of personal values and their employment choices. In: Future Focused Thinking - DRS International Conference
57. Miyake N (1986) Constructive interaction and the iterative process of understanding. Cognit Sci 10(2):151–177
58. Laban G, Kappas A, Morrison V, Cross ES (2023) Building long-term human–robot relationships: examining disclosure, perception and well-being across time. Int J Social Rob 1–27
59. Kasap Z, Magnenat-Thalmann N (2012) Building long-term relationships with virtual and robotic characters: the role of remembering. Visual Comput 28:87–97
60. Lee KM, Park N, Song H (2005) Can a robot be perceived as a developing creature? effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it. Hum Commun Res 31(4):538–563
61. Roccas S, Sagiv L (2010) Personal values and behavior: taking the cultural context into account. Soc Personal Psychol Compass 4(1):30–41
62. Maslova OV, Shlyakhta DA, Yanitskiy MS (2020) Schwartz value clusters in modern university students. Behav Sci 10(3):66
63. Schwartz SH, Cieciuch J, Vecchione M, Davidov E, Fischer R, Beierlein C, Ramos A, Verkasalo M, Lönnqvist J-E, Demirutku K, et al. (2012) Refining the theory of basic individual values. J Personality Social Psychol 103(4):663
64. Hinds PJ, Roberts TL, Jones H (2004) Whose job is it anyway? a study of human-robot interaction in a collaborative task. Hum Comput Interact 19(1-2):151–181

65. Goetz J, Kiesler S, Powers A (2003) Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. Roman 2003, IEEE, pp 55–60
66. Perugia G, Lisy D (2023) Robot's gendering trouble: a scoping review of gendering humanoid robots and its effects on HRI. Int J Social Rob 15(11):1725–1753