

Generating Diverse Personas for User Simulators to Test Interview Dialogue Systems

Mikio Nakano^{1,2}, Kazunori Komatani², Hironori Takeuchi³

¹C4A Research Institute, Inc., Setagaya, Tokyo, Japan

²SANKEN, University of Osaka, Ibaraki, Osaka, Japan

³Musashi University, Nerima, Tokyo, Japan

mikio.nakano@c4a.jp, komatani@sanken.osaka-u.ac.jp

h.takeuchi@cc.musashi.ac.jp

Abstract

This paper addresses the issue of the significant labor required to test interview dialogue systems. While interview dialogue systems are expected to be useful in various scenarios, like other dialogue systems, testing them with human users requires significant effort and cost. Therefore, testing with user simulators can be beneficial. Since most conventional user simulators have been primarily designed for training task-oriented dialogue systems, little attention has been paid to the personas of the simulated users. During development, testing interview dialogue systems requires simulating a wide range of user behaviors, but manually creating a large number of personas is labor-intensive. We propose a method that automatically generates personas for user simulators using a large language model. Furthermore, by assigning personality traits related to communication styles when generating personas, we aim to increase the diversity of communication styles in the user simulator. Experimental results show that the proposed method enables the user simulator to generate utterances with greater variation.

1 Introduction

Interview dialogue systems, which can efficiently gather information from many users, are a promising application of dialogue system technology and have attracted increasing research interest in recent years (Lang and Eskenazi, 2025; Hashimoto et al., 2025; Zeng et al., 2023). However, building dialogue systems, including interview dialogue systems, incurs significant costs. One major cost is the effort required for testing. The goal of our study is to alleviate the manual effort for testing interview dialogue systems.

To reduce testing costs, *user simulation* can be used instead of manual testing. However, existing user simulators have primarily been developed for the training and evaluation of task-oriented dialogue systems. Thus, we aim to develop a simu-

lator that is useful for testing interview dialogue systems.

When testing interview dialogue systems, simulating a wide variety of users is crucial. To increase variation, it is beneficial to have a large number of *user personas* (Zhang et al., 2018; Jiang et al., 2024; Hong et al., 2025). However, manually creating numerous personas is labor-intensive. While using a large set of pre-constructed personas has been considered (Mazaré et al., 2018), this approach is unsuitable for interview dialogues, where personas need to be aligned with the specific interview topics.

Therefore, this paper proposes a method to generate diverse personas using *large language models* (LLMs). In this method, a few manually created personas are provided as examples, and the LLM is tasked with generating a large number of additional personas. To ensure diversity in speaking styles and attitudes toward the system, personality traits are specified during generation.

Evaluation experiments demonstrated that the proposed method enables the creation of diverse dialogues through the generation of many personas. This increases the likelihood of discovering system issues. Note that detecting system issues from simulated dialogues is beyond the scope of our study. Manual inspection is one possible approach, but automated methods using LLMs are also conceivable (Finch et al., 2023).

The contributions of our study are as follows:

- We discuss that diverse personas are necessary for user simulators used in testing interview dialogue systems.
- We propose a method for generating diverse personas using LLMs.
- We introduce a technique to enhance persona diversity by specifying personality traits related to communication style.

- Through evaluation experiments, we demonstrate that the proposed persona generation method enables the creation of diverse dialogues.

2 Related Work

2.1 Interview Dialogue Systems

Interview dialogue systems that extract information from humans through natural language interaction have attracted significant research attention due to their high practical value. Examples of the application domains include: course rating surveys at universities (Stent et al., 2006), telephone surveys (Johnston et al., 2013; Lang and Eskenazi, 2025), mental health assessments (DeVault et al., 2014), dietary intake recording (Kobori et al., 2016), job interviews (Su et al., 2019; Rao et al., 2020; Inoue et al., 2020; Jones and Sabouret, 2012; Hoque et al., 2013; Gebhard et al., 2014), dietary preference surveys (Zeng et al., 2023), career counseling for nurses (Hashimoto et al., 2025), frailty diagnosis (Asao et al., 2020), collecting people’s beliefs about the future of robots (Skantze et al., 2012), facilitating user review writing (Tanaka and Inaba, 2024), and customer interviews (Sidaoui et al., 2020). We propose a method for efficiently testing such interview dialogue systems.

2.2 Testing Dialogue Systems

As mentioned earlier, testing dialogue systems requires significant effort, and various methods and tools have been proposed to facilitate this process (Li et al., 2022). For example, there are tools that verify whether a dialogue system behaves as expected using corpora (Degerstedt and Jönsson, 2006), and methods that test the system using predefined test cases (Atefi and Alipour, 2019; Guo et al., 2024; Gómez-Abajo et al., 2024). Since these testing methods require the preparation of test cases in advance, they are effective for detecting issues within the expected scope. However, in interactions with diverse users, unexpected problems may arise, and thus, methods capable of detecting such unforeseen issues are desired.

2.3 User Simulation for Dialogue Systems

We therefore use user simulators. User simulators for dialogue systems have traditionally been used to evaluate dialogue strategies (Eckert et al., 1997; Niimi and Nishimoto, 1999). However, with advances in statistical modeling of dialogue strate-

gies, research has increasingly focused on training dialogue strategy models using reinforcement learning (Schatzmann et al., 2006; Pietquin and Du-toit, 2006; Schatzmann et al., 2007; Ai and Litman, 2009; Li et al., 2017; Pietquin and Hastie, 2013).

Recently, with the development of large language models (LLMs), user simulators based on LLMs have been proposed (Terragni et al., 2023; Luo et al., 2024; Algherairy and Ahmed, 2025; Sekulic et al., 2024; Hu et al., 2023; Sun et al., 2023; Di Bratto et al., 2024; Liu et al., 2023).

In studies on user simulators for task-oriented dialogue systems, evaluation metrics such as similarity to human user behavior and the quality of the learned dialogue strategies have been used (Pietquin and Hastie, 2013).

Our study differs from these previous studies in that it proposes a method for constructing a user simulator specifically for testing interview dialogue systems. Different evaluation metrics are also required.

2.4 Using Personas in User Simulation

Research has also been conducted on using personas in user simulators. Georgila et al. (2008) and Georgila et al. (2010) demonstrated that different dialogue strategies can be learned by preparing simulators for older and younger users. Hashimoto et al. (2025) manually created nurse personas for a user simulator used in evaluating a career counseling dialogue system for nurses, under the supervision of nursing administrators. Our study differs from these studies in that it automatically generates personas.

2.5 Persona Generation

As for research on automatic persona generation, there are methods that extract personas from Reddit (Mazaré et al., 2018). They collect sentences that express a persona (e.g., those containing "I" or "my") and use them for persona creation. Methods for generating personas using LLMs have also been proposed. Ge et al. (2024) proposed a method for generating persona-representing passages from text using an LLM. However, with these methods, it is difficult to generate interviewee personas tailored to the domain of interview dialogue systems.

3 Proposed Method

In the proposed method, we generate simulated user personas using an LLM. It is necessary to

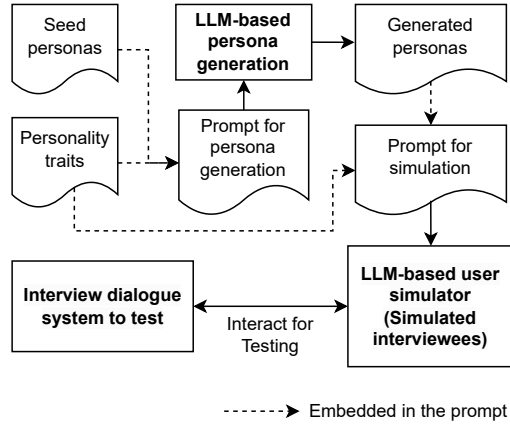


Figure 1: Overview of the proposed method.

ensure that the generated personas include information relevant to the domain of the target interview dialogue system. To achieve this, we perform in-context learning using a small number of manually written personas (called *seed personas* hereafter) as few-shot examples. Furthermore, to facilitate the discovery of potential issues by introducing greater variation in user utterances during simulation, we provide personality traits that define communication styles when generating personas with the LLM. By embedding the generated personas and personality traits into the prompt of an LLM-based user simulator, we can generate diverse dialogues. Figure 1 illustrates the proposed method.

We use the following two axes as personality traits.

Degree of Anthropomorphism One axis concerns whether the user treats the system as an object or as a human. Users who treat the system as an object aim to use the dialogue system efficiently and tend to use expressions that are easy for the system to understand. In contrast, users who treat the system as a human speak to it as if they were speaking to a person, without considering whether the system can understand them. This behavior can be interpreted as users anthropomorphizing the system (Reeves and Nass, 1996). We refer to this axis as the degree of anthropomorphism.

Degree of Elaborateness The other axis concerns whether the user engages in redundant communication or direct communication. Pragst et al. (2019) and Miehle et al. (2020) classify users’ communication styles into elaborateness and directness.

- Whether the user often travels
- Places the user has recently visited
- Favorite places among the user’s past travel destinations
- Activities the user enjoys doing while traveling
- Reasons for not traveling often (if applicable)
- Places the user would like to visit next
- How the user books trips (e.g., visiting a travel agency, calling a travel agency, using a website)

Figure 2: Topics the travel interview dialogue system asks the user about.

We adopt the same classification in our study.

Although the Big Five personality traits are well-known, we did not use them because it is considered more effective to directly specify users’ speaking styles is considered more effective for our purpose.

4 Evaluation

We conducted an evaluation experiment to investigate whether the proposed method is effective for testing interview dialogue systems. Note that this experiment deals only with Japanese systems, and that all subsequent figures are translations from Japanese.

4.1 Compared Methods

We compared the following five conditions:

BL (BaseLine): Only seed personas are used without generating new personas.

noPT (no Personality Traits): Proposed method without prompts concerning the two kinds of personality traits. Only seed personas are given when personas are generated.

APM (AnthroPoMorphism): Proposed method with degree of anthropomorphism given as a personality trait for persona generation. The degree is either High or Low.

EL (ELaborateness): Proposed method with degree of elaborateness given as a personality trait for persona generation. The degree is either High or Low.

APM+EL Proposed method with both degree of anthropomorphism and degree of elaborateness given as personality traits for persona generation.

4.2 Interview Dialogue Systems Used

We used the following two systems:

The first is a Japanese text-based interview dialogue system that conducts interviews about travel.

- Name: Risako Machiyama
 - 49 years old
 - Female
 - Lives in Tokyo
 - Lives with husband and two daughters (a senior high school second-year student and a junior high school third-year student)
 - Works part-time in an office
 - Speaks cheerfully and talks rapidly with many words
 - Travels about once a year
 - Most recently visited Taiwan with her family
 - Enjoys shopping while traveling
 - Would like to visit Los Angeles
 - Usually books trips by visiting a travel agency
- Name: Yuma Yamanaka
 - 32 years old
 - Male
 - Lives in Hokkaido
 - Lives alone
 - Chef
 - Speaks clearly and often talks about himself
 - Travels about once a year
 - Most recently visited Osaka alone
 - Enjoys food tours during travel for professional inspiration
 - Would like to visit Singapore
 - Usually books trips online

Figure 3: Examples of the seed personas for the travel interview dialogue system.

It was built using the ChatGPT dialogue built-in block of DialBB (Nakano and Komatani, 2024), where dialogues are conducted based on a single prompt template. The interview topics extracted from users are shown in Figure 2.

The second is a Japanese text-based interview dialogue system that conducts interviews focusing on the user’s preferences for sweets. It asks users questions such as whether they often eat sweets, what kinds of sweets they like, and where they usually buy sweets. This system was also built using DialBB, employing ChatGPT for language understanding, named entity extraction, and dialogue management using a state transition network. ChatGPT is used for evaluating transition conditions and for utterance generation within dialogue management. The state transition network consists of 32 states and 54 transitions.

4.3 Procedure

We conducted persona generation and simulation according to the following procedure.

First, for each system, we manually created 10 seed personas. Figure 3 shows examples of seed

Task

- Create 25 personas for interviewees being surveyed about their travel experiences and desired travel destinations, generating them in JSON format as shown in the example.
- Ensure as much variation among the personas as possible.
- The interviewees have the following personality traits; create personas consistent with these traits.
- Do not reuse any of the personas shown in the example.

Personality Traits

{personality}

Examples

{examples}

Figure 4: Prompt template for persona generation. {personality} is replaced with a description of personality traits, and {examples} is replaced with the seed personas converted into JSON format. In the noPT condition, the template without the personality trait description is used.

personas for the travel system. In this paper, we show only examples and prompt templates for the travel interview system for the lack of space. In the noPT, APM, EL, and APM+EL conditions, these seed personas were embedded into the prompt as few-shot examples to generate new personas. In the APM, EL, and APM+EL conditions, personality traits were assigned in a balanced manner, and personas consistent with those traits were generated. The template used for persona generation for the travel system is shown in Figure 4, and the descriptions of the personality traits used in the experiments are shown in Figures 5 and 6. The third item under "APM: High" in Figure 5 indicates that users who anthropomorphize the system attempt to test whether the system can engage in conversation as flexibly as a human. For each condition, 100 personas were generated.¹

The interview dialogue systems interacted with user simulators based on the generated personas. The prompt template used for simulation is shown in Figure 7. Each dialogue session consisted of 15 user utterances. For conditions other than BL, each persona was used for only one dialogue session, resulting in 100 dialogues per condition. For BL,

¹Personas were generated in batches of about 25 at a time, and when the total exceeded 100, 100 personas were randomly selected.

APM: High

- Views the system as a human rather than an object
- Speaks to the system as if speaking to a person
- Tries to test the system, abruptly changing topics or speaking at length

APM: Low

- Views the system as an object rather than a human
- Aware of the system's limitations in understanding, speaks as simply as possible
- Shows a cooperative attitude toward the system

Figure 5: Personality trait descriptions used in the prompt for persona generation by varying the degree of anthropomorphism.

EL: High

- Speaks redundantly
- Talks about things not explicitly asked
- Sometimes uses indirect expressions

EL: Low

- Speaks directly
- Speaks concisely
- Does not talk about unasked topics

Figure 6: Personality trait descriptions used in the prompt for persona generation by varying the degree of elaborateness.

each of the 10 seed personas was used 10 times.

The LLM used by the persona generator and the simulator was OpenAI’s gpt-4o-2024-11-20,² while the LLM used by the interview dialogue systems was OpenAI’s gpt-4o-mini-2024-07-18.³ Since our objective is to evaluate the simulator, we used a high-performance model for the simulator. The temperature parameter was set to 0.7 for all LLM usages.

Prompt tuning was conducted separately from the main experiments, using a different system (a sweets interview dialogue system that uses a single prompt template) and different seed personas from those used in the main evaluation described in Section 4.2.

4.4 Evaluation Metrics

Our goal is to automatically expose issues in interview dialogue systems. One direct way to evaluate

²<https://platform.openai.com/docs/models/gpt-4o>

³<https://platform.openai.com/docs/models/gpt-4o-mini>

Task Description

- You are acting as a user of a dialogue system and are being interviewed by the system. Based on the flow of the conversation so far, generate your next utterance.

Notes

- The utterance should consist of either 1, 2, or 3 sentences.
- Do not produce utterances of the same length consecutively.
- Each utterance must be within 100 characters.
- Speak in accordance with your assigned persona below.
- Do not include your name or the word "User" at the beginning of the utterance.
- Answer if you are asked for your name.
- Do not enclose the utterance in quotation marks.

Your Persona

{persona}

Dialogue History

{dialogue_history}

Figure 7: Prompt template used by the simulator. {persona} is replaced with the generated persona, and {dialogue_history} is replaced with the dialogue history at runtime.

the proposed method would be to build a faulty interview dialogue system, interact with it using the user simulator generated by the proposed method, and measure how many issues can be detected. However, faulty systems may cause dialogues to collapse once a problem occurs. In practice, using the user simulator to improve a system would involve repeatedly fixing detected issues and running new simulations, but evaluating this iterative process is impractical.

Therefore, as a second-best approach, we measure the diversity of user simulator utterances. The rationale is that the more diverse the utterances, the higher the probability of uncovering system issues.

Following previous studies on user simulators for task-oriented dialogue systems (Terragni et al., 2023; Algherairy and Ahmed, 2025; Sekulic et al., 2024), we used the following metrics to evaluate utterance diversity. For Japanese word tokenization, we used Sudachi (Takaoka et al., 2018) in C mode (a mode that does not split compound words).

Average and S.D. of utterance lengths The average and standard deviation of the number of words

Condition	Personality	#Dialogues	Ave. utterance length (S.D.)		Total words	Unique words	Unique bigrams	TTR
BL		100	28.2	(7.7)	42,260	15,747	31,620	.373
noPT		100	28.5	(7.0)	42,784	16,362	32,730	.382
APM	All	100	28.4	(7.2)	42,563	16,089	32,334	.378
	High	50	30.1	(7.6)	22,547	8,406	17,098	.373
	Low	50	26.7	(6.3)	20,016	7,683	15,236	.384
EL	All	100	36.0	(18.2)	53,951	18,556	39,010	.344
	High	50	50.7	(13.9)	38,001	12,175	26,883	.320
	Low	50	21.3	(5.7)	15,950	6,381	12,127	.400
APM+EL	All	100	31.4	(13.4)	47,111	16,856	34,839	.358
	High+High	25	46.2	(12.3)	17,308	5,624	12,309	.325
	High+Low	25	23.6	(4.8)	8,836	3,451	6,787	.391
	Low+High	25	35.2	(9.9)	13,184	4,651	9,796	.353
	Low+Low	25	20.8	(5.9)	7,783	3,130	5,947	.402

Condition	Personality	Unique CW	CW-TTR	SE	CE	MTLD	MSTTR	Ave. TTR (S.D.)	
BL		1,917	.106	7.52	3.44	49.6	0.719	.373	(.023)
noPT		2,282	.122	7.62	3.46	51.4	0.725	.384	(.023)
APM	All	2,050	.112	7.59	3.47	50.4	0.725	.379	(.020)
	High	1,543	.161	7.58	3.37	52.6	0.733	.374	(.020)
	Low	1,418	.162	7.46	3.20	48.9	0.719	.385	(.020)
EL	All	2,428	.104	7.64	3.62	51.7	0.731	.362	(.046)
	High	2,127	.131	7.67	3.61	56.5	0.746	.323	(.024)
	Low	1,204	.171	7.27	2.96	42.8	0.696	.401	(.024)
APM+EL	All	2,257	.112	7.62	3.60	50.4	0.727	.369	(.037)
	High+High	1,440	.198	7.59	3.42	54.6	0.741	.327	(.020)
	High+Low	908	.242	7.38	3.00	48.3	0.720	.391	(.021)
	Low+High	1,243	.220	7.50	3.26	51.8	0.731	.356	(.026)
	Low+Low	803	.235	7.15	2.74	43.4	0.698	.403	(.024)

Table 1: Diversity metrics of the simulations for the travel interview system. For example, the row where Condition is APM and Personality is All represents metrics calculated from all the dialogues under the APM condition. The subsequent row, where Personality is High, shows metrics calculated from the dialogues under the APM condition using a high degree of anthropomorphism in the personality setting. The row where Condition is APM+EL and Personality is High+High represents metrics calculated from the dialogues with a high degree of anthropomorphism and a high degree of elaborateness. Bold numbers are mentioned in the main text.

per user utterance across all dialogues.

Total words The total number of words across all user utterances in all dialogues.

Unique words The number of unique words across all user utterances in all dialogues.

Unique bigrams The number of unique bigrams across all user utterances in all dialogues.

TTR Type-Token Ratio: ((# of unique words) / (# of total words)) for all user utterances in all dialogues.

Unique CW The number of Unique Content Words appearing in all user utterances.

CW-TTR The type-token ratio for content words across all user utterances.

SE Shannon Entropy calculated from all user utterances across all dialogues.

CE Conditional bigram Entropy calculated from all user utterances across all dialogues.

MTLD Measure of Textual Lexical Diversity (McCarthy and Jarvis, 2010) calculated by concatenating all user utterances (threshold: 0.72).

MSTTR Mean Segmental Type-Token Ratio (McCarthy and Jarvis, 2010) calculated by concatenating all user utterances.

Average and S.D. of TTR The average and standard deviation of type-token ratios calculated per dialogue.

4.5 Results

Tables 1 and 2 respectively show the diversity metrics for the travel interview system and the sweets interview system. Results for each personality trait condition are also presented. These results suggest the following two points: (1) noPT exhibits greater

Condition	Personality	#Dialogues	Ave. utterance length (S.D.)		Total words	Unique words	Unique bigrams	TTR
BL		100	25.9	(7.8)	25,152	11,138	20,534	.443
noPT		100	25.6	(8.0)	25,509	11,105	20,634	.435
APM	All	100	25.6	(8.2)	25,680	11,124	20,862	.433
	High	50	26.5	(8.2)	13,269	5,723	10,824	.431
	Low	50	24.6	(8.0)	12,411	5,401	10,038	.435
EL	All	100	33.9	(17.7)	34,020	13,214	26,117	.388
	High	50	47.5	(14.8)	23,755	8,578	17,808	.361
	Low	50	20.3	(6.1)	10,265	4,636	8,309	.452
APM+EL	All	100	28.1	(12.2)	28,067	11,549	22,182	.411
	High+High	25	39.4	(13.1)	9,857	3,743	7,643	.380
	High+Low	25	22.2	(6.7)	5,560	2,458	4,533	.442
	Low+High	25	30.6	(10.5)	7,642	3,087	6,015	.404
	Low+Low	25	20.0	(5.9)	5,008	2,261	3,991	.451

Condition	Personality	Unique CW	CW-TTR	SE	CE	MTLD	MSTTR	Ave. TTR (S.D.)	
BL		1,157	.109	7.35	3.09	53.4	0.732	.444	(.023)
noPT		1,458	.133	7.43	3.18	53.9	0.736	.437	(.029)
APM	All	1,547	.141	7.51	3.33	53.2	0.737	.435	(.024)
	High	1,090	.193	7.45	3.15	56.6	0.740	.432	(.020)
	Low	1,081	.202	7.41	3.11	51.3	0.724	.438	(.028)
EL	All	1,838	.127	7.58	3.45	54.2	0.740	.409	(.053)
	High	1,553	.156	7.58	3.41	57.2	0.753	.364	(.027)
	Low	933	.208	7.24	2.90	47.3	0.715	.454	(.028)
APM+EL	All	1,604	.134	7.51	3.38	52.7	0.731	.421	(.038)
	High+High	981	.240	7.49	3.16	56.5	0.750	.383	(.026)
	High+Low	649	.272	7.20	2.85	49.8	0.729	.443	(.025)
	Low+High	808	.246	7.33	2.94	52.6	0.726	.407	(.032)
	Low+Low	628	.284	7.08	2.62	47.0	0.709	.452	(.024)

Table 2: Diversity metrics of the simulations for the sweets interview system.

content variation compared to BL. In other words, persona generation using LLMs can increase the variety of content in utterances. (2) EL shows greater stylistic variation than noPT. That is, by varying redundancy levels, one can increase stylistic diversity in utterances.

(1) is suggested by the following findings: Comparing BL and noPT, especially in the travel domain, CW-TTR increases from .106 to .122 in the travel domain, and in the sweets domain, from .109 to .133. This suggests that the generated personas contain different content words from the seed personas. This may help uncover issues caused by diverse utterance content. Although SE, CE, and MTLD also show slight improvements, the differences are not significant. TTR slightly decreases in the sweets domain system, possibly due to limited stylistic changes, leading to low variation in function words.

(2) is suggested by the following findings: Compared to noPT, EL (All) shows an increase in the standard deviation of utterance length—from 7.0 to 18.2 in the travel domain and from 8.0 to 17.7 in the sweets domain. This suggests that varying the

degrees of elaborateness enables the generation of utterances with diverse lengths, potentially exposing problems caused by such diversity. TTR and CW-TTR are lower, likely because elaborate utterances under the EL High condition tend to include fixed phrases.

There is little difference between APM and noPT. Also, comparing APM Low and High settings reveals no significant variation. Furthermore, APM+EL does not differ much from EL. Even when examining the dialogues, major differences were not observed. This suggests that the personality traits described in APM may not be sufficient to alter dialogue style. Since personas generated in noPT already contain variation in personality, APM might not have introduced additional diversity beyond that.

Note that in settings like EL High, where personality traits are narrowly defined, TTR and CW-TTR increase. This is likely because the number of dialogues is small, reducing word repetition.

noPT:

- Name: Erina Takahashi
- 29 years old
- Female
- Lives in Kyoto Prefecture
- Lives alone
- Works for an IT company
- Reserved, but talkative with close friends
- Takes several trips a year
- Most recently visited Kanazawa; traveled alone
- Enjoys visiting local art museums and cafes while traveling
- Would like to visit Paris, France
- Books flights and hotels online

EL High:

- Name: Ken Nakamura
- 34 years old
- Male
- Live in Hiroshima Prefecture
- Lives with his wife and 2-year-old son
- Public servant
- Speaks gently and is family-oriented
- Travels less frequently since having a child, but still goes on one trip per year
- Most recently visited Kagawa; a family trip focused on touring udon restaurants
- Chooses travel destinations that can be enjoyed by the whole family
- Would like to visit Hokkaido
- Plans trips together with his wife

Figure 8: Examples of generated personas (1).

4.6 Generated Personas

Examples of the generated personas are shown in Figures 8, 9, and 10. These examples were randomly selected from the personas generated under each condition. Figures 9 and 10 feature only male personas, but this is purely coincidental. While the male personas shown speak in a reserved manner, some of the other generated male personas are sociable and talkative. From these examples, it can be inferred that the generated personas are consistent with their respective personality traits. The APM+EL examples are omitted, as they did not yield any notable results with respect to the diversity metrics. Additional examples of generated personas are shown in Figures 17 and 18 in the Appendix.

5 Discussion

As discussed above, the experimental results suggest that generating personas with an LLM enables the generation of dialogues with greater variation. Although the quantitative metrics do not indicate this very clearly, you can see the variation by looking at the generated dialogues. This could lead to the discovery of more problems in dialogue systems. Furthermore, specifying personality traits related to communication style can introduce greater variation in utterance lengths, which may help uncover additional problems.

In the current user simulator prompt we used, relatively similar utterances tend to be generated consecutively. This might be different from the behaviors of human users. It is possible that generating various types of utterances within a single dialogue could help reveal different types of issues. In future work, we will explore methods for increas-

EL Low:

- Name: Yusuke Ogawa
- 45 years old
- Male
- Lives in Hiroshima Prefecture
- Lives with his wife and two children (a son in 8th grade and a daughter in 5th grade)
- Public servant
- Speaks calmly and avoids unnecessary talk
- Goes on one trip per year
- Most recently visited Kyoto on a family trip
- Visits temples and shrines during travels
- Would like to visit Nara
- Makes travel reservations by calling a travel agency

Figure 9: Examples of generated personas (2).

ing such variation and investigate to what extent the behaviors of the user simulator cover those of human users.

Although our study focused on APM and EL traits, there may be other personality traits that could introduce further variation. Identifying and incorporating such traits will be considered in future research.

The ultimate goal of our study is to identify as many issues in interview dialogue systems as possible. However, in this work, we only measured the diversity of simulated users. Whether the proposed method can help discover issues across various dialogue systems remains a topic for future work. In the future, we aim to integrate this method into tools for building interview dialogue systems and validate it through its application in the development of diverse systems.

Note that, although the systems used in our evaluation experiments did not exhibit obvious issues such as completely incoherent utterances, we found that under the EL High condition, simulated users tended to engage in extended small talk. This often

APM High:

- Name: Kenta Omura
- 24 years old
- Male
- Lives in Osaka Prefecture
- Lives alone
- Graduate student
- Quiet and not very talkative, but polite
- Takes about one trip per year
- Most recently visited Nara; did some sightseeing after attending an academic conference
- Enjoys exploring nature at travel destinations
- Would like to visit the Rocky Mountains in Canada
- Books flights online by himself

APM Low:

- Name: Naoto Yamashita
- 33 years old
- Male
- Lives in Hiroshima Prefecture
- Lives alone
- Designer
- Speaks calmly and politely
- Takes two trips a year
- Most recently visited Kyoto alone
- Enjoys visiting architectural sites and art museums while traveling
- Would like to visit Paris
- Plans and books trips meticulously online

Figure 10: Examples of generated personas (3).

caused the system to fail in extracting the necessary information, revealing a potential issue.

6 Concluding Remarks

This paper proposed a persona generation method for user simulation using large language models (LLMs). We also introduced a method to incorporate personality traits related to communication style during persona generation. This enables diverse testing of interview dialogue systems without human involvement, significantly reducing development costs. Although the evaluation experiments were limited and some issues remain, the results suggest the effectiveness of the proposed method. Therefore, we believe it is worthwhile to share our proposed method and experimental findings.

In the future, we plan to develop user simulators for interview dialogue systems with speech input/output and multimodal input/output. We also aim to extend our approach to build user simulators for various types of dialogue systems beyond interview dialogues.

Acknowledgments

This work was partially inspired by the collaboration between Ekai Hashimoto, Takayoshi Sakurai, and Shun Shiramatsu of Nagoya Institute of Technology, Toshitake Komazaki of Tokyo Healthcare University, Shiho Tsuchiya of Kitasato University Hospital, and the first author. We sincerely thank them.

Limitations

In addition to what was discussed in Section 5, there are several other limitations in our evaluation experiments. First, we used only a single model as the LLM for both persona generation and the simulator. Second, we used only Japanese systems. Furthermore, the number of generated personas per condition was fixed at 100, the number of seed personas was set to 10 for both systems, and the temperature parameter was set to 0.7 throughout the experiment; other configurations have not been tested. The impact of varying these conditions remains unexamined.

In future work, we plan to address these issues as we apply the proposed method to the development of various interview dialogue systems.

Ethical Considerations

In our study, human participants were not involved in system testing; therefore, there is no risk of collecting or leaking personal information from individuals.

The biases potentially inherent in LLMs (Gallegos et al., 2024) may be reflected in automatically generated personas or user utterances. As a result, certain types of personas or utterances may be excluded from testing, which could hinder the identification of potential issues. Verifying whether this problem occurs and addressing it if necessary remain future challenges.

References

- Hua Ai and Diane Litman. 2009. [Setting up user action probabilities in user simulations for dialog system development](#). In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 888–896, Suntec, Singapore. Association for Computational Linguistics.
- Atheer Algherairy and Moataz Ahmed. 2025. [Prompting large language models for user simulation in task-](#)

- oriented dialogue systems. *Computer Speech & Language*, 89:101697.
- Yoshihiko Asao, Julien Kloetzer, Junta Mizuno, Dai Saiki, Kazuma Kadowaki, and Kentaro Torisawa. 2020. [Understanding user utterances in a dialog system for caregiving](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 653–661, Marseille, France. European Language Resources Association.
- Soodeh Atefi and Mohammad Amin Alipour. 2019. [An automated testing framework for conversational agents](#). *Preprint*, arXiv:1902.06193.
- Lars Degerstedt and Arne Jönsson. 2006. [Lintest: a development tool for testing dialogue systems](#). In *Proceedings of Interspeech 2006*, pages 225–235.
- David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, Gale Lucas, Stacy Marsella, Fabrizio Morbini, Angela Nazarian, Stefan Scherer, Giota Stratou, Apar Suri, David Traum, Rachel Wood, Yuyu Xu, Albert Rizzo, and Louis-Philippe Morency. 2014. [SimSensei kiosk: a virtual human interviewer for healthcare decision support](#). In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14*, pages 1061–1068, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.
- Martina Di Bratto, Antonio Origlia, Maria Di Maro, and Sabrina Mennella. 2024. [Linguistics-based dialogue simulations to evaluate argumentative conversational recommender systems](#). *User Modeling and User-Adapted Interaction*, 34(5):1581–1611.
- W. Eckert, E. Levin, and R. Pieraccini. 1997. [User modeling for spoken dialogue system evaluation](#). In *1997 IEEE Workshop on Automatic Speech Recognition and Understanding Proceedings*, pages 80–87.
- Sarah E. Finch, Ellie S. Paek, and Jinho D. Choi. 2023. [Leveraging large language models for automated dialogue analysis](#). In *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 202–215, Prague, Czechia. Association for Computational Linguistics.
- Isabel O. Gallegos, Ryan A. Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck Dernoncourt, Tong Yu, Ruiyi Zhang, and Nesreen K. Ahmed. 2024. [Bias and fairness in large language models: A survey](#). *Computational Linguistics*, 50(3):1097–1179.
- Tao Ge, Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. 2024. [Scaling synthetic data creation with 1,000,000,000 personas](#). *Preprint*, arXiv:2406.20094.
- Patrick Gebhard, Tobias Baur, Ionut Damian, Gregor Mehlmann, Johannes Wagner, and Elisabeth André. 2014. [Exploring interaction strategies for virtual characters to induce stress in simulated job interviews](#). In *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS '14*, pages 661–668, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.
- Kallirroi Georgila, Maria Wolters, and Johanna Moore. 2008. [Simulating the behaviour of older versus younger users when interacting with spoken dialogue systems](#). In *Proceedings of ACL-08: HLT, Short Papers*, pages 49–52, Columbus, Ohio. Association for Computational Linguistics.
- Kallirroi Georgila, Maria Wolters, and Johanna Moore. 2010. [Learning dialogue strategies from older and younger simulated users](#). In *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 103–106, Tokyo, Japan. Association for Computational Linguistics.
- Pablo Gómez-Abajo, Sara Pérez-Soler, Pablo C. Cañizares, Esther Guerra, and Juan de Lara. 2024. [Mutation testing for task-oriented chatbots](#). In *Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering, EASE '24*, pages 232–241, New York, NY, USA. Association for Computing Machinery.
- Guoxiang Guo, Aldeida Aleti, Neelofar Neelofar, and Chakkrit Tantithamthavorn. 2024. [Mortar: Metamorphic multi-turn testing for llm-based dialogue systems](#). *Preprint*, arXiv:2412.15557.
- Ekai Hashimoto, Mikio Nakano, Takayoshi Sakurai, Shun Shiramatsu, Toshitake Komazaki, and Shiho Tsuchiya. 2025. [A career interview dialogue system using large language model-based dynamic slot generation](#). In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 1562–1584, Abu Dhabi, UAE. Association for Computational Linguistics.
- Mengze Hong, Chen Jason Zhang, Chaotao Chen, Rongzhong Lian, and Di Jiang. 2025. [Dialogue language model with large-scale persona data engineering](#). In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 3: Industry Track)*, pages 961–970, Albuquerque, New Mexico. Association for Computational Linguistics.
- Mohammed (Ehsan) Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W. Picard. 2013. [MACH: my automated conversation coach](#). In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '13*, pages 697–706, New York, NY, USA. Association for Computing Machinery.
- Zhiyuan Hu, Yue Feng, Anh Tuan Luu, Bryan Hooi, and Aldo Lipani. 2023. [Unlocking the potential of user feedback: Leveraging large language model as user simulators to enhance dialogue system](#). In *Proceedings of the 32nd ACM International Conference*

- on Information and Knowledge Management, CIKM '23, pages 3953–3957, New York, NY, USA. Association for Computing Machinery.
- Koji Inoue, Kohei Hara, Divesh Lala, Kenta Yamamoto, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara. 2020. [Job interviewer android with elaborate follow-up question generation](#). In *Proceedings of the 2020 International Conference on Multimodal Interaction, ICMI '20*, pages 324–332, New York, NY, USA. Association for Computing Machinery.
- Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy, and Jad Kabbara. 2024. [PersonaLLM: Investigating the ability of large language models to express personality traits](#). In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 3605–3627, Mexico City, Mexico. Association for Computational Linguistics.
- Michael Johnston, Patrick Ehlen, Frederick G. Conrad, Michael F. Schober, Christopher Antoun, Stefanie Fail, Andrew Hupp, Lucas Vickers, Huiying Yan, and Chan Zhang. 2013. [Spoken dialog systems for automated survey interviewing](#). In *Proceedings of the 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 329–333, Metz, France. Association for Computational Linguistics.
- Hazaël Jones and Nicolas Sabouret. 2012. [An affective model for a virtual recruiter in a job interview context](#). *Procedia Computer Science*, 15:312–313. 4th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES'12).
- Takahiro Kobori, Mikio Nakano, and Tomoaki Nakamura. 2016. [Small talk improves user impressions of interview dialogue systems](#). In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 370–380, Los Angeles. Association for Computational Linguistics.
- Max M. Lang and Sol Eskenazi. 2025. [Telephone surveys meet conversational AI: Evaluating a llm-based telephone survey system at scale](#). *Preprint*, arXiv:2502.20140.
- Xiaomin Li, Chuanqi Tao, Jerry Gao, and Hongjing Guo. 2022. [A review of quality assurance research of dialogue systems](#). In *2022 IEEE International Conference On Artificial Intelligence Testing (AITest)*, pages 87–94.
- Xiujun Li, Zachary C. Lipton, Bhuwan Dhingra, Lihong Li, Jianfeng Gao, and Yun-Nung Chen. 2017. [A user simulator for task-completion dialogues](#). *Preprint*, arXiv:1612.05688.
- Yajiao Liu, Xin Jiang, Yichun Yin, Yasheng Wang, Fei Mi, Qun Liu, Xiang Wan, and Benyou Wang. 2023. [One cannot stand for everyone! Leveraging multiple user simulators to train task-oriented dialogue systems](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1–21, Toronto, Canada. Association for Computational Linguistics.
- Xiang Luo, Zhiwen Tang, Jin Wang, and Xuejie Zhang. 2024. [DuetSim: Building user simulator with dual large language models for task-oriented dialogues](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 5414–5424, Torino, Italia. ELRA and ICCL.
- Pierre-Emmanuel Mazaré, Samuel Humeau, Martin Raison, and Antoine Bordes. 2018. [Training millions of personalized dialogue agents](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2775–2779, Brussels, Belgium. Association for Computational Linguistics.
- Philip M McCarthy and Scott Jarvis. 2010. [MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment](#). *Behavior research methods*, 42(2):381–392.
- Juliana Miehe, Isabel Feustel, Julia Hornauer, Wolfgang Minker, and Stefan Ultes. 2020. [Estimating user communication styles for spoken dialogue systems](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 540–548, Marseille, France. European Language Resources Association.
- Mikio Nakano and Kazunori Komatani. 2024. [DialBB: A dialogue system development framework as an educational material](#). In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 664–668, Kyoto, Japan. Association for Computational Linguistics.
- Yasuhisa Niimi and Takuya Nishimoto. 1999. [Mathematical analysis of dialogue control strategies](#). In *Proceedings of the 6th European Conference on Speech Communication and Technology*, pages 1403–1406.
- O. Pietquin and T. Dutoit. 2006. [A probabilistic framework for dialog simulation and optimal strategy learning](#). *IEEE Transactions on Audio, Speech, and Language Processing*, 14(2):589–599.
- Olivier Pietquin and Helen Hastie. 2013. [A survey on metrics for the evaluation of user simulations](#). *The knowledge engineering review*, 28(1):59–73.
- Louisa Pragst, Wolfgang Minker, and Stefan Ultes. 2019. [Exploring the Applicability of Elaborateness and Indirectness in Dialogue Management](#), pages 189–198. Springer International Publishing, Cham.
- Pooja S. B. Rao, Manish Agnihotri, and Dinesh Babu Jayagopi. 2020. [Automatic follow-up question generation for asynchronous interviews](#). In *Proceedings of the Workshop on Intelligent Information Processing and Natural Language Generation*, pages 10–20, Santiago de Compostela, Spain. Association for Computational Linguistics.
- Byron Reeves and Clifford Nass. 1996. *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.

- Jost Schatzmann, Blaise Thomson, Karl Weilhammer, Hui Ye, and Steve Young. 2007. [Agenda-based user simulation for bootstrapping a POMDP dialogue system](#). In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, pages 149–152, Rochester, New York. Association for Computational Linguistics.
- Jost Schatzmann, Karl Weilhammer, Matt Stuttle, and Steve Young. 2006. [A survey of statistical user simulation techniques for reinforcement-learning of dialogue management strategies](#). *The Knowledge Engineering Review*, 21(2):97–126.
- Ivan Sekulic, Silvia Terragni, Victor Guimarães, Nghia Khau, Bruna Guedes, Modestas Filipavicius, Andre Ferreira Manso, and Roland Mathis. 2024. [Reliable LLM-based user simulator for task-oriented dialogue systems](#). In *Proceedings of the 1st Workshop on Simulating Conversational Intelligence in Chat (SCI-CHAT 2024)*, pages 19–35, St. Julians, Malta. Association for Computational Linguistics.
- Karim Sidaoui, Matti Jaakkola, and Jamie Burton. 2020. [AI feel you: customer experience assessment via chatbot interviews](#). *Journal of Service Management*, 31(4):745–766.
- Gabriel Skantze, Samer Al Moubayed, Joakim Gustafson, Jonas Beskow, and Björn Granström. 2012. Furhat at robotville: A robot head harvesting the thoughts of the public through multi-party dialogue. In *Proceedings of the Workshop on Real-time Conversations with Virtual Agents in conjunction with the International Conference on Intelligent Virtual Agents*.
- Amanda Stent, Svetlana Stenchikova, and Matthew Marge. 2006. [Dialog systems for surveys: the rate-a-course system](#). In *2006 IEEE Spoken Language Technology Workshop*, pages 210–213.
- Ming-Hsiang Su, Chung-Hsien Wu, and Yi Chang. 2019. [Follow-up question generation using neural tensor network-based domain ontology population in an interview coaching system](#). In *Interspeech 2019*, pages 4185–4189.
- Weiwei Sun, Shuyu Guo, Shuo Zhang, Pengjie Ren, Zhumin Chen, Maarten de Rijke, and Zhaochun Ren. 2023. [Metaphorical user simulators for evaluating task-oriented dialogue systems](#). *ACM Trans. Inf. Syst.*, 42(1).
- Kazuma Takaoka, Sorami Hisamoto, Noriko Kawahara, Miho Sakamoto, Yoshitaka Uchida, and Yuji Matsumoto. 2018. [Sudachi: a Japanese tokenizer for business](#). In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).
- Yoshiki Tanaka and Michimasa Inaba. 2024. [User review writing via interview with dialogue systems](#). In *Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 428–439, Kyoto, Japan. Association for Computational Linguistics.
- Silvia Terragni, Modestas Filipavicius, Nghia Khau, Bruna Guedes, André Manso, and Roland Mathis. 2023. [In-context learning user simulators for task-oriented dialog systems](#). *Preprint*, arXiv:2306.00774.
- Jie Zeng, Yukiko Nakano, and Tatsuya Sakato. 2023. [Question generation to elicit users’ food preferences by considering the semantic content](#). In *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 190–196, Prague, Czechia. Association for Computational Linguistics.
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. [Personalizing dialogue agents: I have a dog, do you have pets too?](#) In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia. Association for Computational Linguistics.

A Example Dialogues from the Travel Interview System

Example dialogues from the travel interview system are shown in Figures 11–16. Figure 11 shows a result from testing with a simulator using the lower persona from Figure 3. Figures 12–16 show results from testing with simulators using the generated personas in Figures 8, 9, and 10.

B Additional Examples of Generated Personas

Figures 17 and 18 show additional examples of generated personas. These were also randomly selected.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User About once a year, I'd say. I recently went to Osaka. I strolled around on my own and enjoyed trying out different foods.

System Eating your way through Osaka sounds fun! Was there any dish that left a strong impression on you?

User Definitely the takoyaki. I had some at a local shop—the outside was crispy and the inside was soft and gooey. It was amazing. The kushikatsu was also delicious.

System Takoyaki and kushikatsu both sound tasty! Have there been any other travel destinations that you particularly enjoyed?

User Kyoto was great too. I really liked the historic atmosphere of the streets, and I tried a variety of traditional Japanese sweets.

System The historical streets of Kyoto sound wonderful! Were there any sweets that especially stood out to you?

User The dorayaki left the biggest impression. The fluffy pancake and the mildly sweet bean paste were a perfect match. I also enjoyed yatsushashi.

System Dorayaki and yatsushashi—both sound delicious! What do you enjoy most when traveling?

User Food tours are my favorite. I love discovering local flavors unique to each area. I do a bit of sightseeing too.

System So eating local food is the most fun part! Where would you like to go next?

User I'd love to visit Singapore. I want to try chicken rice and hawker stall dishes. I'm also looking forward to experiencing the atmosphere of the city.

System Hawker food in Singapore is definitely appealing! How do you usually make your travel arrangements?

User I usually book everything online. It's convenient to search for flights and hotels all in one place.

System Online booking really is convenient! Which sites do you often use?

User Mostly the major travel websites. They make it easy to compare options, and I always check for campaign deals. I go with the ones I find easiest to use.

System Comparing travel sites is important! How do you keep track of special deals and offers?

User I use site notifications and subscribe to email newsletters. I check them regularly so I don't miss any good deals.

System It's great that you keep on top of special offers! Is there anything else you pay attention to when planning a trip?

User I make a list of places I want to visit, and I try to plan carefully around time and budget. I also check the weather.

System Making a list sounds like a smart idea! Have you started planning your next trip?

User Not in detail yet, but I'm hoping to go to Singapore next year. I'm looking forward to planning it out!

System Sounds exciting! Are there any other countries or regions you'd like to visit?

User I'm also interested in Europe. I'd especially like to try authentic cuisine in Italy and Spain.

System European cuisine is certainly appealing! Are there any specific dishes you'd like to try in Italy or Spain?

User In Italy, I want to try pizza and pasta. In Spain, definitely paella and tapas. I'm especially interested in dishes that use local ingredients.

System Local ingredient-based dishes are wonderful! How do you choose where to eat when you travel?

User I usually look up popular places or restaurants loved by locals. I also rely on reviews and word-of-mouth recommendations.

System Choosing popular local spots sounds like a great strategy! Is there anything else you look forward to when traveling?

User I enjoy experiencing the local scenery and culture too. I love immersing myself in the unique atmosphere of each place.

System Experiencing the scenery and culture is definitely rewarding! Was there a place that left a particularly strong impression on you?

Figure 11: Example simulated dialogue in condition BL.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User I travel a few times a year. I recently took a solo trip to Kanazawa.

System Kanazawa is a lovely place! Was there anything in particular that left an impression?

User Yes, the exhibition I saw at a local art museum was very memorable. I also really enjoyed visiting different cafes.

System Visiting art museums and cafes sounds wonderful! Have there been other destinations you enjoyed?

User Yes, though it's close to Kyoto, I also liked Nara. I appreciated the calm and quiet atmosphere.

System The peaceful atmosphere of Nara is indeed charming! What do you enjoy doing when you travel?

User I like visiting places where I can experience local culture and history. I also enjoy taking photographs.

System Experiencing local culture and history is great! Where would you like to travel next?

User I'd like to visit Paris, France. I want to fully enjoy the museums and the cityscape.

System Paris is a fascinating destination! How do you usually make your travel reservations?

User I usually book flights and hotels online. It's convenient and makes it easy to compare options.

System Online booking is certainly convenient! Which websites do you use most often?

User I usually use major travel booking sites. I appreciate the abundance of user reviews.

System Having plenty of reviews is important! Do you consult any other sources for information?

User I also refer to travel blogs and social media posts. Photos in particular help me get a better sense of what to expect.

System Travel blogs and social media are great sources! Which social media platforms do you use most?

User I mainly use Instagram. There are so many travel photos that really spark my imagination.

System Instagram is perfect for travel inspiration! Are there any other countries you'd like to visit?

User I'd like to visit Italy too. Especially to tour museums in Florence and Rome.

System Touring museums in Italy sounds fantastic! Are there any specific museums you're looking forward to?

User The Uffizi Gallery is at the top of my list. I'd love to take my time viewing Renaissance masterpieces.

System The Uffizi is a great choice! Are there any other artworks you're interested in?

User I'm also interested in Impressionist works by Monet and Renoir. Their use of light is especially beautiful.

System Monet and Renoir's works are lovely indeed! Are there any other artists you like?

User I also like Klimt. The use of gold leaf in "The Kiss" is particularly striking.

System Klimt's "The Kiss" is truly a beautiful piece! Are there other museums you'd like to visit?

User I'd like to go to the Orsay Museum as well. They have many Impressionist works, and I'd love to explore them thoroughly.

System The Orsay Museum is also a wonderful choice! Besides museum visits, is there anything else you're looking forward to?

User I'd like to stroll through the streets of Paris at a relaxed pace. I'm also looking forward to spending some leisure time at cafes.

System Spending time at cafes sounds lovely! Are there any cafes in particular you'd like to visit?

Figure 12: Example simulated dialogue in condition noPT.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User Not so frequently these days since our child is still small, but we try to take a family trip about once a year.

System A yearly family trip sounds wonderful! Where did you go most recently?

User We went to Kagawa recently. We toured several udon restaurants as a family, and our son really seemed to enjoy it. The scenery in Kagawa was beautiful and very relaxing as well.

System A tour of udon restaurants in Kagawa sounds fun! Were there any other places you've enjoyed visiting?

User Yes, a little while ago we went to Kyoto. It has many historical sites, and my wife and I enjoyed strolling around. That was before our son was born, so we were able to take it easy.

System The historical sites in Kyoto are wonderful! What do you find most enjoyable when you travel?

User The most important thing is doing something that the whole family can enjoy together. Lately, we've been choosing destinations and activities our son would like. We enjoy sightseeing, but discovering delicious local food is also a big part of the fun.

System Family-friendly activities and good food—that sounds lovely! Is there somewhere you'd like to go next?

User I'd like to visit Hokkaido. The vast natural landscapes and delicious food are really appealing. I also hope we can experience the snowy scenery as a family someday.

System Hokkaido is a great choice! How do you usually plan your trips?

User I usually plan the trip together with my wife. We research online and check reviews. We try to be thorough so that everyone in the family can enjoy the trip.

System Online research is helpful! Reviews are especially valuable. Which websites do you use?

User We often use travel review sites and accommodation booking sites. We tend to look for sites that have a lot of family-friendly information. We also find blogs and social media posts from people who've actually been there quite helpful.

System Review sites and social media are very useful! Which site has been the most helpful?

User Recently, we've been using Jalan and Rakuten Travel. They offer many family-friendly plans, and the reviews are detailed and informative. Reading about others' experiences really builds anticipation.

System Jalan and Rakuten Travel are great options! Are there any other destinations you're interested in?

User Yes, we're also interested in Okinawa. The ocean is beautiful, and it seems like a great place to relax as a family. Once our son is a bit older, we'd love to try snorkeling too.

System Okinawa sounds like a wonderful choice! Snorkeling sounds fun. Is there a specific beach you'd like to visit?

User We haven't decided on a specific beach yet, but we'd like to go somewhere with clear water. A calm beach where our son can play in the sand would be ideal. My wife and I are planning to look into places that are kid-friendly.

System A calm beach sounds perfect for a relaxing family time! Do you have any other memorable travel experiences?

User I remember when our son took his first bullet train ride on a trip. He was a bit nervous at first, but he really enjoyed looking out the window. We also shared a boxed lunch on the train—it was a fun memory for all of us.

System What a lovely memory on the bullet train! Any other travel stories that left an impression?

User When we did the udon tour in Kagawa, our son said, "I want to eat more!" That really stuck with me. It made us all laugh, and the atmosphere was so warm and cheerful. It's often those small, everyday moments during travel that become the most memorable.

System Those little moments during travel really stay with you! Are there any other countries or regions you'd like to visit?

User I'd love to visit Europe, especially Italy and France. I'm drawn to the beautiful cityscapes and historic architecture. I think it would be great to take a leisurely family trip there.

System Europe, especially Italy and France, sounds wonderful! Do you have any specific cities in mind?

User Yes, in Italy, I'd like to visit Florence and Venice. In France, Paris is definitely a dream destination. Both places are rich in history and culture, and I think walking around as a family would be really enjoyable.

System Florence and Venice are excellent choices! In Paris, is there a particular attraction you're most interested in?

User I'd really like to visit the Eiffel Tower and the Louvre Museum. They're famous spots, but I think they're places you have to see at least once. I also think a stroll along the Seine with my family would be lovely.

System The Eiffel Tower and the Louvre sound amazing! Are there any other activities you'd like to try?

Figure 13: Example simulated dialogue in condition EL High.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User I travel about once a year.

System A yearly trip sounds great! Where did you go most recently?

User I went to Kyoto. It was a family trip where we visited various temples and shrines.

System Visiting temples and shrines in Kyoto sounds wonderful! Was there a place that left a strong impression?

User Kiyomizu-dera Temple stood out the most. The scenery was beautiful, and my whole family enjoyed it.

System The view from Kiyomizu-dera is lovely! Were there any other enjoyable activities during the trip?

User Choosing souvenirs with my children was a lot of fun. My daughter especially seemed very happy.

System Picking out souvenirs is always enjoyable! Where would you like to travel next?

User I'm thinking of visiting Nara. I'd like to tour the historic temples and shrines there.

System Nara is a great destination too! How do you usually make your travel reservations?

User I make reservations by calling a travel agency. I feel more secure speaking directly with someone.

System Speaking directly does provide peace of mind. How is the service at the travel agency?

User They're polite and explain things clearly. They're also very flexible with our requests.

System Good service at a travel agency is important! Are there any other points you prioritize when traveling?

User I focus on making sure the whole family has fun, and that transportation and accommodation are smooth and stress-free.

System Ensuring the whole family enjoys the trip is very important! Do you make any special arrangements for transport or lodging?

User We choose transportation that minimizes stress, and we prefer places where the family can relax comfortably together.

System Those are thoughtful considerations! Do you have any memorable travel experiences?

User I remember how excited my kids were when they rode the bullet train for the first time.

System Their first bullet train ride—that must have been special! Any other memorable travel moments?

User One time we all wore yukata and strolled through the garden of our inn. It felt fresh and different from our everyday life.

System Walking in yukata sounds like a unique experience! Are there other places you'd like to visit?

User I'd like to go to Hokkaido. It seems to have rich nature and many places the whole family can enjoy.

System Hokkaido is a great choice! Are there any activities you're especially looking forward to?

User I'd like to go hiking with my family in nature. I've heard there are many scenic places.

System Hiking sounds fantastic! Are there any foods you'd like to try or experiences you'd like to have?

User I'd like to try fresh seafood—especially crab and sea urchin.

System Fresh seafood is very appealing! Would you like help finding recommended restaurants in Hokkaido?

User Thank you, but I prefer to look for restaurants once I'm there. I only do a little research in advance.

System Exploring on-site sounds fun too! Is there anything in particular you pay attention to when preparing for a trip?

Figure 14: Example simulated dialogue in condition EL Low.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User About once a year. I recently went to Nara. It was for a conference, but I did some sightseeing as well.

System Nara is a historic place! Was there anything that particularly impressed you during your sightseeing?

User Todai-ji Temple left a strong impression. I was overwhelmed by the sheer presence of the Great Buddha. I also found the deer very soothing.

System The Great Buddha at Todai-ji really is impressive! Were there any other enjoyable experiences during your travels?

User I enjoy walking in nature, so strolling through Nara's green parks was a lot of fun.

System Walking in nature is wonderful! Where would you like to go next?

User I'd like to visit the Rocky Mountains in Canada. I want to hike through the vast natural landscapes.

System The Rocky Mountains are a fantastic choice! How do you usually book your trips?

User I usually arrange flights and accommodation online by myself. Which websites do you often use?

System So you typically book online. I often use comparison websites myself. Do you have any recommended sites?

User I don't stick to any particular site, but I often use Skyscanner or Expedia.

System Skyscanner and Expedia are quite handy! Is there anything else you pay attention to when planning a trip?

User I try to plan within my budget and avoid overdoing it. I also make sure to check local transportation in advance.

System Checking your budget and transport options is important! How often do you usually travel?

User About once a year. I can't go that often, so I try to plan thoroughly when I do.

System It's great that you plan thoroughly! Is there a reason you can't travel more often?

User I don't have much free time. My graduate research keeps me pretty busy.

System It sounds like your research takes a lot of your time. It must be hard to secure quality time. How do you plan your trips?

User I plan them when my research schedule is more relaxed. I try to decide on the destination and dates early.

System Planning early is a good idea! Do you already have specific dates in mind for your next trip?

User Not yet, but I'll start thinking about it once things calm down with my research.

System Once your research calms down, how long are you thinking of traveling for?

User I'm thinking about a week. If it's too long, it could affect my research, so a balanced length is ideal.

System A week-long trip sounds great! Do you have a budget in mind?

User I haven't decided exactly, but I'm estimating around 200,000 yen, including flights and accommodation.

System Around 200,000 yen—got it. Do you have any specific activities in mind for the trip?

User I think the main activity will be hiking in nature. I'd also like to try some local food.

System Enjoying local food sounds wonderful! Is there anything specific you'd like to eat?

User Nothing in particular, but I'd like to try local specialties and dishes made with fresh ingredients.

System Enjoying local specialties is a great idea! Is there anything else you look forward to when traveling?

Figure 15: Example simulated dialogue in condition APM High.

System Hello, my name is Shota Watanabe. I'd like to interview you today about travel. Do you travel often?

User I travel about twice a year. Most recently, I went to Kyoto.

System Kyoto is a wonderful place! Was there anything that left a particular impression on you?

User I toured various architectural sites and museums alone, and the Five-Storied Pagoda at To-ji Temple was especially memorable.

System The Five-Storied Pagoda at To-ji is indeed impressive! Were there any other places you enjoyed visiting?

User I also liked Onomichi, which is relatively close to Hiroshima. The sloping streets and old townscape were very charming.

System The slopes and townscape of Onomichi sound lovely! What do you enjoy doing when you travel?

User I enjoy visiting architectural sites and museums. It's fascinating to experience the history and design.

System Engaging with history and design is definitely appealing! Is there a place you'd like to visit next?

User I'd like to go to Paris. I especially want to see the Louvre Museum and the Eiffel Tower.

System Paris is a wonderful choice! How do you usually make your reservations?

User I research online and make detailed plans before booking. It helps me avoid wasting time.

System So you're a well-organized traveler! Which websites do you usually use?

User Mainly travel booking sites and official museum websites. I also gather information from review sites.

System Sounds like you plan your trips thoroughly! Which review sites have been particularly helpful?

User It depends on the destination, but I often use TripAdvisor. There's a wealth of information there.

System TripAdvisor is definitely useful! Are there any other sites you recommend?

User I also find Expedia convenient and easy to use. I often book flights and accommodations together.

System Expedia is a good choice too! How do you choose your accommodations?

User I prioritize location and cleanliness. I always make sure to read reviews carefully.

System Choosing accommodations wisely is important! What kind of reviews do you find most helpful?

User Reviews that include detailed impressions and photos from guests are especially useful. I also check the average rating.

System Guest impressions are certainly valuable! Do you already have your next destination in mind?

User Not yet, but I'd like to choose a place where I can engage with history and design again.

System A place rich in history and design sounds exciting! Do you have any specific candidates?

User I'm interested in other European cities, such as Florence and Barcelona.

System Florence and Barcelona are both great choices! Why do you want to visit them?

User I want to experience Renaissance art and architecture in Florence. As for Barcelona, Gaudí's architecture is very appealing. Both cities seem rich in design and history, which fascinates me.

System Florence and Barcelona are both wonderful! Are there any other countries you'd like to visit?

User I'm also interested in Switzerland. I'd like to see landscapes where nature and architecture blend together.

System Switzerland's fusion of nature and architecture sounds amazing! Is there any particular place you'd like to visit there?

Figure 16: Example simulated dialogue in condition APM Low.

noPT:

- Name: Yusuke Matsumoto
- 33 years old
- Male
- Lives in Nagano Prefecture
- Lives with his parents
- Farmer
- Quiet but sincere; good at detailed tasks
- Goes on a trip once a year
- Recently visited Nagoya alone
- Visits markets and direct-from-farm produce stands when traveling
- Would like to visit Spain
- Does his own research and makes reservations online

EL High:

- Name: Kaori Ishikawa
- 48 years old
- Gender: Female
- Lives in Niigata Prefecture
- Lives with her husband and their junior high school-aged daughter
- Nursery school teacher
- Speaks in a gentle tone and is polite in conversation
- Goes on a trip about once a year
- Recently visited Sendai on a family sightseeing trip
- Enjoys soaking in hot springs while traveling
- Would like to visit Kyoto
- Plans trips using travel magazines as a reference

EL Low:

- Name: Misaki Yamauchi
- 47 years old
- Female
- Lives in Nagasaki Prefecture
- Lives with her husband
- Homemaker
- Speaks in a gentle and calm manner
- Travels once a year
- Recently visited Kagoshima with her husband
- Buys local specialties when traveling
- Would like to visit Shikoku
- Uses a travel agency

Figure 17: Additional examples of generated personas (1).

APM High:

- Name: Miho Uchida
- 38 years old
- Female
- Lives in Hyogo Prefecture
- Lives with her husband and their daughter, who is in the first grade of elementary school
- Nurse
- Cheerful and a good listener
- Goes on a trip once a year
- Recently visited Kyoto
- Enjoys taking photos while traveling
- Would like to visit Sapporo
- Makes reservations online

APM Low:

- Name: Kana Kawamura
- 26 years old
- Female
- Lives in Tokyo
- Single and lives alone
- Freelance illustrator
- Has a free-spirited and creative personality
- Travels frequently
- Recently visited Kyoto alone
- Enjoys visiting art museums while traveling
- Would like to visit Florence
- Makes reservations based on information found on social media

Figure 18: Additional examples of generated personas (2).