RaPSIL: A Preference-Guided Interview Agent for Rapport-Aware Self-Disclosure

Kenta Hama, Atsushi Otsuka, Masahiro Mizukami, Hiroaki Sugiyama, Makoto Nakatsuji
NTT Human Informatics Laboratories
Correspondence: kenta.hama@ntt.com

Abstract

Facilitating self-disclosure without causing discomfort remains a difficult task—especially for AI systems. In real-world applications such as career counseling, wellbeing support, and onboarding interviews, eliciting personal information like concerns, goals, and personality traits is essential. However, asking such questions directly often leads to discomfort and disengagement. We address this issue with **RaPSIL** (**R**apport-**a**ware **P**reference-guided **S**elf-disclosure **I**nterview **L**earner), a two-stage LLM-based system that fosters natural, engaging conversations to promote self-disclosure. In the first stage, RaPSIL selectively imitates interviewer utterances that have been evaluated by LLMs for both strategic effectiveness and social sensitivity. It leverages LLMs as multi-perspective judges in this selection process. In the second stage, it conducts self-play simulations, using the Reflexion framework to analyze failures and expand a database with both successful and problematic utterances. This dual learning process allows RaPSIL to go beyond simple imitation, improving its ability to handle sensitive topics naturally by learning from both successful and failed utterances. In a comprehensive evaluation with real users, RaPSIL outperformed baselines in enjoyability, warmth, and willingness to re-engage, while also capturing self-descriptions more accurately. Notably, its impression scores remained stable even during prolonged interactions, demonstrating its ability to balance rapport building with effective information elicitation. These results demonstrate that RaPSIL enables socially aware AI interviewers capable of eliciting sensitive personal information while maintaining user trust and comfort—an essential capability for real-world dialogue systems.

1 Introduction

Simulating humans accurately is an important issue in the NLP and AI fields, especially for applications that aim to replicate human attitudes and behavior (Shengli, 2021; Park et al., 2024). These applications often require deep personal information, such as individual values and life philosophies, which play a key role in improving personalization and supporting self-understanding (Safdari et al., 2023; Otsuka et al., 2023). However, collecting such information can be challenging. Traditional questionnaires impose a cognitive burden on those with limited self-analytical skills (Eckstein et al., 2021), and often yield superficial responses due to social desirability bias (van de Mortel, 2008). While dialogue-based methods such as interviews can elicit more authentic self-disclosures (Leong et al., 2024), they require skilled interviewers and significant time investment, which limits their scalability. This has created a need for scalable, socially aware systems that can promote self-disclosure in natural dialogue.

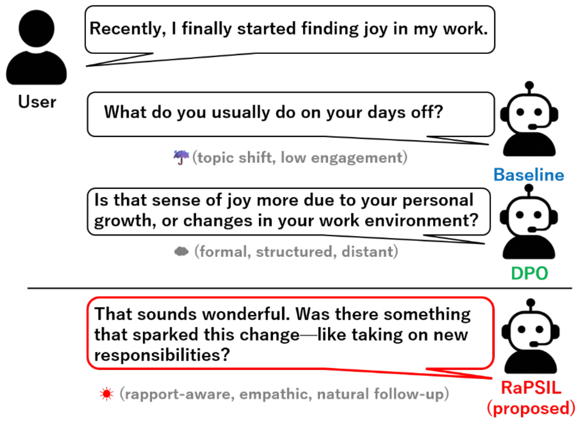


Figure 1: Example dialogue system responses: RaPSIL offers warm, rapport-aware follow-ups encouraging self-disclosure, unlike DPO and baseline.

While dialogue systems can address these scalability challenges, their effective implementation presents unique obstacles. Human interviewers naturally employ strategies of beginning with superficial topics and gradually transitioning to deeper subjects after establishing rapport (Leong et al., 2024). When systems lack this capability, users ex-

Proceedings of the 26th Annual Meeting of SIGDIAL, pages 355–369
Aug 25–27, 2025.

perience discomfort, resulting in decreased engagement and compromised information accuracy (Baihaqi et al., 2024; Kobori et al., 2016). An ideal dialogue system must maintain positive user impressions while successfully eliciting deep self-recognition information, particularly in initial interactions.

Current dialogue systems use various approaches to elicit information: some implement strategic patterns like interspersing casual conversation between questions or gradually progressing from superficial to deeper topics (Kobori et al., 2016; Leong et al., 2024), while others employ imitation learning from human dialogue logs (Sun et al., 2023; Rieser and Lemon, 2011). However, these methods often struggle to adapt to general information-gathering scenarios. While recent advances in LLMs have enabled prompt-based dialogue system design (Qiao et al., 2023), creating systems that can flexibly adapt strategies to maintain user engagement while eliciting deeper personal information remains an area requiring further exploration (Baihaqi et al., 2024; Fiksdal, 1988).

To address this challenge, we propose RaPSIL, a preference-guided interview agent that balances positive user impressions with effective information elicitation through two approaches: (1) selective imitation learning from human-to-human interview dialogue datasets, using an LLM to identify and learn from only highly-rated utterances (Gu et al., 2024), and (2) simulation-enhanced learning that generates diverse LLM-to-LLM dialogues to acquire adaptable conversation strategies despite limited real data (Ulmer et al., 2024; Wang et al., 2024). In addition, failed utterances in the simulations are analyzed using Reflexion (Zhang et al., 2024), and improved versions are stored in a database of few-shot examples. These examples are retrieved in similar contexts to guide the model, enabling it to learn not only from good responses but also from failures in challenging scenarios. Figure 1 illustrates an example of the dialogue strategy adopted by RaPSIL to promote self-disclosure.

We evaluated our approach through experiments with 30 participants, tasking the system to elicit information across five personal categories (interests, dislikes, work, personality traits, and life goals). Among these, the "dislikes," "personality traits," and "life goals" categories are often sensitive in nature and typically require some level of rapport before users feel comfortable disclosing them. Simply asking direct questions about these topics may lead to reduced engagement, even if information is successfully obtained. This makes the task particularly challenging, as systems must manage both strategic questioning and user comfort. The study compared four models: a baseline prompt-based model, a simple imitation learning model, a DPO-optimized model (Rafailov et al., 2023), and RaPSIL. Evaluation included (1) subjective engagement assessment, (2) objective information extraction accuracy using self-description reproducibility, and (3) turn-by-turn strategic quality assessment.

The results strongly supported our hypotheses. RaPSIL successfully elicited personal information while maintaining engagement and establishing rapport. It showed statistically significant improvements in subjective metrics (particularly re-engagement willingness and warmth) and achieved the highest recall and F1 scores in RAGchecker (Ru et al., 2024) evaluations. Unlike other methods that declined in quality over time, our approach maintained high dialogue quality throughout extended interactions, demonstrating its potential for real-world long-term user engagement applications.

The contributions of this study are as follows:

1. We propose RaPSIL, a dialogue system that integrates selective imitation learning and simulation-based refinement to elicit personal information without sacrificing engagement.

2. We design an evaluation framework that jointly measures rapport building and information elicitation in qualitative interviews.

3. We demonstrate that our two-stage approach achieves superior engagement and extraction accuracy compared to existing methods.

2 Related Work

This section explores interview dialogue systems and learning methods for socially adaptive conversational agents.

2.1 Dialogue Systems for Interview Tasks

Traditional interview systems have mimicked semi-structured techniques used in medical counseling and job interviews, relying on pre-designed question flows (Leong et al., 2024; Kobori et al., 2016). A previous study demonstrated that dialogue systems can reduce social pressure compared to human interviewers, thus facilitating self-disclosure (Lucas et al., 2014). While the emergence of large

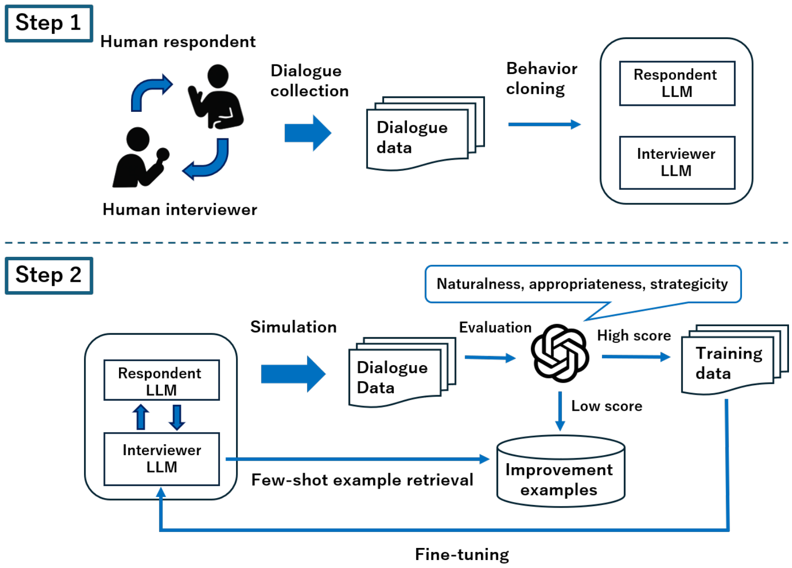


Figure 2: Overview of RaPSIL. In Step 1, an initial model is trained on human interview dialogue data. In Step 2, dialogues generated through simulation are evaluated by an LLM acting as a judge. High-rated utterances are utilized as training data, while low-rated utterances are stored in an improved utterance database and referenced as few-shot examples within the prompt.

language models (LLMs) has enabled more natural conversations, prompt-based implementations have exhibited limitations in capturing complex personality traits and dynamically adapting to individual users (Wang et al., 2023).

Task-oriented dialogue systems prioritize information extraction, while more conversational approaches that incorporate rapport-building strategies have been shown to enhance user engagement (Baihaqi et al., 2024). However, free-form dialogue approaches enhance naturalness but introduce uncertainty in reliability of the acquired information (Fiksdal, 1988). For example, a study on diet tracking interviews showed that mechanical questioning discourages long-term user retention (Tang et al., 2015). Prior research comparing chat-based and question-based dialogue strategies also found the latter more effective in structured interviews (Kobori et al., 2016). Additionally, research on relational agents has emphasized the importance of designing systems for long-term relationship building rather than short-term task completion (Bickmore and Picard, 2005).

In terms of evaluation methodology, traditional approaches have relied either on slot-filling success rates (Qin et al., 2023) or subjective assessments (Gu et al., 2024), but comprehensive techniques that jointly evaluate consistency with self-descriptions and the maintenance of user engagement remain scarce. Recently, RAG-based evaluation methods have shown promise for quantitatively assessing the accuracy of information extraction in qualitative interviews (Ru et al., 2024).

2.2 Learning Approaches for Dialogue Models

While various approaches have been proposed for training dialogue models, each faces distinct limitations. Imitation learning remains a foundational method, but it suffers from the inclusion of both effective and ineffective strategies in training data, which can degrade dialogue quality (Sun et al., 2023; Rieser and Lemon, 2011). Self-reinforcement learning frameworks, such as SOTOPIA-π (Wang et al., 2024), aim to balance task completion with social appropriateness, but the uneven difficulty across personal information categories in interviews makes optimization with a single reward function particularly challenging.

Preference optimization methods like Direct Preference Optimization (DPO) (Rafailov et al., 2023) and failure-based learning mechanisms such as Reflexion (Zhang et al., 2024) have been introduced to address reward design and adaptability issues, though their effectiveness in qualitative interview settings remains underexplored. Similarly, the LLM-as-a-judge approach has demonstrated correlation with human evaluation in dialogue contexts (Gu et al., 2024), but its application to information elicitation tasks is still under investigation.

Simulation-based learning and data augmentation are also gaining traction. Dialogic (Li et al., 2022) leveraged LLMs to simulate dialogue scenarios in a controlled fashion, producing interactions with fluency and diversity comparable to human-collected data.

While these works highlight important develop-

Turn	Speaker	Utterance	Strategy
1	A	Oh, that's quite long! Have you always been fond of Takarazuka?	Empathy / Probing
2	B	For about the past ten years, I've gone once a year, or sometimes not at all. I used to be a hardcore fan, attending up to six times a week in the past.	–
3	A	Six times a week is impressive! By the way, have you ever considered performing on stage yourself?	Complimenting / Topic Transition
4	B	I once dreamed of being an actor and even attended a training school, but I realized that I prefer being a fan to performing. I even skipped rehearsals to follow my favorite actors.	–

Table 1: Sample dialogue annotated with interviewer strategies.

Item	Value
# Interviewers	3
# Interviewees	52
# Dialogues	52
# Utterances	4,111
# Interviewer utterances	2,219
# Interviewee utterances	1,892
# Avg. characters per utterance	28.13
# Avg. utterances per dialogue	79.06

Table 2: Statistics of the collected dialogue dataset.

ments, few studies have simultaneously addressed both adaptive acquisition of personal information and sustained engagement of users in qualitative interviews. Our research fills this gap by proposing a two-stage training approach that combines selective imitation learning guided by LLM-as-a-judge with simulation-enhanced strategy acquisition. This method aims to learn dynamic dialogue strategies that balance reliable information extraction with engagement over extended interactions.

3 Proposed Method

We propose **RaPSIL**, an LLM-based dialogue system for continuous interviews, using a two-stage learning approach (Fig. 2).

1. **Imitation Learning of Human Dialogue Strategies**: Acquire dialogue strategies (expressed as textual descriptions of the intent behind each utterance) employed by human interviewers during natural small talk using actual interview dialogue data.

2. **Simulation-Based Strategy Improvement**: Based on the initial model obtained via imitation learning, perform dialogue simulations between an interviewer LLM and a respondent LLM, and construct a self-improvement loop by employing an LLM as a judge (Gu et al., 2024) to dynamically refine the strategy.

Before detailing our learning approach, we first describe the construction of the interview dataset used for behavior cloning, followed by implementation details of each learning step.

3.1 Interview Dataset Construction

To build our dialogue dataset, we conducted chat-based interviews with three interviewers and 52 respondents. Each dialogue consists of 25 full turns, resulting in over 4,000 utterances in total—a relatively long format for chat-based interviews. The goal was to elicit a wide range of personal information—from superficial details to deeper self-disclosures—by focusing on five predefined personal characteristic categories: (1) Likes/Interests, (2) Dislikes/Weaknesses/Concerns, (3) Social Status/Work and Pride, (4) Personality, and (5) Life Goals/Future Plans. These categories are commonly used in persona-based dialogue datasets and digital twin research, and include deeply personal topics that may cause discomfort if addressed without sufficient rapport. Detailed examples of each category are provided in Appendix A.

During data collection, interviewers were instructed to engage in natural, spontaneous chit-chat as would be expected in first-time encounters. Rather than using direct questions, they were guided to organically integrate queries related to the five categories into the conversation. After each dialogue, interviewers documented the information they gathered for each category and annotated the intent behind each of their utterances (e.g., probing for further details, expressing empathy, or facilitating topic transitions). Notably, the next-utterance intent was recorded by the interviewer themselves, providing a rare form of fine-grained supervision. This explicit annotation of conversational intent is critical for our learning process, as it offers concrete linguistic cues that help the model learn strategic dialogue behavior.

In total, we collected 52 dialogues (4,111 utterances). Sample dialogues are provided in Table 1 and the statistics of the collected data are listed in Table 2. All dialogues were conducted in Japanese with Japanese participants.

3.2 Selective Behavior Cloning

In the first stage, we acquire human-like dialogue strategies through selective behavior cloning using our constructed dialogue dataset. Rather than learning from all dialogue data indiscriminately, we evaluate the quality of collected dialogue data and selectively learn from high-quality dialogue examples to achieve a balance between efficient information collection and natural dialogue.

Specifically, each dialogue is evaluated using GPT-4o[1] (OpenAI, 2024) as a judge, based on three key criteria: naturalness, appropriateness to the interlocutor, and strategicity. Each aspect is rated on a 10-point scale (detailed evaluation criteria are provided in Appendix B), and only high-quality dialogue examples exceeding a threshold score (24 points out of 30, corresponding to approximately the top 20%) are selected as training data, following the threshold-setting approach proposed in SOTOPIA-π (Wang et al., 2024). This enables learning focused on dialogue patterns that maintain human-like dialogue strategies while achieving more efficient information collection.

During training, the model takes the dialogue history of the 20 most recent utterances as input and simultaneously outputs both the content of the next utterance to be generated and its conversational intent. When the dialogue history exceeds 20 utterances, older utterances are discarded in chronological order. The specific format of the input prompts and examples are provided in Appendix C.

We adopt Llama-3.1-Swallow-8B-Instruct-v0.1[2] as our base model, which has demonstrated high performance in Japanese dialogue tasks (Fujii et al., 2024). We perform supervised fine-tuning (SFT) with a batch size of 2, gradient accumulation steps of 8, and a learning rate of 2e-6 (linear scheduling). The selected high-quality dataset is split into training and validation sets with a ratio of 8:2, and early stopping is implemented based on the validation loss. These values were chosen based on GPU memory constraints and preliminary tuning on the validation set, aiming for stable convergence. Unlike prior behavior cloning methods that use only dialogue-response pairs, our model is trained to generate both the response and its associated strategic intent, leveraging high-quality intent annotations for more strategy-aware generation.

[1]https://platform.openai.com/docs/models/gpt-4o

[2]https://huggingface.co/tokyotech-llm/Llama-3.1-Swallow-8B-Instruct-v0.1

For the subsequent simulation step, we train a respondent LLM using the same selective dataset, with the five categories of personal information included in the prompt. This phase establishes both interviewer and respondent models that reflect efficient and natural dialogue patterns.

3.3 Simulation-Based Strategy Improvement

In the second stage, we further refine the initial model obtained through selective imitation learning by implementing strategy improvement through dialogue simulation. Specifically, dialogue simulations are conducted between an interviewer LLM and a respondent LLM, generating 30 dialogues, consisting of 40–50 utterances each. The respondent LLM, which uses a model pre-fine-tuned via SFT, is provided with randomly generated personal characteristic information for the five categories. These persona descriptions are generated using the base Llama model to ensure diverse simulation scenarios. Similar to the first stage, each simulated dialogue is evaluated by GPT-4o using the same evaluation criteria, but with a focus on improving adaptability to more diverse dialogue situations. Simulation allows us to generate diverse dialogue scenarios beyond the limited coverage of human-collected data, which is especially important for learning adaptive strategies across sensitive topics.

Based on evaluation results, the following two improvement processes are carried out in parallel:

1. **Reconstruction of Training Data and Re-fine-tuning**: We reused the previously defined threshold (24/30) to select high-quality simulated dialogues. These high-scoring dialogues are selected as new training data and split into training and validation sets at an 8:2 ratio. To ensure the model remains aligned with human-like dialogue, the validation data includes, in addition to the collected dialogue data, an equal portion of dialogues from an existing human-to-human dataset. During re-fine-tuning, the previous model is used as the initial value, with model selection based on the validation loss and early stopping.

2. **Construction and Utilization of Few-shot Example Database**: For dialogues with low evaluation scores, we apply a Reflexion approach, where GPT-4o analyzes issues and simultaneously generates improvement examples along with the LLM-as-a-judge scores. Each example, consisting of an utterance and

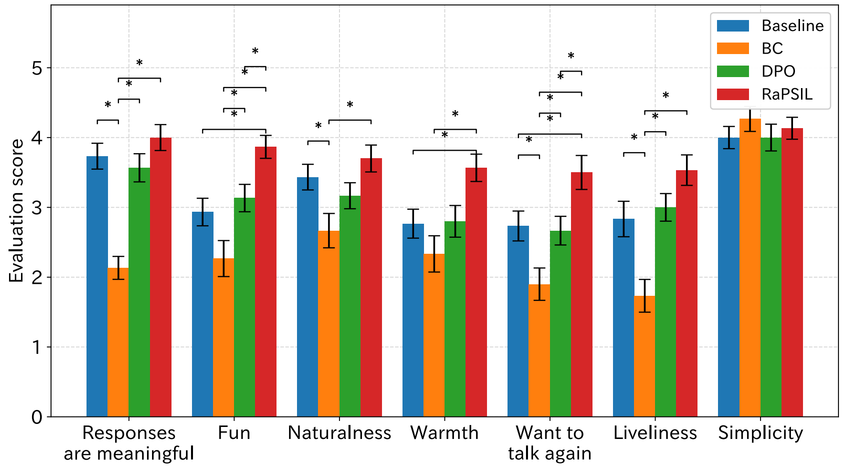


Figure 3: Comparison of impression evaluations for each method. Q1: Responses are Meaningful, Q2: Fun, Q3: Naturalness, Q4: Warmth, Q5: Want to Talk Again, Q6: Liveliness, and Q7: Simplicity. Error bars indicate the standard error, and asterisks denote significant differences after FDR correction ($p < 0.05$).

its corresponding intent, is embedded using Sentence-BERT[3]. During both training and inference, the dialogue history is similarly embedded, and the top two most similar examples based on cosine similarity are retrieved from the database and included in the prompt as few-shot examples.

Through this self-reinforcement loop, the model is able to learn from both successful and failure cases, acquiring a dynamic and optimal dialogue strategy that moves beyond the fixed behavior cloning model. RaPSIL enables continuous rapport building and effective information extraction by combining behavior cloning with simulation-based refinement.

4 Evaluations and Results

In this section, we evaluate the effectiveness of RaPSIL across three dimensions: user impressions, information extraction accuracy, and dialogue quality over time.

4.1 Experimental Setup

This experiment aimed to evaluate whether RaPSIL can elicit deeper personal information without reducing user engagement, compared to other methods. To validate the effectiveness of our approach, we conducted evaluation experiments with 30 participants who engaged in actual conversations with our dialogue systems. We evaluated four types of models. First, the baseline model only provides interview instructions as prompts to Llama-3.1-Swallow-8B-Instruct-v0.1. Second, the Behavior Cloning (BC) model was fine-tuned using the entire collected human dialogue dataset, with training parameters identical to those described in Section 3.2. Third, a DPO model was implemented, which was trained using GPT-4o Reflexion-improved utterances as preferred examples and original Llama outputs as rejected examples. Finally, RaPSIL, our proposed method underwent five iterations of the learning process described in Section 3.

The DPO model utilized a beta value of 0.1 and was trained with a learning rate of 5e-7 using a cosine scheduler. These hyperparameters were selected through minimal empirical tuning and were found to yield stable performance across five simulation-based iterations. Similar to RaPSIL, the DPO approach first trained on human dialogue data before proceeding through five iterations of simulation-based improvement, with each iteration using the previous model as initialization. For the experiment, each participant engaged in a dialogue consisting of 50 utterances (25 turns) with each of the four models. Participants were asked to provide self-description texts at least one day before the dialogue, covering five personal characteristic categories (likes, dislikes, work, personality, and goals). These self-descriptions served as ground truth for evaluating information extraction capabilities. Sample dialogues generated by each model are provided in Appendix D.

4.2 Subjective Evaluation

In the subjective impression evaluation, participants were asked to rate each dialogue system on a 5-point Likert scale across seven dimensions—responses are meaningful, fun, natural-

[3] https://huggingface.co/sonoisa/sentence-bert-base-ja-mean-tokens-v2

ness, warmth, willingness to talk again, liveliness, and conciseness—based on dimensions adapted from prior dialogue evaluation study (Kobori et al., 2016). Consistent with the statistical analysis approach in previous studies, non-parametric methods were employed. First, we conducted a Friedman test to examine significant differences among the four models (with a significance level of $p < 0.05$), followed by post-hoc comparisons using the Wilcoxon signed-rank test with FDR correction using the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

As shown in Figure 3, RaPSIL achieved the highest scores in six out of the seven dimensions, with only "simplicity" showing no significant difference across models. The BC model received the highest score for conciseness (4.27), likely reflecting its imitation of concise human speaking styles. However, this led to monotonous agreement or short replies, which negatively affected other dimensions such as "liveliness" (1.73) and "want to talk again" (1.90). The DPO model followed the baseline in overall performance. Although it slightly reflected user input and utilized more formal language due to GPT-4o-refined chosen examples, its dialogue behavior often resembled the baseline, with repeated questions and limited personalization. This may explain why it received similar scores to the baseline across most dimensions.

RaPSIL demonstrated statistically significant superiority over other models in multiple dimensions. It outperformed all others in "fun" (3.87) and "want to talk again" (3.50), and showed significant improvements over the BC model in "responses are meaningful" (4.00), "naturalness" (3.70), "warmth" (3.57), and "liveliness" (3.53). These results indicate that RaPSIL effectively addresses the challenge of maintaining user engagement while eliciting personal information. In particular, significant improvements in warmth, fun, and willingness to talk again highlight the system's potential for sustained long-term interaction in practical use cases.

4.3 Information Extraction

To objectively evaluate the information extraction capabilities of each model, we employed RAGChecker (Ru et al., 2024), using GPT-4o-mini[4] as the underlying LLM. This tool measures precision, recall, and F1 score by comparing information extracted from dialogues against participants' self-description texts as reference. We extracted claims from each dialogue and checked them against the participants' self-descriptions across the five personal characteristic categories. Table 3 presents the evaluation results.

Model	Recall	Precision	F1
Baseline	<u>26.52</u>$\pm$17.00	16.85$\pm$7.84	18.95$\pm$9.37
BC	16.67$\pm$17.21	<u>19.93</u>$\pm$14.61	16.23$\pm$15.72
DPO	24.89$\pm$16.26	15.34$\pm$8.62	<u>17.98</u>$\pm$10.51
RaPSIL	**29.25**$\pm$13.40	**21.49**$\pm$6.57	**23.29**$\pm$8.57

Table 3: Comparison of recall, precision, and F1 scores on the information extraction task (mean $\pm$ std).

RaPSIL achieved the highest Recall (29.25%), Precision (21.49%), and F1 score (23.29%) among all models, indicating superior information extraction capability. Interestingly, while the baseline model showed the second-highest Recall value (26.52%), its relatively lower Precision (16.85%) suggests that it gathers a wide range of information but with less accuracy. The BC model, in contrast, demonstrated higher Precision (19.93%) but the lowest Recall (16.67%), indicating it extracts more accurate but less comprehensive information. The DPO model showed performance comparable to the baseline across all metrics, suggesting moderate information breadth and precision.

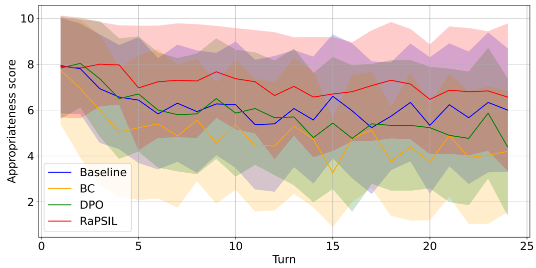
These objective evaluation results align with and complement the subjective impression findings. The RaPSIL method not only creates more engaging and natural conversations (as shown in subjective evaluations) but also extracts more comprehensive and accurate information. This demonstrates that our approach effectively balances the dual objectives of maintaining user engagement and gathering valuable personal information, addressing the key challenge in interview-based information collection systems.

4.4 Dialogue Quality Over Time

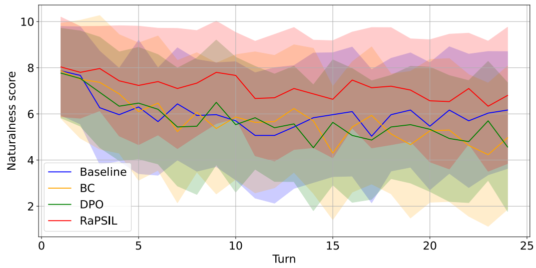
To understand how dialogue quality evolves over the course of an interaction, we conducted a temporal evaluation in which participants rated each system utterance on a 10-point scale across three dimensions: appropriateness to the interlocutor, naturalness, and strategic value. These evaluation criteria were aligned with those used in the LLM-as-a-judge process during training in RaPSIL.

As shown in Figure 4, RaPSIL consistently maintained higher scores throughout dialogue, with a more gradual decline in quality compared to other

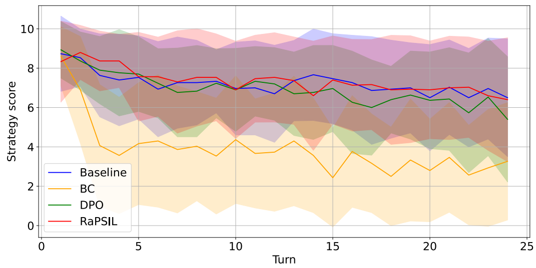
[4]https://platform.openai.com/docs/models/gpt-4o-mini



(a) Appropriateness



(b) Naturalness



(c) Strategy

Figure 4: Turn-by-turn evaluation of dialogue quality across three dimensions: (a) appropriateness to the interlocutor, (b) naturalness, and (c) strategic value.

models. The BC model exhibited the most pronounced degradation over time, while the baseline and DPO models showed moderate and similar decreases. Strategic value remained at a comparable level across all models except BC, which lagged behind in later turns. These results highlight a critical advantage of RaPSIL: the ability to sustain dialogue quality over extended interactions—a particularly important feature for interview scenarios that require maintaining engagement while eliciting comprehensive information.

5 Discussion

To further analyze the types of information elicited by each system, we categorized the claims extracted by RAGChecker using GPT-4o-mini into five semantic categories. The aggregated results are shown in Table 4. Notably, RaPSIL produced the highest number of claims in the "Other" category, which includes casual or off-topic statements that are not directly related to the main interview objective. This indicates that RaPSIL frequently initiates

Model	Interests	Dislikes	Work	Traits	Goals	Other
Baseline	**12.80**	**3.10**	3.50	4.73	3.20	1.83
BC	9.43	2.33	1.33	1.20	0.60	1.20
DPO	11.77	2.70	2.90	**5.33**	**5.60**	1.30
RaPSIL	9.70	1.57	**4.30**	3.17	4.07	**3.13**

Table 4: Average number of claims per user across categories. RaPSIL elicited more claims in the Other category, which includes casual or off-topic utterances.

small talk or tangential conversations—an interaction pattern known to promote rapport and engagement in dialogue systems (Kobori et al., 2016). Representative examples of extracted claims for each category are shown in Appendix E

A potential concern with such off-topic exchanges is the risk of reduced strategic quality. However, Figure 4c shows that RaPSIL maintained high strategy scores throughout the dialogue, comparable to or better than other methods. This suggests that RaPSIL engages in off-topic conversation not arbitrarily, but as a deliberate strategy to enhance engagement while preserving information elicitation goals. In short, RaPSIL demonstrates the ability to integrate small talk in a context-aware and goal-sensitive manner, contributing to both user engagement and effective information gathering.

6 Conclusion

This study demonstrated a method for building dialogue systems that can naturally elicit deep personal information—such as concerns, personality, and life goals—without reducing user engagement. By combining imitation learning with simulation-enhanced refinement, our approach enables the model to acquire conversational techniques like casual chatting and empathetic phrasing, which contribute to both user comfort and strategic information gathering.

Experimental results confirmed that our system successfully extracted accurate, self-consistent information aligned with users' self-descriptions while maintaining high engagement scores. These findings highlight the potential of our approach for improving the personalization accuracy of dialogue systems by fostering user self-disclosure through enjoyable and natural conversations. In particular, it shows strong potential as a foundational technology for practical systems such as interview agents and personalized assistants that require sustained user engagement to build a deeper understanding of individuals.

7 Limitations

While RaPSIL demonstrated promising performance in eliciting personal information without compromising user engagement, several directions remain for future improvement and extension.

First, our current evaluation focuses on reproducing information aligned with five predefined categories from user self-description texts. While this approach enables quantitative comparison, it does not fully capture the potential of interview systems to uncover deeper or previously unarticulated insights. Moreover, as our study integrates multiple components—selective imitation, reflexion-based few-shot prompting, and intent annotations—we did not conduct component-wise ablation analysis. This is partly due to the scale of human evaluation and the difficulty of isolating individual effects in a statistically robust manner. Future work should develop new evaluation metrics that assess not only fidelity to self-reported content but also the depth, novelty, and specific contribution of each component in promoting effective dialogue.

Second, our system was implemented and tested in Japanese, due to practical considerations such as participant recruitment and ensuring natural communication. However, the framework itself is language-agnostic and can be adapted to English or other languages. Future studies should examine how cultural and linguistic differences affect rapport-building strategies and disclosure patterns.

Third, while our method already utilizes a relatively lightweight model (8B) suitable for many real-time applications, response latency remains a critical factor in user experience. Optimizing for faster response generation—through model distillation, caching strategies, or reinforcement learning approaches like GRPO (DeepSeek-AI et al., 2025)—will further improve the practicality of deployment in live systems.

Finally, our dialogue setting is currently limited to single-turn, text-based interactions. A natural extension would be to incorporate multimodal signals such as gaze and facial expressions, which play a key role in trust and engagement. Furthermore, expanding to multi-agent scenarios—where multiple interviewer agents coordinate—may unlock more dynamic and context-sensitive dialogue strategies.

Taken together, these directions aim to build on the strengths of RaPSIL and move toward more adaptive, efficient, and socially intelligent dialogue agents for personalized interaction.

8 Ethical Considerations

All dialogue data used in this study were collected with written informed consent, after clearly explaining the research purpose, data to be collected, and scope of use. Participation was entirely voluntary, and participants were informed that they could withdraw at any time.

All collected data were fully anonymized and contained no personally identifiable information (PII). The self-description texts and dialogue logs were used strictly for research purposes, with no third-party distribution or commercial use involved.

Even in future work involving more long-term or personally sensitive interactions, we recognize the importance of maintaining user privacy through appropriate data management and transparency.

References

Muhammad Yeza Baihaqi, Angel Garc'ia Contreras, Seiya Kawano, and Koichiro Yoshino. 2024. Rapport-driven virtual agent: Rapport building dialogue strategy for improving user experience at first meeting. *ArXiv*, abs/2406.09839.

Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.

Timothy W. Bickmore and Rosalind W. Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput.-Hum. Interact.*, 12(2):293–327.

DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, and Z. F. Wu. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *Preprint*, arXiv:2501.12948.

Grant Eckstein, Dana Ferris, and Katherina Sibbald. 2021. What do students think about their own writing? insights for teaching new college writers. *Writing & Pedagogy*, 12(2-3):307–339.

Susan R. Fiksdal. 1988. Verbal and nonverbal strategies of rapport in cross-cultural interviews. *Linguistics and Education*, 1:3–17.

Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Hiroki Iida, Masanari Ohi, Kakeru Hattori, Hirai Shota, Sakae Mizuki, Rio Yokota, and Naoaki Okazaki. 2024. Continual pre-training for cross-lingual llm adaptation: Enhancing japanese language capabilities. In *Proceedings of the First Conference on Language Modeling*, COLM, page (to appear), University of Pennsylvania, USA.

Jiawei Gu, Xuhui Jiang, Zhichao Shi, Hexiang Tan, Xuehao Zhai, Chengjin Xu, Wei Li, Yinghan Shen, Shengjie Ma, Honghao Liu, Yuanzhuo Wang, and Jian Guo. 2024. A survey on llm-as-a-judge. *CoRR*, abs/2411.15594.

Takahiro Kobori, Mikio Nakano, and Tomoaki Nakamura. 2016. Small talk improves user impressions of interview dialogue systems. In *Proceedings of the SIGDIAL 2016 Conference, The 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 13-15 September 2016, Los Angeles, CA, USA*, pages 370–380. The Association for Computer Linguistics.

Chee Wee Leong, Navaneeth Jawahar, Vinay Basheerabad, Torsten Wörtwein, Andrew Emerson, and Guy Sivan. 2024. Combining generative and discriminative ai for high-stakes interview practice. In *Companion Proceedings of the 26th International Conference on Multimodal Interaction*, ICMI Companion '24, page 94–96, New York, NY, USA. Association for Computing Machinery.

Zekun Li, Wenhu Chen, Shiyang Li, Hong Wang, Jing Qian, and Xifeng Yan. 2022. Controllable dialogue simulation with in-context learning. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 4330–4347. Association for Computational Linguistics.

Gale M Lucas, Jonathan Gratch, Adam King, and Louis-Philippe Morency. 2014. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37:94–100.

OpenAI. 2024. Gpt-4o system card. *Preprint*, arXiv:2410.21276.

Atsushi Otsuka, Kenta Hama, Narichika Nomoto, Ryo Ishii, Atsushi Fukayama, and Takao Nakamura. 2023. Learning user embeddings with generating context of posted social network service texts. In *Social Computing and Social Media - 15th International Conference, SCSM 2023, Held as Part of the 25th HCI International Conference, HCII 2023, Copenhagen, Denmark, July 23-28, 2023, Proceedings, Part I*, volume 14025 of *Lecture Notes in Computer Science*, pages 106–115. Springer.

Joon Sung Park, Carolyn Q. Zou, Aaron Shaw, Benjamin Mako Hill, Carrie J. Cai, Meredith Ringel Morris, Robb Willer, Percy Liang, and Michael S. Bernstein. 2024. Generative agent simulations of 1,000 people. *CoRR*, abs/2411.10109.

Shuofei Qiao, Yixin Ou, Ningyu Zhang, Xiang Chen, Yunzhi Yao, Shumin Deng, Chuanqi Tan, Fei Huang, and Huajun Chen. 2023. Reasoning with language model prompting: A survey. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Toronto, Canada, July 9-14, 2023*, pages 5368–5393. ACL.

Libo Qin, Wenbo Pan, Qiguang Chen, Lizi Liao, Zhou Yu, Yue Zhang, Wanxiang Che, and Min Li. 2023. End-to-end task-oriented dialogue: A survey of tasks, methods, and future directions. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pages 5925–5941. Association for Computational Linguistics.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, volume 36, pages 53728–53741. Curran Associates, Inc.

Verena Rieser and Oliver Lemon. 2011. Learning and evaluation of dialogue strategies for new applications: Empirical methods for optimization from small data sets. *Computational Linguistics*, 37(1):153–196.

Dongyu Ru, Lin Qiu, Xiangkun Hu, Tianhang Zhang, Peng Shi, Shuaichen Chang, Cheng Jiayang, Cunxiang Wang, Shichao Sun, Huanyu Li, Zizhao Zhang, Binjie Wang, Jiarong Jiang, Tong He, Zhiguo Wang, Pengfei Liu, Yue Zhang, and Zheng Zhang. 2024. Ragchecker: A fine-grained framework for diagnosing retrieval-augmented generation. In *Advances in Neural Information Processing Systems (NeurIPS) Datasets and Benchmarks Track*.

Mustafa Safdari, Greg Serapio-García, Clément Crepy, Stephen Fitz, Peter Romero, Luning Sun, Marwa Abdulhai, Aleksandra Faust, and Maja J. Mataric. 2023. Personality traits in large language models. *CoRR*, abs/2307.00184.

Wei Shengli. 2021. Is human digital twin possible? *Computer Methods and Programs in Biomedicine Update*, 1:100014.

Zhoujian Sun, Chenyang Zhao, Zhengxing Huang, and Nai Ding. 2023. Replicating complex dialogue policy of humans via offline imitation learning with supervised regularization. *Preprint*, arXiv:2305.03987.

J. Tang, C. Abraham, and E. Stamp. 2015. How can weight-loss app designers best engage and support users? a qualitative investigation. *British Journal of Health Psychology*, 20(4):781–791.

Dennis Ulmer, Elman Mansimov, Kaixiang Lin, Lijia Sun, Xibin Gao, and Yi Zhang. 2024. Bootstrapping LLM-based task-oriented dialogue agents via self-talk. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 9500–9522, Bangkok, Thailand. Association for Computational Linguistics.

Thea F. van de Mortel. 2008. Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, 25(4):40–48.

J. Wang, J. Sun, X. Yuan, G. Dong, and 1 others. 2023. Prompting frameworks for large language models: A survey. *arXiv preprint arXiv:2311.12785*.

Ruiyi Wang, Haofei Yu, Wenxin Sharon Zhang, Zhengyang Qi, Maarten Sap, Graham Neubig, Yonatan Bisk, and Hao Zhu. 2024. Sotopia-π: Interactive learning of socially intelligent language agents. *CoRR*, abs/2403.08715.

Wenqi Zhang, Yongliang Shen, Linjuan Wu, Qiuying Peng, Jun Wang, Yueting Zhuang, and Weiming Lu. 2024. Self-contrast: Better reflection through inconsistent solving perspectives. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3602–3622, Bangkok, Thailand.

A Personal Characteristic Categories

Personal characteristic categories were uniquely selected by referencing existing persona dialogue datasets and digital twin research on personal information. Table 5 shows detailed examples for each of the five categories used in our study.

These five categories were also used in the simulation phase to construct respondent personas. Specifically, we instructed the base Llama model to generate an original set of five personal characteristics in natural Japanese, using few-shot examples randomly sampled from the dialogue dataset collected in this study. This helped guide the model to produce diverse yet realistic personas. Outputs were manually checked to filter out implausible or inconsistent responses.

B Dialogue Evaluation Criteria

For dialogue quality evaluation, we used a structured scoring system with three main dimensions, as shown in Table 6. These criteria were utilized by GPT-4o in the role of a judge to evaluate both human dialogues during the selective behavior cloning phase and simulated dialogues during the strategy improvement phase.

C Prompt Details

In this work, we adopt a structured prompt format for both training and inference. The prompt directs the system—acting as an experienced interviewer—to gather information on five topics: (i) Likes/Interests, (ii) Dislikes/Weaknesses/Concerns, (iii) Social Status/Work and Pride, (iv) Personality Traits, and (v) Life Goals/Future Plans.

Table 7 summarizes these prompts:

1. **System Message**: Defines the interviewer's role and objectives, emphasizing that the system must consistently remain in character.

Category	**Example**
Interests and Preferences	Enjoys reading on Japanese history, traveling, and listening to music. These interests have long remained consistent and can be enjoyed alone.
Dislikes and Concerns	Feels that a lack of stamina prevents the accomplishment of desired tasks and leads to often falling short compared to others. Although there is a desire to take on more work, being too tired to care for family renders it meaningless.
Work, Status, and Pride	Feels a certain level of satisfaction with the work as a registered dietitian, and would like to engage in work that is truly well-suited if the income were more stable.
Personality Traits	Tends to be anxious and sensitive; although there are goals, the slow pace of achievement may be partly due to a lack of stamina.
Life Goals and Future Plans	Aspires to work in English and hopes to increase opportunities to interact with foreigners by becoming proficient in the language.

Table 5: Representative examples for each personal characteristic category.

2. **Reference Examples**: Provides sample interviewer utterances with associated reasoning (Intent), selected based on Sentence-BERT embedding similarity, when available.

3. **Current Dialogue Context**: Presents a history of recent dialogue with speaker attribution, enabling context-aware questioning.

4. **Expected Output Format**: Specifies the output should include the next interviewer utterance along with an explanation of its rationale. In our setup, the intent is generated after the utterance as a post-hoc explanation, rather than a prior plan; this reflects the conversational nature of intent as an interpretive label, not a control signal.

This concise design clarifies the interviewer's role, maintains natural dialogue flow, and facilitates systematic information collection. For RaPSIL, reference examples include both original and Reflexion-improved utterances, chosen by semantic similarity to ensure fair and consistent comparisons across modeling approaches.

Aspect	Scoring Guidelines
Naturalness	Human-like flow of dialogue *- Natural topic transitions* *- Emotional and colloquial expressions* *- Avoidance of mechanical or templated responses*
Appropriateness	Adaptation to interlocutor *- Stylistic alignment with the interlocutor* *- Psychological and emotional consideration* *- Responses reflecting interlocutor's interests*
Strategicity	Dialogue planning and elicitation *- Elicitation of missing information* *- Using prior utterances in question design* *- Turn-aware dialogue development*

Table 6: Evaluation criteria for dialogue strategies. Each aspect is scored on a 10-point scale.

D Dialogue Examples

Table 8 presents characteristic dialogue patterns observed for each method. Clear differences were found in how each model developed topics and elicited information during conversations.

The baseline model asked about the types of leisure facilities the respondent liked but only engaged in surface-level confirmations without developing the topic further. It remained at the level of simple question-and-answer exchanges, showing little ability to expand the dialogue.

The BC model reacted to the user's mention of children by providing basic responses but failed to further deepen the conversation. Although it demonstrated some degree of acknowledgment, it did not successfully build on the information to drive the dialogue forward.

The DPO model focused intensively on a specific topic by asking detailed questions about the user's travel planning and then further exploring aspects of the trip the user was looking forward to. However, its tendency to continuously fire successive questions sometimes created a sense of pressure, potentially overwhelming the respondent.

RaPSIL smoothly expanded the conversation from a topic about a skiing trip, naturally shifting to snowboarding, and then to vivid descriptions of the children's free and playful skiing. By flexibly building on the user's responses, RaPSIL succeeded in eliciting rich personal episodes and maintaining high engagement throughout the dialogue.

E Representative Claim Examples

Table 9 presents randomly sampled examples of knowledge claims extracted during user interactions with each system. The categorization was automatically conducted using GPT-4o-mini. Each example reflects actual information elicited from users, illustrating the types of knowledge captured across different categories.

Notably, the "Other" category includes casual or off-topic knowledge, indicative of spontaneous conversational dynamics. Since RaPSIL generated more claims in the "Other" category compared to the other systems (see Table 4), this suggests that RaPSIL naturally facilitated more casual and exploratory conversations during interactions.

[System Message]
You are an experienced interviewer. Your role is to extract information about the following 5 topics through natural conversation:

- Likes/Interests
- Dislikes/Weaknesses/Concerns
- Social Status/Work and Pride
- Personality Traits
- Life Goals/Future Plans

Key interviewer guidelines:
1. Always maintain the interviewer role, and never respond as an interviewee.
2. Carefully track dialogue history to avoid repetitive questions.
3. Guide topic transitions naturally through appropriate questions.
4. Ensure questions are contextually relevant to previous responses.
5. Balance active listening with information gathering.
6. Conclude with gratitude once sufficient information is collected.

[Reference Examples] (If available)
*(This section may contain **N** few-shot examples. N can be 0 or more.)*
Example N:
Utterance: [Interviewer's question]
Intent: [Strategic reasoning]

[Current Dialogue Context]
Speaker: [Interviewer/Respondent]
Content: [utterance]
(Shows recent dialogue history)

[Expected Output Format]
Utterance: [Next appropriate interviewer response]
Intent: [Explanation of question choice and information target]

Table 7: Overview of the prompt format utilized in this work.

Model	Speaker	Utterance
Baseline	User	I like visiting leisure facilities.
	System	What kind of facilities do you enjoy, such as amusement parks or theme parks?
	User	Disneyland and DisneySea.
	System	You enjoy both parks, not just one, that's wonderful. If you have a favorite attraction or show, could you tell me?
	User	I like all attractions.
	System	Then, could you tell me more about how you usually enjoy Disney? For example, do you prefer meeting characters, watching shows and parades, or something else?
	(Characteristic: Surface-level confirmation without topic development)	
BC	User	Yes, I'm looking forward to it. My child really wanted to go, so I'm working hard on planning.
	System	Oh, you have children.
	User	Yes, two children.
	System	That must be tough.
	(Characteristic: Simple back-and-forth, does not deepen the topic)	
DPO	User	We reserved accommodations and transportation, and also booked some activities. I also researched recommended local restaurants because I wanted to enjoy good local food.
	System	How did you go about planning this USJ trip? Did you plan it with someone, or based on your own research? I'd love to hear the specifics.
	User	I'm going with my daughter, so I asked her what she wanted to do at USJ. I made the detailed plans myself. I checked hotel reviews on Google Maps and chose a hotel close to USJ that was affordable.
	System	What are you most looking forward to during this trip?
	(Characteristic: Aggressive deep-diving into a single topic, continuous questioning flow)	
RaPSIL	User	I went skiing in Niigata Prefecture. Although the weather was cloudy, the slopes were just right for my level, and I had a great time skiing.
	System	Do you like skiing?
	User	Yes, although nowadays I snowboard. When I was a child, I used to ski.
	System	I see! So your children also enjoy skiing with you?
	User	Yes, they do. When we go skiing, they just keep skiing non-stop. They love jumping over bumps and skiing through little forest paths, enjoying it freely.
	(Characteristic: Natural flow leading to rich personal episodes with high engagement)	

Table 8: Comparison of characteristic dialogue patterns across models.

Category	Example Claims
Interests and Preferences	(User, enjoys, hot springs and theme parks) (User, enjoys, 3D modeling) (User, likes, sirloin beef) (User, enjoys, traveling abroad) (User, likes, dynamic dance moves)
Dislikes and Concerns	(User, struggles with, workload and time) (User, feels frustrated, with strict compliance in society) (User, experienced, significant losses) (User, avoids, sharing existing character illustrations due to potential issues) (User, believes, moderate rest is necessary)
Work, Status, and Pride	(User, finds fulfillment, in contributing opinions to product development) (User, researches, support needed for children with developmental disabilities in regular classes) (User, considers communication, important) (User, is currently, working on the satisfactory completion of ongoing work) (User, manages, internship operations)
Personality Traits	(User, believes, relaxation comes after completing tasks) (User, feels happy when, thanked by others) (User, feels growth when, overcoming difficult tasks) (User, tries to maintain, a peaceful atmosphere at work) (User, describes self as, easy-going)
Life Goals and Future Plans	(User, wants to contribute to, decarbonized society) (User, plans to visit, Egypt) (User, aims to achieve, career advancement in about 10 years) (User, wants to focus on, design after solving current issues) (User, dreams of living, in a tropical country)
Other	(User, observed, cactus growing suddenly when exposed to direct sunlight) (User, exchanged alcohol using, a Lawson trial coupon service) (User, was shocked by, witnessing a murder trial) (User, felt like, being in an aquarium while scuba diving) (User, found warm peach dessert, having prominent aroma and sweetness)

Table 9: Representative examples of extracted claims for each category. Each entry shows (subject, relation, object).