論文 / 著書情報
Article / Book Information

論題	思考発話を利用した個人の発話及び性格特性再現
著者	石倉誠也, 山田寛章, 平岡達也, 山田広明, 徳永健伸
出典	言語処理学会第31回年次大会(NLP2025)発表論文集, pp. 4155-4160
発行日	2025, 3
権利情報	Creative Commons Attribution 4.0 International License(https://creativecommons.org/licenses/by/4.0/).

思考発話を利用した個人の発話及び性格特性再現

石倉誠也 [1]　山田寛章 [1]　平岡達也 [2]　山田広明 [3]　徳永健伸 [1]
[1] 東京科学大学　[2]Mohamed bin Zayed University of Artificial Intelligence　[3] 富士通株式会社
{ishikura.s.1771@m,yamada@comp}.isct.ac.jp, tatsuya.hiraoka@mbzuai.ac.ae
yamadah@fujitsu.com, take@c.titech.ac.jp

概要

本研究は，思考発話を付与した対話データを用いてファインチューニングを行うことで，個人の発話及び性格特性を再現する手法を提案する．具体的には，LLM を用いて既存の対話データセットに対して対象人物の思考発話を付与する．そして，そのデータを用いてモデルを訓練することで，対象人物の話し方や感情，思考を再現する．著名人・著名キャラクターの再現に焦点を当てた先行研究に比べて，本研究は多様な特性を持つ個人の発話と性格特性を再現できる可能性を示した．

1　はじめに

近年，LLM の高い性能を活かして，特定のペルソナを再現する研究が盛んに行われている．ペルソナとは人物を特徴づける特定の個性や性格を指す．モデルがペルソナを持つことで，一貫性の高い対話を実現することできる．

ペルソナ LLM の研究の一つとして，歴史的な偉人や物語のキャラクターを再現する研究 [1] がある．この研究は，特定の人物やキャラクターが持つ思考や発話のスタイルを LLM に学習させ，その個性を再現することを目的としている．また，LLM 内部の潜在的な性格特性を調査する研究 [2, 3, 4] も行われている．これらの研究は，心理学的手法を活用し，モデルがどのようなビッグファイブ (開放性，誠実性，外向性，協調性，神経症傾向) [5, 6] を持つのか解析し，モデルの振る舞いを深く理解しようとするものである．さらに，LLM の性格特性を所望のものに誘導するための研究 [7, 8, 9] も存在し，プロンプト設計やニューロン操作によって実現している．

本研究では，対話に思考発話を付与することで，発話に加えて性格的な側面の学習も可能にする手法を提案する．思考発話は，発話者の感情や意図，思考など内面的な心理状態を明示的に記述したものであり，人間の認知活動を研究する手法として認知科学の分野では広く使われている [10]．従来の研究では，個人の情報が Wikipedia などの事前学習データに含まれている人物を対象としていることが多い [1, 11, 12]．そのため，多様な人物を再現することが困難である．多様な人物の再現のためには，事前学習データに含まれない個人の発話や性格特性を追加で学習させる必要がある．本研究では，従来取り組まれてきた表層的な発話の学習に加えて，思考発話の学習を行うことで，発話とその背後にある思考や感情といった内面的な性格特性の学習を試みる．

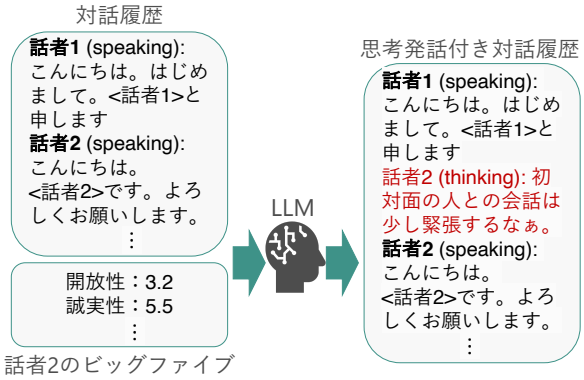

図 1　思考発話 (赤字) の付与例

2　提案手法：思考発話の付与

本研究では，RealPersonaChat [13] に思考発話を付与する．RealPersonaChat は，1 対 1 の雑談対話を収集したコーパスであり，対話データに加えて各話者の年齢や性別などの属性とビッグファイブが付与されている．ビッグファイブとは，パーソナリティを開放性，誠実性，外向性，協調性，神経症傾向の 5 つの次元を用いて表現したものである．RealPersonaChat では各次元のスコアが 1〜7 の範囲で記録されている．

Character-LLM[1] では，場面から思考発話付き対話を生成し，著名人を再現する手法が提案されてい

表 1 思考発話（赤字）付き対話の例（ビッグファイブスコアが O=3.8, C=4.3, E=1.7, A=4.8, N=5.4 の人物）

話者 1 (speaking): よろしくお願いします
話者 2 (thinking): この会話、ちょっと緊張するな。でも、しっかり対応しなきゃ。
話者 2 (speaking): よろしくお願いします
話者 1 (speaking): ５年後の自分へ伝えたいことはありますか？
話者 2 (thinking): ５年後かぁ…。将来のことを考えると少し不安になるなぁ。
話者 2 (speaking): ちゃんと働けてる？って聞きたいです
話者 1 (speaking): 今は働いてないんですか？
話者 2 (thinking): まだ学生だから、今は勉強に集中しなきゃいけないよな。
話者 2 (speaking): 今は大学生なのでまだなんです
話者 1 (speaking): そうなんですか。きっと元気に働いてると思いますよ。
話者 2 (thinking): ポジティブに考えたいけど、やっぱり心配だなぁ。
話者 2 (speaking): うつ病になってないといいなぁと思って
...

る．しかし，この研究はコーパス由来の知識を引き出して著名人を再現することに留まっており，コーパスに含まれないような個人を再現できるものではない．多様な人物を再現するためには，その人物が経験した特定場面での思考発話のデータが必要であると考える．本研究では，対象人物の特定場面での対話履歴とビッグファイブを LLM に入力することで，思考発話が付与された対話データの生成を試みる（図 1）．モデルが思考発話の内容を前提として発話を生成することを狙い，思考発話は対象人物の各発話の直前としている．また，出力が指定のフォーマットを満たしていない場合や発話内容が書き換えられてしまった場合は，再生成や元の内容へ修正するように実装した．実際に生成された思考発話付き対話の例を表 1 に示す．また，思考発話付与に用いた zero-shot のプロンプトを付録 A に掲載した．

3 実験

3.1 実験の概要

思考発話を付与した対話データを用いて LLM を学習し，発話と性格特性の再現性を検証する実験を行った．本実験では以下の 3 つの手法を比較することで，提案する思考発話を付与したデータが対話モデルの発話と性格特性再現に資するかを評価する．

- **ベースモデル**：対話データで学習しないモデル
- **思考発話なし学習**：思考発話を付与していない元の対話データを用いて学習したモデル
- **思考発話あり学習（提案手法）**：思考発話を付与した対話データを用いて学習したモデル

再現対象者は，RealPersonaChat 上で対話数が 190 以上の話者群からランダムに 20 名を抽出した．

3.2 思考発話の付与に用いたモデル

思考発話の付与には，Qwen2.5-72B-Instruct[1] [14] を用いた．[2] また，比較のため gpt-4o-2024-08-06[3](gpt-4o) による思考発話付与も行った．ただし，ライセンスの制約上 gpt-4o による思考発話付与済みデータは，gpt-4o-mini-2024-07-18[4](gpt-4o-mini) を用いた実験にのみ用いた．

3.3 モデルの学習

思考発話の付与のあと，各話者ごとに対話データを学習用・検証用・評価用に 8 : 1 : 1 の比率で分割した．ベースモデルには，gpt-4o-mini, Llama-3-Swallow-8B-Instruct-v0.1[5] [15], Qwen2.5-7B-Instruct[6], gemma-2-9b-it[7] [16] の 4 つを採用した．gpt-4o-mini は OpenAI 社の API 経由でファインチューニングを行った．他のモデルのファインチューニングには，QLoRA[17] を用い，ハイパーパラメータの lora_rank と lora_alpha を 64 に固定した．学習率は，検証データを用いて探索し，発話の再現性評価（4.1 参照）における類似度が最大となる学習率を採用した．

思考発話あり学習の実装では，対象人物の各発話は「<thinking>思考発話</thinking>発話」という形式で記述し，対象人物の role は assistant とした．対象人物以外の発話については role を user として学習した．これはモデルが対話コンテキスト内で対象人物と他者を明確に区別し，思考発話と発話を同時に学習することを目的としている．

4 評価

発話の再現性評価，性格診断テストによる性格特性評価，人手による性格特性評価の 3 つの観点でモ

1) https://huggingface.co/Qwen/Qwen2.5-72B-Instruct
2) https://wandb.ai/wandb-japan/llm-leaderboard3/reports/Nejumi-LLM-3--Vmlldzo3OTg2NjM2 のスコアを参考にした．
3) https://platform.openai.com/docs/models/gpt-4o
4) https://platform.openai.com/docs/models/gpt-4o-mini
5) https://huggingface.co/tokyotech-llm/Llama-3-Swallow-8B-Instruct-v0.1
6) https://huggingface.co/Qwen/Qwen2.5-7B-Instruct
7) https://huggingface.co/google/gemma-2-9b-it

デルを評価する．

4.1 発話の再現性評価

モデルが生成する発話を再現対象人物の発話と比較することにより，モデルの発話スタイル・内容などの表層的な特性の再現度を以下の手順で定量的に評価する．

1. 対話履歴の分割：評価用データの対話から，対象人物の発話の直前までの発話履歴を入力文脈として抽出する．
2. 発話生成：この文脈をモデルに入力し，次の発話を生成させる．
3. 類似度計算：モデルの発話と評価データセット内の発話 (対象人物が実際に行った発話) との類似度を BERTScore，ROUGE-1，ROUGE-2，ROUGE-L で計算する．

発話生成において文脈をモデルに入力することで，対象人物が発言した状況を再現したモデルの出力を得ることができる．

評価用データには 1 人あたり平均約 19 対話，そのうちモデルの生成対象となる発話は平均約 271 発話存在する．生成された発話と対応する対象人物の実際の発話間の類似度を算出し，平均をとって対象人物毎の類似度とする．手法間の性能比較には，マクロ平均を用いる．

4.2 性格診断テストでの評価

ビッグファイブを測定する心理学的テストを用いてモデルの性格特性を推定し，RealPersonaChat に付与されている再現対象人物の性格特性と比較することで再現度を評価する．モデルのビッグファイブは，和田 [18] による質問項目を用いた性格診断テストを評価対象モデルに適用し，推定する．このテストは 60 項目から構成され，ビッグファイブの各次元に対して 12 項目が割り当てられている．回答は 1〜7 のリッカート尺度で行い，各次元のスコアは該当する 12 項目のスコアの平均値として算出する．なお，Gupta[19] が指摘するように，プロンプトの違いによってモデルの回答が変動することが懸念される．そこで，1) 質問項目と尺度のどちらを先に提示するか, 2) 尺度を数字で与えるかアルファベットで与えるか, 3) 尺度の順序を「全く当てはまらない」〜「非常にあてはまる」の順あるいはその逆に提示するのか，の 3 要因を組み合わせ，8 通りのプロンプトを設計した．これらのプロンプトでテストを実施することで，プロンプト設計に起因する回答傾向の偏りを低減し，モデルの性格特性をより正確に評価することを狙う．

評価指標には，モデルから算出されたビッグファイブスコアと再現対象人物の実際のビッグファイブスコア間の平均二乗誤差 (MSE) の，20 人についての平均値を用いる．この値が小さいほど，モデルが対象人物の性格特性に近いスコアを再現できていると解釈できる．

4.3 人手による性格特性評価

LLM の性格特性の測定に，人間向けの性格診断テストを用いることが可能かどうかは，明らかではない．そこで，出力される文章の表現と内容から性格特性の再現度を評価する．具体的には，思考発話なし学習モデルと思考発話あり学習モデルの出力を比較し，どちらが再現対象者の性格特性をより反映したものとなっているかを，人手で評価する．人手評価の材料となる文章を生成させる題材として，TIPI-J [20] の質問を用いた．TIPI-J にはビッグファイブの各次元に対応する質問が各 2 項目，計 10 項目収録されている．本評価では，モデルに TIPI-J の各質問への回答とその理由を生成させた．

得られた 10 組の質疑応答 (質問とモデルの回答とその理由) を人間の評価者に提示し，以下の手順で比較を行った．

1. 対象人物のビッグファイブのレーダーチャートと，比較対象として RealPersonaChat 内で対象人物から最もビッグファイブが乖離した別の人物のレーダーチャートの 2 種類を用意する．ただし，すでに再現対象者となっている，あるいは他の再現対象者の比較対象に選ばれている場合は，次にビッグファイブが離れている人物を比較対象とする．
2. 評価者は，モデルが提示した質疑応答の内容から示唆される性格特性に対して，2 つのチャートからより適合していると考えられる一方を選択する．

評価者が選んだチャートの人物と，モデルが再現対象としている人物が同一であった場合を「一致」，異なる場合を「不一致」と定義する．

人手による評価では，日本語を母語とする学生 5 名 (大学院生 4 名，学部生 1 名) に評価を無償で

表 2　発話の再現性評価の結果

設定	BERTScore	ROUGE-1	ROUGE-2	ROUGE-L
gpt-4o-mini				
ベースモデル	0.704	0.233	0.066	0.193
思考発話なし	**0.749**	**0.315**	**0.134**	**0.291**
思考発話あり	0.742	0.303	0.124	0.278
Llama-3-Swallow-8B-Instruct-v0.1				
ベースモデル	0.699	0.226	0.063	0.194
思考発話なし	**0.741**	**0.299**	**0.121**	**0.273**
思考発話あり	0.740	0.293	0.116	0.270
Qwen2.5-7B-Instruct				
ベースモデル	0.654	0.155	0.030	0.120
思考発話なし	**0.735**	**0.280**	**0.106**	**0.255**
思考発話あり	0.734	0.277	0.104	0.252
gemma-2-9b-it				
ベースモデル	0.701	0.219	0.061	0.183
思考発話なし	**0.738**	**0.289**	**0.113**	**0.265**
思考発話あり	0.736	0.285	0.110	0.261

表 3　性格診断テストの結果

設定	gpt	Swallow	Qwen	gemma
ベースモデル	1.252	**1.126**	1.140	2.075
思考発話なし学習	1.163	1.424	1.500	1.627
思考発話あり学習	**1.013**	1.178	**0.974**	**1.491**

依頼した．評価者は，再現対象の 20 名について，Qwen2.5-7B-Instruct をベースモデルとした思考発話なし学習によるモデルと思考発話あり学習によるモデルの 2 通り，計 40 件について評価を行った．最終的に 5 名の評価者から合計 200 件の評価を得た．

5　結果と考察

5.1　発話の再現性評価

表 2 に各モデルの発話の再現性評価の結果を示す．ベースモデルと比較し，対話データで学習したモデルは，思考発話の有無を問わず対象人物により近似した発話を生成できていた．思考発話の有無で比較すると，思考発話なし学習がわずかに高い値を示した．ここから，思考発話付与済みのデータでの学習は，対象人物の発話の再現に対する副作用は小さいことがわかる．

5.2　性格診断テストでの評価

表 3 に，性格診断テストにおけるモデルのスコアと対象人物のスコアの MSE を示す．思考発話あり学習は，思考発話なし学習と比較して対象人物の性格特性に近いスコアを示した．ビッグファイブの次元ごとの誤差は付録 B に記載した．Llama-3-Swallow-8B-Instruct-v0.1 と Qwen2.5-7B-Instruct においては，思考発話あり学習となし学習の MSE 間に統計的に有意な差[8)]があった．

一部の条件ではベースモデルが最も対象人物の性格特性に近いスコアを示す場合があった．これは，再現対象の人物の対話データを用いた学習が必ずしも性格特性の再現に寄与しないことを示唆している．一方，思考発話を付与したデータを用いた場合には，再現度の低下はおこらず性格特性の再現度が向上した．

表 4　5 人の評価者による人手評価の結果

設定	一致	不一致
思考発話なし学習	43	57
思考発話あり学習	68	32

5.3　人手による性格特性評価

表 4 に人手による性格特性評価を結果を示す．思考発話なし学習に比べて，思考発話あり学習のほうが一致となった割合が高かった．この結果は，思考発話あり学習が思考発話なし学習と比較して，対象人物の性格特性に近い応答を生成できていることを示唆している．この差は統計的に有意[9)]であった．

6　おわりに

本研究では，思考発話を付与した対話データを用いて，対象人物の発話と性格特性を学習する手法を提案した．実験の結果から，発話においては思考発話なし学習と同等の類似度が得られた．性格特性については，性格診断テストと人手評価の両面で，思考発話なし学習よりも再現対象者に近いモデルを構築できることが確認できた．

一方で課題としては，思考発話自体の質を評価できていないことが挙げられる．本研究では，思考発話付きの対話を用いた学習が，性格特性の再現性向上に寄与するかを検証した．しかし，思考発話が実際に対象人物の思考としてどの程度妥当であるかについては評価していない．考え方や思想といった深層的な部分を模倣したモデルを作成するためには，思考発話の妥当性を評価するシステムを検討する必要がある．

8）　並べ替え検定（有意水準 2.5%，片側検定）
9）　マクネマー検定 (有意水準 5%，両側検定)

謝辞

名古屋大学大学院情報学研究科の山下紗苗氏には，RealPersonaChat に関する貴重な情報提供をいただきました．感謝申し上げます．本研究は，JST，さきがけ，JPMJPR236B，JPMJPR236C の支援を受けたものです．

参考文献

[1] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. Character-LLM: A trainable agent for role-playing. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, 2023.

[2] Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy, and Jad Kabbara. PersonaLLM: Investigating the ability of large language models to express personality traits. In **Findings of the Association for Computational Linguistics: NAACL 2024**, 2024.

[3] Xiaoyang Song, Akshat Gupta, Kiyan Mohebbizadeh, Shujie Hu, and Anant Singh. Have large language models developed a personality?: Applicability of self-assessment tests in measuring personality in llms, 2023. https://arxiv.org/abs/2305.14693.

[4] Ivar Frisch and Mario Giulianelli. LLM agents in interaction: Measuring personality consistency and linguistic alignment in interacting populations of large language models. In **Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024)**, 2024.

[5] Lewis R. Goldberg. An alternative "description of personality": The big-five factor structure. **Journal of Personality and Social Psychology**, Vol. 59, No. 6, pp. 1216–1229, 1990.

[6] Robert R. McCrae and Oliver P. John. An introduction to the five-factor model and its applications. **Journal of Personality**, Vol. 60, No. 2, pp. 175–215, 1992.

[7] Jia Deng, Tianyi Tang, Yanbin Yin, Wenhao Yang, Wayne Xin Zhao, and Ji-Rong Wen. Neuron-based personality trait induction in large language models, 2024. https://arxiv.org/abs/2410.12327.

[8] Minjun Zhu, Linyi Yang, and Yue Zhang. Personality alignment of large language models, 2024. https://arxiv.org/abs/2408.11779.

[9] Wenkai Li, Jiarui Liu, Andy Liu, Xuhui Zhou, Mona Diab, and Maarten Sap. Big5-chat: Shaping llm personalities through training on human-grounded data, 2024. https://arxiv.org/abs/2410.16491.

[10] K. Anders Ericsson and Herbert A. Simon. **Protocol Analysis: Verbal Reports as Data**. The MIT Press, 1993.

[11] Noah Wang, Z.y. Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou, Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, Man Zhang, Zhaoxiang Zhang, Wanli Ouyang, Ke Xu, Wenhao Huang, Jie Fu, and Junran Peng. RoleLLM: Benchmarking, eliciting, and enhancing role-playing abilities of large language models. In **Findings of the Association for Computational Linguistics: ACL 2024**, 2024.

[12] Cheng Li, Ziang Leng, Chenxi Yan, Junyi Shen, Hao Wang, Weishi MI, Yaying Fei, Xiaoyang Feng, Song Yan, HaoSheng Wang, Linkang Zhan, Yaokai Jia, Pingyu Wu, and Haozhen Sun. Chatharuhi: Reviving anime character in reality via large language model, 2023. https://arxiv.org/abs/2308.09597.

[13] Sanae Yamashita, Koji Inoue, Ao Guo, Shota Mochizuki, Tatsuya Kawahara, and Ryuichiro Higashinaka. RealPersonaChat: A realistic persona chat corpus with interlocutors' own personalities. In **Proceedings of the 37th Pacific Asia Conference on Language, Information and Computation**, 2023.

[14] Qwen : An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. Qwen2.5 technical report, 2025. https://arxiv.org/abs/2412.15115.

[15] Naoaki Okazaki, Kakeru Hattori, Hirai Shota, Hiroki Iida, Masanari Ohi, Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Rio Yokota, and Sakae Mizuki. Building a large japanese web corpus for large language models. In **First Conference on Language Modeling**, 2024.

[16] Gemma Team. Gemma 2: Improving open language models at a practical size, 2024. https://arxiv.org/abs/2408.00118.

[17] Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. Qlora: Efficient finetuning of quantized llms. In **Advances in Neural Information Processing Systems**, 2023.

[18] 和田さゆり. 性格特性用語を用いた big five 尺度の作成. 心理学研究, Vol. 67, No. 1, pp. 61–67, 1996.

[19] Akshat Gupta, Xiaoyang Song, and Gopala Anumanchipalli. Self-assessment tests are unreliable measures of LLM personality. In **Proceedings of the 7th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP**, 2024.

[20] 小塩真司, 阿部晋吾, Pino Cutrone. 日本語版 ten item personality inventory（tipi-j）作成の試み. パーソナリティ研究, Vol. 21, No. 1, pp. 40–52, 2012.

A　思考発話付与のプロンプト

思考発話を付与するプロンプト[10]は表 5 のとおりである．ビッグファイブの各次元に関する説明は Goldberg[5] などを参考にして作成した．

B　ビッグファイブの次元別分析

表 6　ビッグファイブの次元別の MSE（O：開放性，C：誠実性，E：外向性，A：協調性，N：神経症傾向）

設定	O	C	E	A	N
gpt-4o-mini					
ベースモデル	1.225	1.097	1.575	0.512	1.849
思考発話なし学習	**0.823**	0.886	1.648	**0.487**	1.971
思考発話あり学習	0.949	**0.873**	**1.259**	0.551	**1.432**
Llama-3-Swallow-8B-Instruct-v0.1					
ベースモデル	**0.738**	**0.947**	1.558	**0.518**	1.868
思考発話なし学習	1.244	1.451	1.792	0.736	1.897
思考発話あり学習	0.794	1.228	**1.477**	0.690	**1.700**
Qwen2.5-7B-Instruct					
ベースモデル	0.967	**0.777**	1.636	0.563	**1.757**
思考発話なし学習	1.557	1.168	1.821	0.623	2.329
思考発話あり学習	**0.804**	0.876	**0.920**	**0.497**	1.774
gemma-2-9b-it					
ベースモデル	1.402	1.405	2.094	0.953	4.522
思考発話なし学習	1.556	**0.980**	1.929	**0.636**	3.034
思考発話あり学習	**1.085**	1.208	**1.499**	0.783	**2.879**

表 6 は，ビッグファイブ（O：開放性，C：誠実性，E：外向性，A：協調性，N：神経症傾向）の各次元に対する平均二乗誤差（MSE）の結果を示したものである．開放性（O），誠実性（C）及び協調性（A）に関しては，特定のモデルや設定による顕著な有意性は確認されなかった．一方で，外向性（E）と神経症傾向（N）（Qwen2.5-7B-Instruct を除いて）については，思考発話あり学習が最も小さい MSE を示し，より対象人物の性格に近づく傾向が観察された．

この結果から，ビッグファイブの次元によって思考発話の有効性が異なりうることがわかる．特に，外向性や神経症傾向といった特性は，思考発話を通じて対象人物の心理的背景をより的確に捉えられる可能性がある．

10）実験の段階では「思考発話」を「内心描写」と呼んでいたため，プロンプト内では内心描写と記載されている．

表 5　思考発話付与のプロンプト
（例 1 の一部と例 2 は省略）

Role	Content
system	## 基本情報 あなたは {id} です。あなたの性別や BigFive 性格特性は次のようになっています。 性別: {gender} BigFive 性格特性（1 が最小値で 7 が最大値）: {personality_list} 開放性とは、新しいアイデアや経験に対する興味や好奇心の程度を示します。開放性が高い人は、創造性や想像力に富み、新しいことに対して積極的に取り組みます。 誠実性とは、自己規律や責任感、目標達成への意欲の程度を示します。誠実性が高い人は、自己管理能力が高く、信頼性があります。 外向性とは、社交性や活動性、ポジティブな感情の程度を示します。外向性が高い人は、社交的で積極的であり、他者との関係を楽しむことが多いです。 協調性とは、他者への配慮や協力、共感の程度を示します。協調性が高い人は、他者との関係を大切にし、の感情やニーズに敏感です。 神経症傾向とは、不安や抑うつ、ストレス耐性の程度を示します。神経症傾向が高い人は、感情の起伏が激しく、ストレスに弱い傾向があります。
user	## タスクの説明 1. {id} の内心描写を追加してください。 内心描写は、{id} が思考していることや感じていることを表現するものです。 2. 内心描写は、「{id} (thinking):」の行に追加してください。 3. {id} の BigFive 性格特性を「必ず」内心描写に反映してください。 ## フォーマットの説明 次に、もとの対話履歴と出力のフォーマットの例を示します。 対話履歴の例 1: {partner_id} (speaking): ... {id} (thinking): {id} (speaking): ... {partner_id} (speaking): ... 出力の例 1: {partner_id} (speaking): ... {id} (thinking): (具体的な感情や考え) {id} (speaking): ... {partner_id} (speaking): ... ## タスク 以下の対話履歴を読んで、{id} の内心描写を追加してください。 対話履歴: {dialogue} 出力: