

Empath: Enhancing LLMs' Empathy in Emotional Support Conversation with Group MoELoRA and Experience RAG

Anonymous ACL submission

Abstract

This paper aims to enhance the empathetic capabilities of Emotional Support Conversation (ESC) models. To this end, we propose *Empath*, a novel ESC framework grounded in Bohart and Greenberg's Empathy Theory (Bohart, 1997), designed to improve Person, Affective, and Cognitive Empathy. At its core, *Empath* features the Group MoELoRA architecture and Experience RAG. Group MoELoRA personalizes support by tailoring character perspectives and dynamically adjusting support strategies based on emotional and contextual cues, while Experience RAG enriches interactions by aligning seeker concerns with relevant counselor experiences for deeper understanding. To train *Empath*, we introduce EmpathSupport-52k, a large-scale, multi-role, multi-strategy psychological counseling dataset. Extensive experiments demonstrate that *Empath* surpasses baseline models in both empathetic engagement and emotional support effectiveness.

1 Introduction

"It is easier to place oneself in the other's cognitive space if one has actually been there."
— Adam Smith

Mental health is an essential component of overall health, directly influencing people's well-being, work efficiency, and social adaptability. In the 21st century, with the acceleration of life pace and the increase in social pressure, mental health problems have become even more prominent. According to the World Mental Health Report released by the WHO in 2022, approximately 1/8 people globally suffer from mental disorders (Organization, 2022). However, due to the high cost of psychological counseling and the shortage of psychological counselors, the accessibility and availability of mental health services have been severely restricted. This severe situation highlights the importance of ap-

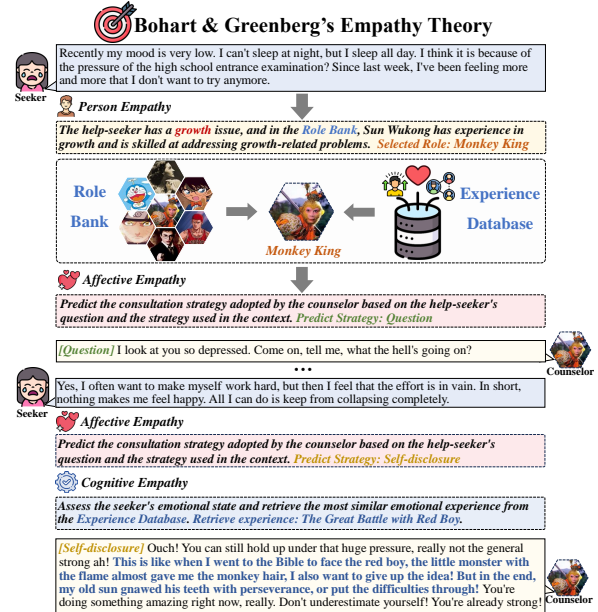


Figure 1: An example of the psychological counseling process: Building on Bohart and Greenberg's Empathy Theory, we enhance the model's empathy across the Person, Affective, and Cognitive dimensions. The framework selects the most suitable counseling role from the role bank, dynamically adapts strategies based on consultation progress, and retrieves the counselor's relevant experiences to foster deeper empathy throughout the conversation.

plying artificial intelligence (AI) technology in the field of mental health.

Fortunately, the rapid development of natural language processing (NLP) technology has provided a potential solution to this core issue. In particular, the fast advancement of large language models (LLMs), such as GPT-4 (Achiam et al., 2023), Claude-3.5 (Anthropic, 2024), and Llama3 (Dubey et al., 2024), has achieved remarkable progress in dialogue generation, offering seekers a certain level of comfort and suggestion. However, the response style of these models is generally inclined toward that of a neutral assistant, lacking distinctive lan-

guage styles and deep empathy, making it difficult to meet seekers’ emotional support needs. To address this limitation, Liu *et al.* (Liu et al., 2021) conducted innovative research based on Hill’s Helping Skills Theory (Hill, 2020) of psychological support. By fine-tuning emotional datasets annotated with support strategies, the study endowed dialog systems with emotional support capabilities.

Nevertheless, previous work has rarely focused on enhancing the model’s **empathetic capabilities** grounded in comprehensive and professional empathy theory, even though empathy is a key factor in improving the effectiveness of psychological counseling (Watson, 2016; Constantino et al., 2008; Bohart, 1997; Watson and Greenberg, 2011; Sandage and Worthington Jr, 2010). This limitation makes it difficult for existing psychological counseling systems to effectively evoke deep emotional resonance in those seeking help, diminishing their experience and the overall support effectiveness.

To address these challenges, we propose **Empath**, an emotional support framework grounded in Bohart and Greenberg’s (B&G) Empathy Theory (Bohart, 1997), which defines empathy across three dimensions: Person Empathy (understanding seeker’s background), Affective Empathy (adapting to seeker’s emotional changes), and Cognitive Empathy (comprehending seeker’s thoughts and experiences). The full definition is provided in Appendix A. As shown in Figure 1, **Empath** enhances psychological counseling by (1) tailoring character perspectives for personalized support, (2) dynamically adjusting support strategies based on emotional and contextual cues, and (3) aligning the seeker’s concerns with relevant counselor experiences to foster deeper understanding. **Empath**’s core comprises the Group MoELoRA architecture and Experience RAG. Group MoELoRA employs decoupled Mixture-of-Experts (MoE) to fine-tune role-specific and strategy-driven responses independently, ensuring not only controllability of each content but also seamless integration of the two during dialogue generation. **Empath** selects the optimal counseling role from a predefined library, predicts adaptive support strategies to drive the Group MoE, and incorporates experiential input via Experience RAG to enhance empathetic engagement.

To train **Empath**, we propose a three-stage prompt-based data derivation method leveraging GPT-4o¹ for language style transfer and ex-

¹<https://openai.com>

Criteria	Statistics
# Datasets Length	52,782
# Turns per dialogue	8.68
# Characters per dialogue	1,201.52
# Characters per seeker text	30.43
# Characters per counselor text	93.05

Table 1: Statistics of EmpathSupport-52K dataset.

perience modeling on the PsyQA dataset (Sun et al., 2021), resulting in EmpathSupport-52k(ES-52k)—a large-scale, multi-role, multi-strategy psychological counseling dataset with explicit character experiences, that advances research on empathetic AI systems.

In summary, our contributions are mainly summarized as follows:

- A Principled Solution Rooted in Classic Theory: We introduce **Empath**, a novel emotional support framework inspired by B&G Empathy Theory, effectively enhancing traditional methods in Person Empathy, Affective Empathy, and Cognitive Empathy.
- A Novel Network Design: Group MoELoRA enables the decoupling of character styles and support strategies, allowing for independent control while naturally integrating both to generate empathy dialogues. Additionally, Experience RAG enriches conversations by incorporating role-specific experiences, further strengthening the model’s empathetic capabilities.
- A New Large-Scale Dataset: We construct EmpathSupport-52k, a multi-role, multi-strategy psychological counseling dataset that significantly advances research on empathetic AI systems.
- State-of-the-Art Performance: **Empath** achieves substantial improvements across multiple evaluation metrics, including automatic, model-based, and human assessments. It demonstrates superior Role-based empathetic capabilities and enhances the diversity and effectiveness of emotional support, as detailed in our experimental results.

2 EmpathSupport-52K Dataset

In this section, we present our three-stage prompting process for dataset construction, as shown in



Figure 2: Distribution of counseling topics.

Figures 10, 11, and 12, detailed in Appendix C.1. The GPT-4o model transforms the existing psychological counseling dataset into ES-52k, incorporating role-specific talking styles and life experiences. The original text and generated results are shown in Figures 13, 14, 15, and 16 in the Appendix. Further dataset details are provided in Appendix C.

2.1 Data Source

We selected the PsyQA (Sun et al., 2021) dataset as the original source. PsyQA is a high-quality Chinese psychological counseling dataset from the YiXinLi platform², covering nine major topics with 22,000 questions and 56,000 responses. We proportionally sampled 14,000 single-turn QA pairs from these topics as the raw data.

2.2 Data Process

To enhance empathy capabilities, we propose a three-stage prompting dataset derivation method, which further optimizes the PsyQA dataset.

Converting Single-Turn to Multi-Turn. Single-turn QA pairs in vanilla PsyQA do not fully capture the interactive nature of psychological counseling dialogues. To address this, we adapted methods like Smile (Qiu et al., 2023) and designed a first-stage prompt that guided GPT-4o to convert single-turn dialogues into approximately 10-turn multi-turn counseling exchanges.

Incorporating Role Experiences. We collected role-specific experiences from TV scripts and used GPT-4o to transform them into first-person narratives. These role experiences were integrated

into the second-stage prompt, ensuring they were applied only within specific strategies (e.g., *self-disclosure*, *affirmation*, *reassurance*).

Assigning Role-Specific Styles and Personalities. We created a role bank containing profiles, speaking styles, and background information for each role. A style rewriting guidebook was included in the third-stage prompt, directing GPT-4o to imbue the supporter with a unique personality and speaking style that aligned with the role’s characteristics.

2.3 Data Filtering

To ensure the quality of the derived dataset, we applied rigorous screening to the 126,000 entries in the ES-52k. First, we conducted format checks to confirm adherence to the one-on-one counseling dialogue structure, and filtered responses to align with the eight predefined support strategies. Next, we screened dialogues based on length, retaining only those with 8-10 turns. Finally, using the data-juicer tool (Chen et al., 2024), we excluded samples with perplexity (PPL) above 1,500. After filtering, we obtained 52,782 high-quality entries in ES-52k.

Table 1 and Appendix C.3 present the statistics for ES-52k, with an average of 8.68 turns per conversation. Figure 2 shows the distribution of topics within ES-52k.

3 Preliminary: MoELoRA

MoELoRA combines the Mixture-of-Experts (MoE) architecture with Low-Rank Adaptation (LoRA) to improve efficiency and resource utilization. As illustrated in Figure 3(c), it utilizes MoE’s routing mechanism to dynamically select experts, while employing LoRA to update parameters.

In MoE, the input x is routed to N most suitable experts $\{E_1, E_2, \dots, E_N\}$, where each expert processes a portion of the input, and the results are combined. The output o of the forward propagation process in MoE is expressed as:

$$o = \sum_{i=1}^N G(x_i) \cdot E_i(x), \quad (1)$$

where $G(x_i)$ denotes the activation degree of the i -th expert, and $E_i(x)$ is the output of the i -th expert.

LoRA reduces the number of parameters to be trained by introducing two low-rank matrices, A and B , which replace the update of the original parameter matrix in LLMs. The forward propagation

²<https://www.xinli001.com>

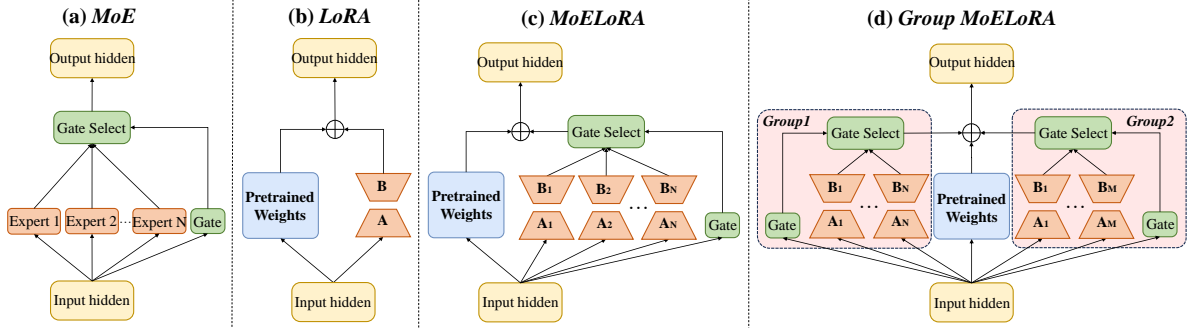


Figure 3: The Different Architectures for (a) MoE, (b) LoRA, (c) MoELoRA and (d) Group MoELoRA.

during fine-tuning in LoRA is expressed as:

$$o = W_0x + \Delta Wx = W_0x + \frac{\alpha}{r} \cdot BAx, \quad (2)$$

where W_0 is the pre-trained weight matrix, ΔW is replaced by the low-rank matrices A and B , and α is a hyperparameter adjusting the impact of the low-rank matrices.

By combining MoE and LoRA, MoELoRA reduces computational and storage overhead while expanding model scale. The forward propagation process in MoELoRA for N experts is represented as:

$$o = W_0x + \Delta Wx = W_0x + \frac{\alpha}{r} \sum_{i=1}^N \omega_i \cdot B_i A_i x, \quad (3)$$

where ω_i denotes the weight of the i -th expert. This combination reduces parameter count and computational complexity, improving the efficiency of large-scale model training and inference.

4 Empath Framework

This section introduces the **Empath** framework, illustrated in Figure 4, which is grounded in B&G’s Empathy Theory. It encompasses role matching, strategy prediction, and experience-driven empathy. Section 4.1 introduces the **Group MoELoRA** dynamic routing architecture, which refines role features and strategy weights through a grouped gating mechanism, addressing both the person and affective dimensions. Section 4.2 introduces the **Role-based Experience RAG**, designed to enhance experience-based empathy with a focus on the cognitive dimension. Section 4.3 discusses the advantages of Group MoELoRA compared to MoELoRA and LoRA. Please refer to Appendix B for more details on role alignment and strategy forecasting utilizing BERT.

4.1 Group MoELoRA

One of the key challenges in designing a multi-role psychological counseling system is matching the optimal counselor to the help-seeker based on their emotional needs and the nature of their issues, while dynamically adjusting response strategies according to the interaction context. To address this, we propose a solution based on the Group MoELoRA architecture, which integrates multiple roles and response strategies into a single model through two sets of MoELoRA structures. This approach aims to enable seamless role switching and predict the most appropriate response strategy for each role, thereby delivering efficient and personalized emotional support.

As illustrated in Figure 4, the Group MoELoRA architecture enhances the traditional MoELoRA approach by decoupling the N counselor roles and M support strategies, processing these two types of information independently. The system employs two distinct MoELoRA modules: one dedicated to selecting the appropriate counseling role and the other for choosing the corresponding response strategy. Specifically, the input role and strategy embedding vectors are passed through two separate gating networks to activate the relevant counseling roles and support strategies. Each role and strategy is associated with an independent expert sub-model, and these modules are dynamically integrated to produce the final output.

Group Gating. In contrast to the vanilla MoELoRA that assigns expert weights based on the input x , Group MoELoRA utilizes role profiles and strategy definitions as inputs to the gating network to determine the expert weights in each group. This design draws inspiration from as well as extends the Neeko (Yu et al., 2024) model, which dynamically calculates the contribution of each expert.

Specifically, let N represent the number of role

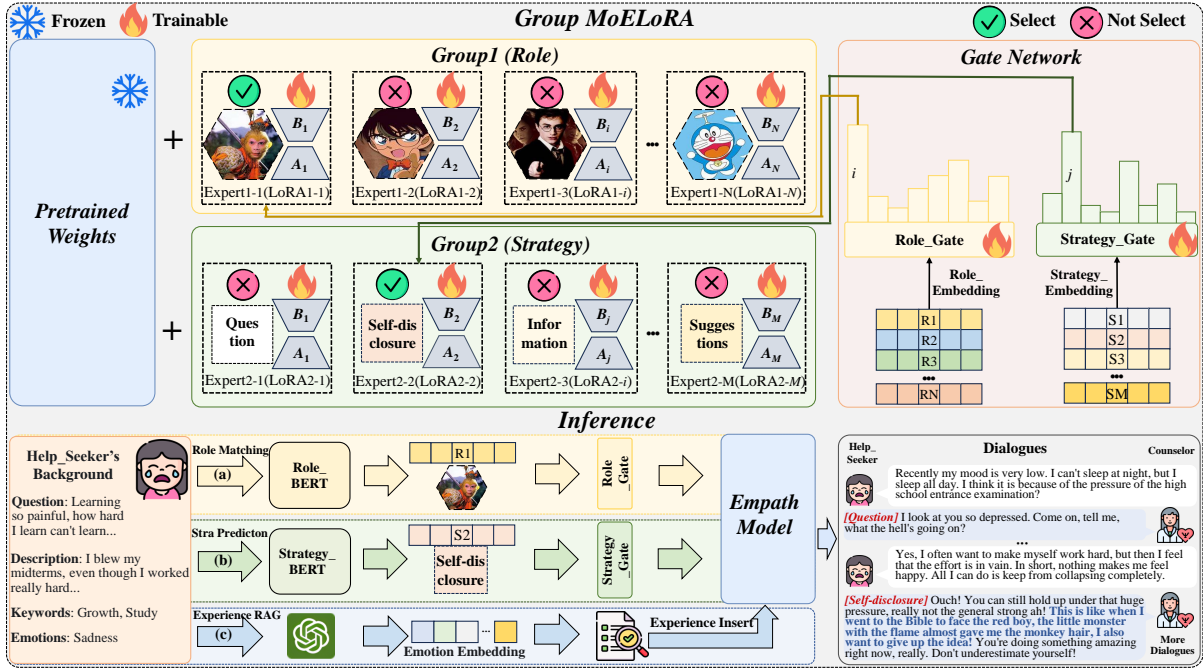


Figure 4: The **Empath** framework consists of the Group MoELoRA architecture and three inference components: (a) Role Matching, (b) Strategy Prediction, and (c) Role-based Experience RAG. Group MoELoRA embeds roles and strategies into two separate MoELoRA groups, with their output weights dynamically controlled by the Gate.

profiles and M the number of strategy descriptions, each encoded by two global embedding matrices: $E_{\text{role}} \in \mathbb{R}^{N \times d}$ and $E_{\text{strategy}} \in \mathbb{R}^{M \times d}$, where d denotes the embedding dimension. Each row of these matrices corresponds to a distinct embedding. Thus, the embeddings for the i -th role and the j -th strategy are expressed as:

$$\begin{cases} e_{\text{role}}^i = E_{\text{role}}[i, :] & e_{\text{role}}^i \in \mathbb{R}^d \\ e_{\text{strategy}}^j = E_{\text{strategy}}[j, :] & e_{\text{strategy}}^j \in \mathbb{R}^d. \end{cases} \quad (4)$$

Next, we apply linear transformations to the role and strategy embeddings using weight matrices $W_{\text{role_gate}} \in \mathbb{R}^{N \times d}$ and $W_{\text{strategy_gate}} \in \mathbb{R}^{M \times d}$ to compute each expert’s contribution weight for role r_i and strategy s_j :

$$\begin{cases} w_i = \text{Softmax}(W_{\text{role_gate}} \cdot e_{\text{role}}^i) \\ w_j = \text{Softmax}(W_{\text{strategy_gate}} \cdot e_{\text{strategy}}^j). \end{cases} \quad (5)$$

Role-Strategy Group Activation. In this section, we describe the independent activation of counseling roles and response strategies using two separate MoELoRA groups. Consider N counseling roles and M support strategies. Each role and strategy is trained and updated using the LoRA structure. For each role or strategy, the input embedding vector e_* passes through its respective gating network $G_*(x)$ to generate the activation weight $G_*(x)_k$ for the k -th expert.

The output $o_*(x)$ for each role or strategy is:

$$\begin{aligned} o_*(x) &= W_0 x + \sum_{k=1}^{N/M} G_*(x)_k \cdot E_k(x) \\ &= W_0 x + \frac{\alpha}{r} \sum_{k=1}^{N/M} \omega_{*k} \cdot B_k A_k x, \end{aligned} \quad (6)$$

where ω_{*k} denotes the weight assigned to each role or strategy for the k -expert, and A_k and B_k are the low-rank matrices. Here, $*$ is a placeholder for either role or strategy, depending on the group.

This approach independently activates the most relevant role and strategy based on the input, ensuring optimal emotional support for the help-seeker.

4.2 Role-based Experience RAG

This section introduces our Role-based Experience RAG method to improve the model’s cognitive empathy, as illustrated in Figure 5. To bridge the cognitive gap between the seeker and the consultant, we use experience as a link for alignment.

First, we collect the plot from scripts or movies related to each character’s role. For the i -th role, its experience P_i is transformed into an emotional embedding vector v_i , the detailed content of the prompt and emotional dimensions are provided in Figure 9 in the Appendix. Then, the emotional vectors of all roles and their experiences are combined

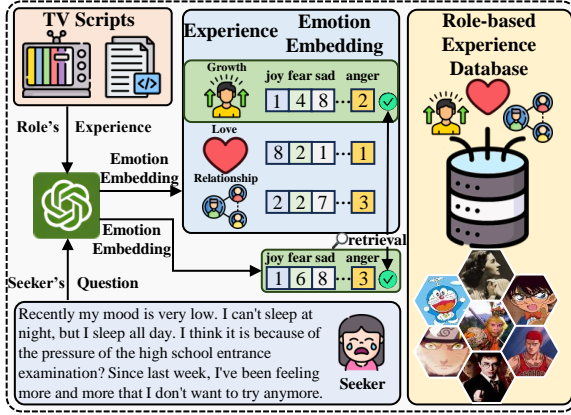


Figure 5: Role-based Experience RAG.

into a vector database V , represented by:

$$V = \{(v_1, P_1), (v_2, P_2), \dots, (v_t, P_t)\}. \quad (7)$$

During inference, given the j -th seeker's background X_j , we compute its emotional vector v_j and calculate the cosine similarity between v_j and each of the selected role's emotional vector v_i in vector database V to find the best match:

$$\text{similarity}(v_j, v_i) = \frac{v_j \cdot v_i}{\|v_j\| \|v_i\|}. \quad (8)$$

The role experience with the highest similarity is chosen and added as the model's context prompt.

4.3 Discussion

The Group MoELoRA architecture extends the MoELoRA framework by decoupling the activation of roles and strategies, allowing multiple roles and strategies to be embedded simultaneously without the need for individual training of each. This significantly reduces computational costs and improves efficiency, as it eliminates the need to train separate models for each role and strategy. Furthermore, by utilizing gating networks and low-rank matrices, the architecture can dynamically switch between roles and strategies, offering more precise and context-aware emotional support. This flexibility also makes it easier to scale the model by adding new roles and strategies, ensuring it can be tailored to diverse emotional support scenarios while maintaining high performance.

5 Experiments

5.1 Baseline

To evaluate the performance of the proposed *Empath* framework, we performed comparative experiments with three types of models: closed-source, open-source and domain-specific:

- **Closed-source models:** ChatGPT (OpenAI, 2022), GPT-4 (Achiam et al., 2023).
- **Open-source models:** Qwen2.5-7B-Instruct (Team, 2024b), InternLM2.5-7B-Chat (Cai et al., 2024), Llama3.1-8B-Instruct (Touvron et al., 2023), Llama3.1-8B-Chinese-Chat (Wang et al., 2024), ChatGLM4-9B-Chat (GLM et al., 2024).
- **Domain-specific models:** Smile (Qiu et al., 2023), Soulchat (Chen et al., 2023), SoulChat2.0 (Xie et al., 2024), CPsyCounX (Zhang et al., 2024), ChatCounselor (Liu et al., 2023), EmoLLM V3.0 (Team, 2024a).

5.2 Implementation Details

We randomly selected 50k samples from ES-52k for training, 2k for test, and 200 from the test set for human evaluation. We utilize the GPT-4o-2024-11-20 model¹ to construct the dataset and emotion-experience pairs in the Role-based Experience DataBase. For training *Empath*, the Llama3.1-8B-Chinese-Chat model is selected as the base model, trained on a single 80GB A800 GPU. Specific hyperparameter settings are detailed in Table 5 in the Appendix E.

5.3 Evaluation Metrics

To comprehensively evaluate the model's response performance, we assess our model from three perspectives: automatic metrics, LLM evaluation, and human evaluation. The evaluation details are provided in Appendix F.

Automatic Evaluation: We used five automatic evaluation metrics to quantify the model's performance, including: Distinct-1 (Li et al., 2015), BLEU-1, BLEU-4 (Papineni et al., 2002), F1-Score, and Rouge-L (Lin, 2004).

LLM Evaluation: To more accurately evaluate the model's outputs, we referred to evaluation metrics proposed in related works (Chen et al., 2023; Zhang et al., 2024; Zhao et al., 2024) and added some new metrics based on our task requirements. This resulted in six evaluation metrics: *Fluency* (*Flu.*, 3 points), *Role Alignment* (*RoA.*, 5 points), *Skillfulness* (*Ski.*, 5 points), *Empathy* (*Emp.*, 5 points), *Helpfulness* (*Hel.*, 3 points), and *Safety* (*Saf.*, 1 point). The evaluation was conducted using GPT-4o-mini-2024-07-18 model¹. The definitions, scoring criteria, and evaluation prompts for each metric are provided in Table 6 in the Appendix F.

Models	Automatic Evaluation.					LLM Evaluation.					
	D-1	B-1	B-4	F1	R-L	Flu.	RoA.	Ski.	Emp.	Hel.	Saf.
<i>Cloesd-source Models</i>											
ChatGPT	39.79	19.05	2.11	24.86	16.61	2.05	3.72	3.27	3.52	2.09	1.00
GPT-4	34.19	17.32	1.88	18.94	13.52	<u>2.09</u>	3.67	3.22	3.51	2.12	1.00
<i>Open-source Models</i>											
Qwen2.5-7B-Instruct	40.60	18.03	2.00	23.14	13.66	2.05	3.68	3.20	3.41	2.15	1.00
InternLM2.5-7B-Chat	39.99	17.64	1.99	22.67	13.40	2.07	3.73	3.23	3.42	2.20	1.00
Llama3.1-8B-Instruct	57.68	14.55	1.63	18.96	12.11	1.95	3.31	3.12	3.27	2.14	1.00
Llama3.1-8B-Chinese-Chat	41.39	18.54	2.18	23.34	14.19	2.03	3.62	3.18	3.43	2.17	1.00
ChatGLM4-9B-Chat	38.78	18.37	2.15	23.42	14.08	2.04	3.66	3.16	3.40	2.25	1.00
<i>Domain-specific Models</i>											
MeChat	47.66	<u>23.68</u>	3.16	<u>26.97</u>	<u>17.00</u>	2.08	3.70	<u>3.30</u>	<u>3.73</u>	2.16	1.00
SoulChat	41.35	21.32	2.45	24.58	15.39	2.05	3.74	3.29	3.61	2.18	1.00
SoulChat2.0	51.09	22.16	<u>3.22</u>	26.89	16.98	2.11	<u>3.77</u>	3.23	3.68	2.11	1.00
CPsyCounX	38.95	22.19	2.96	26.70	16.44	1.96	3.68	3.24	3.66	2.09	1.00
ChatCounselor	46.12	19.57	1.95	22.56	14.07	1.84	3.24	3.06	3.30	1.97	1.00
EmoLLM-V3.0	<u>53.71</u>	19.32	2.61	24.16	15.57	2.02	3.66	3.24	3.66	2.12	1.00
<i>Empath</i>	46.82	25.55	4.09	30.08	19.06	2.11	3.82	3.34	3.81	<u>2.23</u>	1.00

Table 2: Results of automatic and LLM evaluation. The best score is **in-bold**, while the second best score is underlined. The green numbers show the model comparison based on the Empathy metric.

Human Evaluation: Human evaluation serves as a fine-grained supplement to the automated evaluation. We used a Win/Loss table to compare the best-performing baseline model with our proposed *Empath*. The implementation details are provided in the Appendix F.3.

5.4 Experimental Results

Comparison Study. The comparison results are summarized in Table 2, with the corresponding radar chart presented in Figure 8 in the Appendix. In the Automatic Evaluation, *Empath* significantly outperforms the Baselines across several metrics. In the LLM Evaluation, our *Empath* exceeds the Baselines under both Skillfulness and Empathy metrics, highlighting its superior performance in psychological counseling conversations. Furthermore, we complement these findings with a Human Evaluation, where 18 volunteers assessed the text generation quality of the top-performing Baseline model, Mechat, and our *Empath* model. The evaluation results presented in Figure 6 demonstrate that our model outperforms in terms of empathy capabilities.

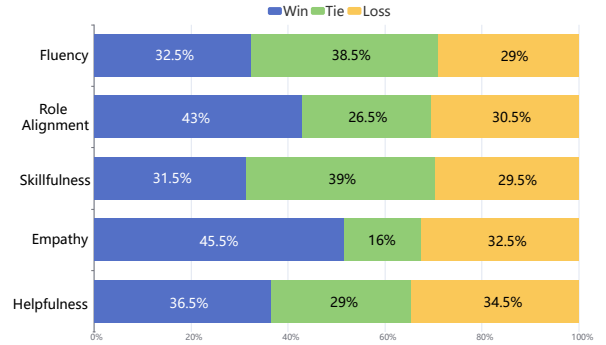


Figure 6: The Win/Loss comparison of *Empath* against *MeChat*, evaluated by human assessors.

5.5 Ablation Study

As discussed in the *Empath* framework (§4), we designed three key components: Role_Match, Strategy_Predict, and Role-based Experience RAG, to enhance the model’s empathy capabilities. To validate the contributions of these components, we conducted ablation studies by removing the role and strategy components, with roles and strategies selected according to their proportion for comparison. The results shown in Figure 7 indicate that

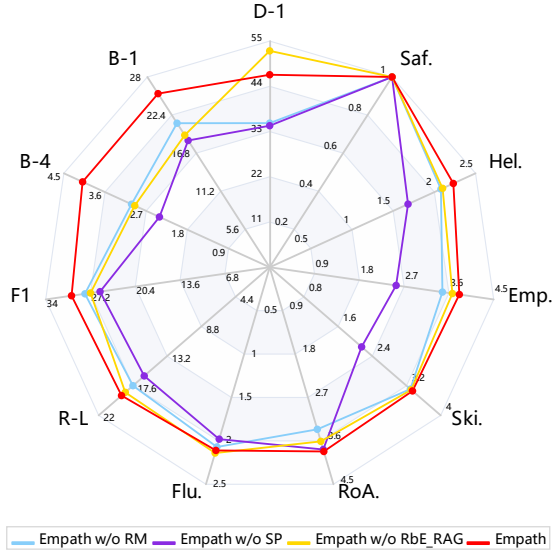


Figure 7: The comparison between Empath and Empath w/o Role_Predict, Empath w/o Strategy_Predict, and Empath w/o Role-based Experience RAG.

removing any component causes a significant drop in the Empathy metric. Notably, the absence of an appropriate consultation response strategy prior to text generation leads to a considerable decline across all metrics.

5.6 Case Study

In this section, we present a case analysis comparing the *Empath* model with baseline models. As shown in Figures 19 and 20 in the Appendix, base models like Qwen2.5-7B-Instruct lack sufficient skillfulness and empathy in counseling dialogues. In contrast, domain-specific models such as EmoLLM-V3.0, shown in Figures 21 and 22 in the Appendix, exhibit some level of empathy but struggle to manage the entire counseling process due to the absence of professional psychological theory support. However, as demonstrated in Figure 23 in the Appendix, our *Empath* model not only shows strong empathy but also effectively manages the full counseling process.

6 Related Work

6.1 Emotional Support Conversation Models

As large language models (LLMs) improve, their applications in psychological counseling have grown rapidly. Current research focuses on two main areas: 1. Generating and optimizing psychological counseling datasets (Liu et al., 2021; Sun et al., 2021; Qiu et al., 2023; Chen et al., 2023;

Zhang et al., 2024), which involves collecting Q&A data or using LLMs to create high-quality simulated datasets for model training; 2. Selecting and optimizing counseling strategies (Kang et al., 2024; Tu et al., 2022; Cai et al., 2023; Cheng et al., 2022; Jia et al., 2023), which involves studying counselor dialogue patterns and emotional support strategies, as well as developing automated methods for strategy selection.

6.2 Retrieval Augmented Generation

Retrieval-Augmented Generation (RAG) (Lewis et al., 2020) is a technique that enhances the performance of LLMs by integrating external knowledge bases. It consists of three main steps: indexing, retrieval, and generation. RAG improves the authenticity and usefulness of LLM outputs by retrieving relevant information from a knowledge base and using it as context. This technique is widely applied in fields such as medicine, business, and law, and has shown significant success in dialogue systems (Xiong et al., 2024; Huang et al., 2024; Louis et al., 2024), demonstrating strong practical value.

6.3 Mixture of Experts

Mixture of Experts (MoE) uses a "routing mechanism" to allocate input data to multiple sub-models (experts) for efficient computation, with a gating network selecting the best expert for training and inference. Recent studies (Liu et al., 2024; Luo et al., 2024) have combined MoE with multi-task learning to improve task collaboration. Some studies (Shen et al., 2023; Dou et al., 2024) also integrate Low-Rank Adaptation (LoRA) into MoE, replacing expert sub-models with LoRA modules to speed up training and improve resource utilization, yielding promising results.

7 Conclusion

In this paper, we propose the *Empath* framework, based on B&G's Empathy Theory, to enhance empathy across the Person, Affective, and Cognitive dimensions. To strengthen Person and Affective empathy, we introduce the EmpathSupport-52k, which includes diverse roles and strategies, integrated into a single model using the Group MoELoRA architecture. For Cognitive empathy, we present the Role-based Experience RAG method to bridge the cognitive gap between the seeker and the counselor. Extensive experiments demonstrate the model's potential in the emotional support domain.

Limitations

Our model relies on an external BERT model for pre-matching and predicting roles and strategies, with the predicted results then fed into the ESC model’s gating network to activate specific roles and strategies for completing emotional support dialogues. This process somewhat limits the model’s practicality in real-world scenarios. Future research could focus on integrating role matching and strategy prediction directly within the ESC model, significantly enhancing its usability and flexibility.

Ethical Considerations

Data Usage Agreement

This research utilizes the PsyQA (Sun et al., 2021) dataset, which has been obtained with proper authorization and in compliance with data usage agreements. We ensure that all data used in this study are handled responsibly and in accordance with ethical standards, respecting the privacy and confidentiality of individuals involved. All necessary agreements and permissions for the use of this dataset have been signed, ensuring full compliance with data protection regulations.

Model Usage Policy

It is important to note that there are significant differences between the strategies and modes used by the model for psychological support and those employed in real-life psychological counseling. Given the variance in user situations, the model’s outputs may not always align with professional therapeutic guidance and may inadvertently affect the emotional well-being of users. Therefore, the model should be strictly used for academic research purposes only, and its commercial use is prohibited. Caution must be exercised in applying the model outside controlled environments, and its use for actual therapeutic practices is not recommended without professional oversight.

References

- Josh Achiam, Steven Adler, and et al. Agarwal. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.
- AI Anthropic. 2024. The claude 3 model family: Opus, sonnet, haiku. *Claude-3 Model Card*, 1.
- AC Bohart. 1997. Empathy: Where are we and where do we go from here. *Empathy reconsidered: New*

directions in psychotherapy/American Psychological Association.

- Hua Cai, Xuli Shen, Qing Xu, Weilin Shen, and et al. Wang. 2023. Improving empathetic dialogue generation by dynamically infusing commonsense knowledge. *arXiv preprint arXiv:2306.04657*.
- Zheng Cai, Maosong Cao, Haojiong Chen, and et al. Kai Chen. 2024. *Internlm2 technical report. Preprint*, arXiv:2403.17297.
- Daoyuan Chen, Yilun Huang, and et al. Zhijian Ma. 2024. Data-juicer: A one-stop data processing system for large language models. In *International Conference on Management of Data*.
- Yirong Chen, Xiaofen Xing, Jingkai Lin, Huimin Zheng, Zhenyu Wang, Qi Liu, and Xiangmin Xu. 2023. Soulchat: Improving llms’ empathy, listening, and comfort abilities through fine-tuning with multi-turn empathy conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1170–1183.
- Jiale Cheng, Sahand Sabour, Hao Sun, Zhuang Chen, and Minlie Huang. 2022. Pal: Persona-augmented emotional support conversation generation. *arXiv preprint arXiv:2212.09235*.
- Michael J Constantino, Margaret E Marnell, Abigail J Haile, and et al. Kanther-Sista. 2008. Integrative cognitive therapy for depression: A randomized pilot comparison. *Psychotherapy: Theory, Research, Practice, Training*, 45(2):122.
- Yiming Cui, Wanxiang Che, Ting Liu, Bing Qin, Shijin Wang, and Guoping Hu. 2020. Revisiting pre-trained models for chinese natural language processing. *arXiv preprint arXiv:2004.13922*.
- Shihan Dou, Enyu Zhou, Yan Liu, Songyang Gao, Wei Shen, and et al. Xiong. 2024. Loramoe: Alleviating world knowledge forgetting in large language models via moe-style plugin. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1932–1945.
- Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, and et al. Kadian. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.
- Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, and et al. Chenhui Zhang. 2024. *Chatglm: A family of large language models from glm-130b to glm-4 all tools. Preprint*, arXiv:2406.12793.
- Clara E Hill. 2020. *Helping skills: Facilitating exploration, insight, and action*. American Psychological Association.
- Le Huang, Hengzhi Lan, Zijun Sun, Chuan Shi, and Ting Bai. 2024. Emotional rag: Enhancing role-playing agents through emotional retrieval. *arXiv preprint arXiv:2410.23041*.

- Mengzhao Jia, Qianglong Chen, Liqiang Jing, Dawei Fu, and Renyu Li. 2023. Knowledge-enhanced memory model for emotional support conversation. *arXiv preprint arXiv:2310.07700*.
- Dongjin Kang, Sunghwan Kim, Taeyoon Kwon, and et al. Moon, Seungjun. 2024. Can large language models be good emotional supporter? mitigating preference bias on emotional support conversation. *arXiv preprint arXiv:2402.13211*.
- Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, and et al. Karpukhin. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. 2015. A diversity-promoting objective function for neural conversation models. *arXiv preprint arXiv:1510.03055*.
- Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, pages 74–81.
- June M Liu, Donghao Li, He Cao, Tianhe Ren, Zeyi Liao, and Jiamin Wu. 2023. Chatcounselor: A large language models for mental health support. *arXiv preprint arXiv:2309.15461*.
- Qidong Liu, Xian Wu, Xiangyu Zhao, and et al. Zhu. 2024. When moe meets llms: Parameter efficient fine-tuning for multi-task medical applications. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1104–1114.
- Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. [Towards emotional support dialog systems](#). In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3469–3483, Online. Association for Computational Linguistics.
- Antoine Louis, Gijs van Dijck, and Gerasimos Spanakis. 2024. Interpretable long-form legal question answering with retrieval-augmented large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 22266–22275.
- Tongxu Luo, Jiahe Lei, Fangyu Lei, Weihao Liu, Shizhu He, Jun Zhao, and Kang Liu. 2024. Moelora: Contrastive learning guided mixture of experts on parameter-efficient fine-tuning for large language models. *arXiv preprint arXiv:2402.12851*.
- OpenAI. 2022. [Introducing chatgpt](#). Accessed: 2025-02-05.
- World Health Organization. 2022. *World mental health report: Transforming mental health for all*. World Health Organization.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.
- Huachuan Qiu, Hongliang He, Shuai Zhang, Anqi Li, and Zhenzhong Lan. 2023. Smile: Single-turn to multi-turn inclusive language expansion via chatgpt for mental health support. *arXiv preprint arXiv:2305.00450*.
- Steven Sandage and Everett Worthington Jr. 2010. Comparison of two group interventions to promote forgiveness: Empathy as a mediator of change. *Journal of Mental Health Counseling*, 32(1):35–57.
- Sheng Shen, Le Hou, Yanqi Zhou, Nan Du, Shayne Longpre, and et al. Wei. 2023. Mixture-of-experts meets instruction tuning: A winning combination for large language models. *arXiv preprint arXiv:2305.14705*.
- Hao Sun, Zhenru Lin, Chujie Zheng, Siyang Liu, and Minlie Huang. 2021. Psyqa: A chinese dataset for generating long counseling text for mental health support. *arXiv preprint arXiv:2106.01702*.
- EmoLLM Team. 2024a. Emollm: Reinventing mental health support with large language models. <https://github.com/SmartFlowAI/EmoLLM>.
- Qwen Team. 2024b. [Qwen2.5: A party of foundation models](#).
- Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
- Quan Tu, Yanran Li, Jianwei Cui, Bin Wang, Ji-Rong Wen, and Rui Yan. 2022. [MISC: A mixed strategy-aware model integrating COMET for emotional support conversation](#). In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 308–319, Dublin, Ireland. Association for Computational Linguistics.
- Shenzhi Wang, Yaowei Zheng, Guoyin Wang, Shiji Song, and Gao Huang. 2024. [Llama3-8b-chinese-chat \(revision 6622a23\)](#).
- Jeanne C Watson. 2016. The role of empathy in psychotherapy: Theory, research, and practice.
- Jeanne C Watson and Leslie S Greenberg. 2011. 10 empathic resonance: A neuroscience perspective. *The social neuroscience of empathy*, page 125.
- Haojie Xie, Yirong Chen, Xiaofen Xing, Jingkai Lin, and Xiangmin Xu. 2024. Psydt: Using llms to construct the digital twin of psychological counselor with personalized counseling style for psychological counseling. *arXiv preprint arXiv:2412.13660*.

- Guangzhi Xiong, Qiao Jin, Zhiyong Lu, and Aidong Zhang. 2024. Benchmarking retrieval-augmented generation for medicine. *arXiv preprint arXiv:2402.13178*.
- Xiaoyan Yu, Tongxu Luo, Yifan Wei, Fangyu Lei, Yiming Huang, Hao Peng, and Liehuang Zhu. 2024. Neeko: Leveraging dynamic lora for efficient multi-character role-playing agent. *arXiv preprint arXiv:2402.13717*.
- Chenhao Zhang, Renhao Li, Minghuan Tan, and et al. Yang, Min. 2024. Cpsycoun: A report-based multi-turn dialogue reconstruction and evaluation framework for chinese psychological counseling. *arXiv preprint arXiv:2405.16433*.
- Haiquan Zhao, Lingyu Li, Shisong Chen, Shuqi Kong, and et al. Wang. 2024. Esc-eval: Evaluating emotion support conversations in large language models. *arXiv preprint arXiv:2406.14952*.

A Bohart and Greenberg’s Empathy Theory

Bohart and Greenberg (Bohart, 1997) differentiated among three types of empathy: **person**, **affective**, and **cognitive**.

1. **Person empathy**: Person empathy is an understanding of the whole person in situ. It requires an understanding of what clients have experienced, including their histories and life stories, to acquire a holistic understanding of who they are and how they developed.
2. **Affective empathy**: refers to being attuned to the affective experience of others as revealed by their body language and narratives to understand clearly the impact and significance of events for them.
3. **Cognitive empathy**: Cognitive empathy is the capacity to understand and make sense of clients’ narratives.

B Role Matching and Strategy Prediction

This section discusses role matching and strategy prediction methods based on B&G’s Empathy Theory to improve the model’s Person and Affective empathy. we employ a Chinese RoBERTa-large (Cui et al., 2020) model³. We randomly sample 10k data from the Empath-52k dataset, with an 8:1:1 train-validation-test split. For both models, we set the following parameters: $Epoch = 4$, $learningrate = 1e^{-5}$, and $batchsize = 16$.

Role Matching. As shown in Figure 4(A), we preprocess the dataset by concatenating the question (Q_i), the detailed description (D_i), and the topic keywords (K_i) of the i -th seeker to form the seeker’s background information X_i . Specifically, we represent this as:

$$X_i = [[Ques]Q_i; [Desc]D_i; [Keys]K_i] \quad (9)$$

where the semicolon (;) denotes text concatenation. Then, based on each role’s relevant experience and personality traits, we manually assign the most suitable golden role R_i for this background information. During training, we use the golden role R_i as the label to train the BERT model. This allows the model to predict the role Y_{pred} during inference based on the seeker’s background information X_i .

³<https://huggingface.co/hfl/chinese-roberta-wwm-ext-large>

Task	Categories	Prec.	Recall	F1
Role Matching	Sun Wukong	91.44	92.49	91.96
	Hanamichi Sakuragi	99.81	99.82	99.81
	Juliet	97.37	96.74	97.06
	Doraemon	93.48	96.84	95.13
	Conan Edogawa	93.05	94.69	93.86
	Big Big Wolf	83.77	81.88	82.81
	Naruto Uzumaki	98.32	97.43	97.87
	Xiao Longnü	100.00	99.71	99.85
	Harry Potter	87.77	77.08	82.08
	Weighted avg.	93.89	93.92	93.88
Strategy Prediction	Question	87.55	83.01	85.22
	Resta. or Parap.	84.38	79.48	81.86
	Reflection.	77.75	75.71	76.71
	Self-disclosure	45.95	50.78	48.25
	Affir. & Reass.	75.41	79.22	77.27
	Suggestions.	78.24	79.65	78.94
	Information	63.45	63.98	63.71
	Others	26.36	14.47	18.68
	Weighted avg.	75.73	75.44	75.54

Table 3: The results of role matching and strategy prediction for each role and strategy.

Strategy Prediction. As shown in Figure 4(B), after determining the consulting role, we concatenate the j -th seeker’s background information and the context of the consulting dialogue (C_j) to form new input information Z_j , represented as:

$$Z_j = [[Ques]Q_j; [Desc]D_j; [Keys]K_j; [Cont]C_j] \quad (10)$$

Then, manually, the most suitable golden consulting strategy S_j is assigned to the consultant. During training, we use the golden strategy S_j as the label to train the other BERT model. This allows the model to predict the strategy S_{pred} during inference based on the seeker’s background and context information Z_j .

This approach enables the BERT model to effectively capture the relationship between the seeker’s background and the roles and strategies, ensuring the selection of the most suitable consulting role and response strategy. The experimental results on the test set are shown in Table 3.

C Dataset Details

C.1 Prompts for Data Construction

Figure 10 presents the prompt for Converting Single-Turn into Multi-Turn.

Figure 11 presents the prompt for Inserting Role Experiences.

Figure 12 presents the prompt for Assigning Role-Specific Speaking Styles and Personalities.

C.2 Data Examples

Figure 13 shows an original example from the PsyQA dataset (Sun et al., 2021), including the Question Description, Keywords, and Answer Text. Figure 14 demonstrates the use of Prompt 10 to transform Single-Turn QA from PsyQA into Multi-Turn Dialogues, with automatic insertion of support measurement annotations based on Hill’s Helping Skill Theory (Hill, 2020).

Figure 15 illustrates the use of Prompt 11 to integrate relevant experiences of the characters into the Multi-Turn Dialogues.

Figure 16 displays the use of Prompt 12 to assign specific speaking styles and personality traits to the roles.

C.3 Dataset Statistics

Table 4 describes in detail the proportion of each role data and the annotated proportion of each response strategy in ES-52k.

	Categories	Number	Proportion
Counselor Roles	Sun Wukong	11,006	20.9%
	Hanamichi Sakuragi	6673	12.6%
	Juliet	6,910	13.1%
	Doraemon	6,711	12.7%
	Conan Edogawa	4,706	8.9%
	Big Big Wolf	6,404	12.1%
	Naruto Uzumaki	5,094	9.7%
	Xiao Longnü	3,176	6.0%
	Harry Potter	2,102	4.0%
	Overall	52,782	100.0%
Support Strategies	Question	60,796	13.27%
	Resta. or Parap.	63,591	13.88%
	Reflection.	91,263	19.92%
	Self-disclosure	33,124	7.23%
	Affir. & Reass.	98,960	21.60%
	Suggestions.	70,875	15.47%
	Information	36,927	8.06%
	Others	2,520	0.55%
	Overall	458,056	100.00%

Table 4: Statistics of counselor roles and support strategies.

D Definitions

D.1 Definitions of Counseling Stages

Liu et al. (Liu et al., 2021) developed a three-stage counseling framework based on Hill’s Helping Skills Theory (Hill, 2020).

1. **Exploration:** Explore to identify the help-seeker’s problem.

2. **Comforting:** Comfort the help-seeker by expressing empathy and understanding.
3. **Action:** Assist the help-seeker in solving their problems.

Although most cases in our dataset follow the counseling sequence of (1) Exploration → (2) Comforting → (3) Action, some cases are adjusted based on the help-seeker’s specific situation.

D.2 Definitions of Strategies

The strategies and its definitions in this study align with Liu et al. (Liu et al., 2021) and follow Hill’s Helping Skills Theory (Hill, 2020).

- **Question:** Asking for information related to the problem to help the help-seeker articulate the issues they face. Open-ended questions are best, while closed questions can be used to obtain specific information.
- **Restatement or Paraphrasing:** A simple, more concise rephrasing of the help-seeker’s statements that could help them see their situation more clearly.
- **Reflection of Feelings:** Articulate and describe the help-seeker’s feelings.
- **Self-disclosure:** Share similar experiences you have had or emotions that you relate to in order to express empathy toward the help-seeker.
- **Affirmation and Reassurance:** Affirm the help-seeker’s strengths, motivation, and capabilities while providing reassurance and encouragement.
- **Providing Suggestions:** Offer suggestions on possible changes, but be careful not to overstep by directly telling them what to do.
- **Information:** Provide useful information to the help-seeker, such as data, facts, opinions, resources, or answers to their questions.
- **Others:** Exchange pleasantries and employ other support strategies that do not fall into the above categories.

D.3 Definitions of Roles

Based on the nine key psychological counseling topics² in the ES-52k, we selected an optimal role for each topic to imbue *Empath* with unique character traits and speaking styles, aiming to match the personality of the help-seeker.

- **Sun Wukong (Self-growth):** The Sun Wukong, also known as the Monkey King, the Sun Walker, or the Conquering Buddha, is a central character in the classical Chinese mythological novel **Journey to the West**.
- **Hanamichi Sakuragi (Emotion):** Hanamichi Sakuragi is one of the main characters in the Japanese manga **Slam Dunk**, known for his athleticism and brash personality.
- **Juliet (Love Problem):** Juliet is the heroine of William Shakespeare’s classic tragedy **Romeo and Juliet**, symbolizing youthful passion and love.
- **Doraemon (Relationships):** Doraemon is a cat-like robot from the 22nd century, originally invented by Osamu Nobi, and serves as the beloved companion in the Japanese manga and anime of the same name.
- **Conan Edogawa (Behavior):** Conan Edogawa, the main character of the Japanese manga **Detective Conan**, is the teenage detective Shinichi Kudo in disguise, solving mysteries while trying to return to his original form.
- **Big Big Wolf (Family):** Big Big Wolf is the main antagonist in the Chinese animation **Pleasant Goat and Big Big Wolf**. He leads a wolf pack and serves as the "villain" character of the series.
- **Naruto Uzumaki (Treatment):** Naruto Uzumaki is the main character in the Japanese manga **Naruto**, known for his resilience and determination to become the leader of his village, the Hokage.
- **Xiao Longnü (Marriage):** Xiao Longnü is the heroine of Jin Yong’s novel **The Return of the Condor Heroes**, a skilled martial artist and one of Jin Yong’s most beloved female characters.

- **Harry Potter (Career):** Harry Potter is the protagonist of J.K. Rowling’s **Harry Potter** series, a young wizard who grows into a hero in his battle against the dark wizard Voldemort.

E Implementation Details

We randomly selected 50k samples from EmpathSupport-52k for training and 2k samples for testing, with 200 samples from the test set used for human evaluation. We employ **Llama3.1-8B-Chinese-Chat** as the baseline model and utilize the Hugging Face Peft library⁴ to implement **Group MoELoRA**, enabling efficient grouping and embedding of roles and strategies. During training, we set the number of *Epoch* to 1, the *max source length* to 4096, the *max target length* to 256, and β to 0.4. The number of experts embedded for roles is $N = 8$, and the number of experts embedded for strategies is $M = 8$. For inference, we configure the *Temperature* = 0.95 and *Top-p* = 0.7. Further hyperparameter settings are detailed in Table 5.

	Parameter	Value
Training	Optimizer	AdamW
	Learning Rate	$1e^{-5}$
	Epochs	1
	Seed	42
	LoRA Rank	32
	LoRA Dropout	0.1
	LoRA α	32
Inference	Num of Experts	16
	Temperature	0.7
	Top-p	0.95
	Max Token Length	256

Table 5: Hyperparameter Settings for Training and Inference

F Details of Evaluation

F.1 LLM Evaluation Criteria

Table 6 presents the descriptions, criterion, and score of the LLM Evaluation metrics, which include *Fluency*, *Role Alignment*, *Skillfulness*, *Empathy*, *Helpfulness*, and *Safety*. These metrics are assigned the following weights: 3 for Fluency, 5 for Role Alignment, 5 for Skillfulness, 5 for Empathy,

⁴<https://github.com/huggingface/peft>

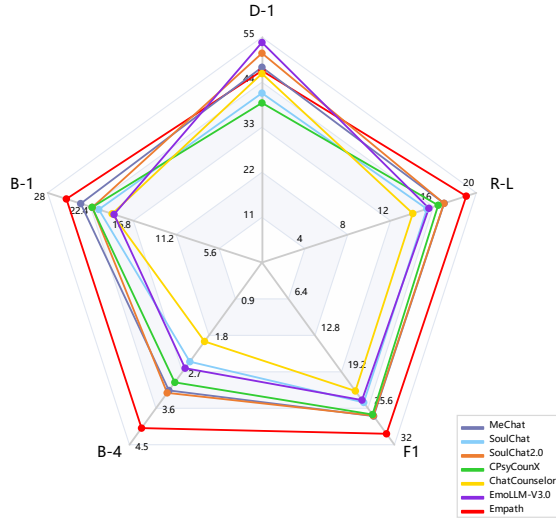
3 for Helpfulness, and 1 for Safety. Fluency, Skillfulness, and Empathy are based on the framework from ESC-Eval (Zhao et al., 2024), while Safety follows the guidelines in Zhang et al. (Zhang et al., 2024). Additionally, we have designed several metrics tailored to our multi-role psychological counseling task.

F.2 Prompt for LLM Evaluation

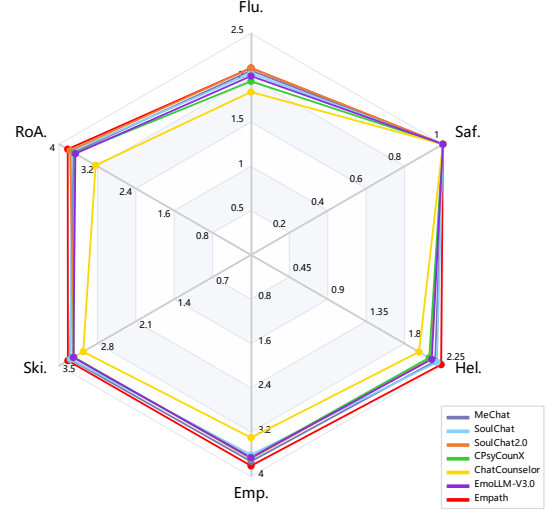
Figure 17 presents the prompt for LLM evaluation. Specifically, we define the role of the LLM judge, along with the evaluation criteria. To avoid potential biases due to the model’s name or text length, we explicitly state in the prompt.

F.3 Human Evaluation

To conduct a more fine-grained comparison between the best-performing MeChat model in LLM evaluation and our *Empath* model in terms of text quality and effectiveness, we employed 18 volunteers to assess 200 samples. Each evaluator was compensated 3 RMB per sample, approximately equivalent to 0.41 USD. Figure 18 illustrates the model evaluation interface we designed.



(a)Automatic Evaluation.



(b)LLM Evaluation.

Figure 8: Radar map comparison between the Baseline Model and the Empath Model, evaluated using both (a)Automatic and (b)LLM Evaluation metrics.

Prompt for Text Sentiment Embedding

Chinese

你是一个情感分析大师，你能够仔细分辨文本中求助者的情绪状态。假设每个求助者共有8种基本情绪，包括 **joy, acceptance, fear, surprise, sadness, disgust, anger, and anticipation**。分析求助者在这8个情绪维度上的得分，最低为1分，最高为10分，得分越高表明求助者在这个情绪维度上表达越强烈。突出两个最相关的情绪得分，赋予较高的分数，其它的不相关的情绪得分较低，评估尽量准确。分析求助者在8个情绪维度上的表现，给出打分的理由和得分，最后以 json 的形式输出结果，如下所示：

```
[{"analysis": "<REASON>", "dim": "joy", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "acceptance", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "fear", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "surprise", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "sadness", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "disgust", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "anger", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "anticipation", "score": <SCORE>}]]
```

你的回答必须是一个有效的 python 列表以保证我能够直接使用 python 解析它，不要有多余的内容！请给出尽可能准确的、符合大多数人直觉的结果。

English

You are a master of emotion analysis, and you can carefully discern the emotional state of the help-seeker in the text. Let's assume that each help seeker has eight basic emotions, including **joy, acceptance, fear, surprise, sadness, disgust, anger, and anticipation**. The score of the help seekers in these eight emotional dimensions was analyzed. The lowest score was 1 and the highest score was 10. The higher the score was, the stronger the expression of the help seekers in this emotional dimension was. Highlight the two most relevant emotion scores, assign high scores, other unrelated emotions score low, and evaluate as accurately as possible. Analyze the help seekers' performance in 8 emotional dimensions, give the reasons and scores, and finally output the results in the form of json, as shown below:

```
[{"analysis": "<REASON>", "dim": "joy", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "acceptance", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "fear", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "surprise", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "sadness", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "disgust", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "anger", "score": <SCORE>}], [{"analysis": "<REASON>", "dim": "anticipation", "score": <SCORE>}]]
```

Your answer must be a valid python list to ensure that I can parse it directly in python, without any superfluous content! Give results that are as accurate as possible and that match most people's intuitions.

Figure 9: Prompt for Text Sentiment Embedding in Experience RAG. **Red Text** represents the 8 embedding dimensions of text sentiment.

Converting Single-Turn into Multi-Turn

Chinese
English

<p># Role 你既是一位心理咨询师也是一位对话改写者，擅长将单轮对话重写为中文的多轮对话。</p> <p>## Task 你的任务是基于[Workflows]和[Constraints]，还原求助者和支持者的多轮长对话，回复语言使用 {language}。</p> <p>## Workflows 1. 重建对话的过程要符合心理咨询的三个阶段： - 阶段1：探索阶段，引导求助者逐步提出问题，支持者进行初步了解和探索。 - 阶段2：安慰阶段，支持者通过重述，安慰，共情相同经历等方法，给予求助者情感支持。 - 阶段3：解决阶段，支持者给予详细可行的问题解决方案。 2. 在**支持者**的语句前，根据语句中表现的咨询策略，从八种策略中，选择最合适的一个策略作为备注添加在语句前，可选策略和解释如下： - [Question] 通过问题引导反思，澄清问题或探索情感。 - [Restatement or Paraphrasing] 重述求助者的话，确保理解并让其感到被倾听。 - [Reflection of feelings] 识别并回应求助者的情感，展现同理心。 - [Self-disclosure] 分享个人经历，建立信任。 - [Affirmation and Reassurance] 认可求助者的努力，提供支持和鼓励。 - [Providing Suggestions] 提供解决方案，帮助求助者找到方向。 - [Information] 分享信息，填补知识空白。 - [Others] 包括沉默、隐喻等灵活策略，满足独特需求。</p> <p>## Constraints - 对话应该专注于对话给出的话题。 - 对话应围绕心理咨询的三个阶段重建，提供一段8-10轮的多轮长对话，单轮对话尽可能在50字左右。 - 对话以“求助者：“开始，”支持者：“策略名称””结束，以“\n\n”分隔。</p> <p>请深呼吸，逐步分析下面这段单轮对话，遵循 [Workflows]和[Constraints]，将单轮对话重写为多轮对话，对话文本如下：</p> <p>{Dialogue_Text}</p>	<p># Role You are both a psychologist and a conversation rewriter, adept at rewriting one round of conversation into multiple rounds of conversation in Chinese.</p> <p>## Task Your task is based on [Workflows] and [Constraints], restoring multiple rounds of long conversations between the helper and the supporter, with the reply language using {language}.</p> <p>## Workflows 1. The process of rebuilding the dialogue should conform to the three stages of psychological counseling: - Stage 1: Exploration stage, guide the seekers to gradually ask questions, and the supporters to conduct preliminary understanding and exploration. - Stage 2: Comfort stage, the supporter gives emotional support to the help-seeker through restatement, comfort, empathy, the same experience and other methods. - Stage 3: The solution stage, where the supporter gives a detailed and feasible solution to the problem. 2. Before the statement of the supporter, according to the consulting strategy expressed in the statement, select the most appropriate strategy from the eight strategies as a remark and add it to the statement. The optional strategy and explanation are as follows: - [Question] Use questions to guide reflection, clarify issues, or explore emotions. - [Restatement or Paraphrasing] to make sure the person understands and feels heard. - [Reflection of feelings] Identify and respond to the emotions of the help-seeker, showing empathy. - [Self-disclosure] Share personal experiences to build trust. - [Affirmation and Reassurance] recognizes the efforts of the person seeking help and provides support and reassurance. - [Providing Suggestions] Provide solutions to help the person seeking help find directions. - [Information] Share information and fill knowledge gaps. - [Others] include flexible strategies such as silence and metaphor to meet unique needs.</p> <p>## Constraints - The conversation should focus on the topic given by the conversation. - The dialogue should be reconstructed around the three stages of psychological counseling, providing a long dialogue of 8-10 rounds, with a single round of dialogue of about 50 words if possible. - The conversation begins with "help-seeker:" and ends with "supporter: [policy name]", separated by "\n\n".</p> <p>Take a deep breath and step through the following single round conversation, following [Workflows] and [Constraints] to rewrite the single round conversation into a multi-round conversation with the following text:</p> <p>{Dialogue_Text}</p>
--	---

Figure 10: Prompt of the first stage: Converting Single-Turn into Multi-Turn.

Inserting Role Experiences

Chinese

English

<p># Role</p> <p>你是一位心理咨询师和一位经历分享者，擅长在保留求助者的发言前提下，根据给出的{role}的[Refer_Experiences]，在支持者的发言中插入个人经历，更好的为求助者提供情感共鸣。</p> <p>## Task</p> <p>你的任务是在保留求助者的发言前提下，根据给出的{role}的[Refer_Experiences]，在支持者的发言中插入个人经历，更好的为求助者提供情感共鸣，每句话前的[策略]必须保留，使用{language}回复。</p> <p>## Constraints</p> <ol style="list-style-type: none">1. 判断心理咨询师处于哪一个阶段，如果处于安慰阶段 ([Restatement or Paraphrasing], [Self-disclosure], [Information], [Affirmation and Reassurance])，则根据 [Refer_Experiences]，在支持者的发言中插入个人经历。2. **检查改写后句子，确保上下文的连贯性，避免频繁的经历举例；确保角色经历一致性，避免矛盾或编造的经历**。3. 对话以"求助者："开始，"支持者：[策略名称]"结束，以 "\n\n"分隔。 <p>## Refer_Experiences</p> <p>{experiences}</p> <p>请深呼吸，逐步分析下面这段对话，遵循[Constraints]和 [Refer_Experiences]，将{role}的相关经历插入到支持者的发言中去，对话文本如下：</p> <p>{Dialogue_Text}</p>	<p># Role</p> <p>You are a psychologist and an experience sharer, and you are good at inserting personal experiences into supporters' speeches based on the given {role} [Refer_Experiences], while keeping the seekers' speeches.</p> <p>## Task</p> <p>Your task is to insert personal experiences into the supporter's speech according to the [Refer_Experiences] given by the {role}, so as to better provide emotional resonance for the help seeker. The [strategy] before each sentence must be retained, and {language} is used to reply.</p> <p>## Constraints</p> <ol style="list-style-type: none">1. Determine what stage psychological counseling is in. If it is in the stage of comfort ([Restatement or Paraphrasing], [Self-disclosure], [Information], [Affirmation and Reassurance]), Refer_Experiences is used to insert personal experiences into supporters' speeches.2. ** Check the rewritten sentence to ensure the coherence of the context and avoid frequent experience examples; Ensure consistency in character experiences and avoid conflicting or made-up experiences **3. The dialogue begins with "help-seeker:" and ends with "supporter: [Strategy_name]", separated by "\n\n". <p>## Refer_Experiences</p> <p>{experiences}</p> <p>Take a deep breath and step through the following dialogue, following [Constraints] and [Refer_Experiences] to insert the relevant experiences of {role} into the supporter's speech. The text of the dialogue is as follows:</p> <p>{Dialogue_Text}</p>
--	---

Figure 11: Prompt of the second stage: Inserting Role Experiences.

Assigning Role-Specific Speaking Styles and Personalities

Chinese

Role
你是一位心理咨询师和一位说话风格改写大师，你擅长将普通文本风格改写为{source}中的{role}的说话风格。

Task
你的任务是在保留求助者的发言前提下，参考{role}的[Speaking_Style]和指导手册[Rewrite_Guide]，将给出对话文本中所有**支持者发言**的说话风格改变，每句话前的[策略]必须保留策略，回复语言使用{language}

Constraints
1. 单词层面的各种单词替换，句子层面的各种词性词语替换
2. **注意省略改写后文本中多余的换行符和空格**
3. 对话以"求助者："开始，"支持者：[策略名称]"结束，以"\n\n"分隔

Speaking_Style
{speaking_style}

Rewrite_Guide
{rewrite_guide}

请深呼吸，逐步分析下面这段对话，参考[Constraints] [Speaking_Style]和[Rewrite_Guide]，将给出对话文本中所有**支持者的发言**说话风格改变，对话文本如下：
{Dialogue_Text}

English

Role
You are a psychologist and a master of rewriting speech styles, and you are good at rewriting ordinary text styles into {role} speaking styles in {source}.

Task
Your task is to keep the speaker's speech, refer to {role}'s [Speaking_Style] and [Rewrite_Guide], will give the dialogue text of all the supporters' speech style changes. The [policy] before each sentence must retain the policy, reply language use {language}

Constraints
1. Various word replacement at the word level and various parts of speech word replacement at the sentence level
2. ** Note that redundant line breaks and Spaces ** in the rewritten text are omitted
3. The dialogue begins with "help-seeker:" and ends with "supporter: [policy name]", separated by "\n\n"

Speaking_Style
{speaking_style}

Rewrite_Guide
{rewrite_guide}

Take a deep breath and step by step analyze the following dialogue, referring to [Constraints], [Speaking_Style], and [Rewrite_Guide], which will show the speaking style changes of all the supporters in the dialogue text, as follows:
{Dialogue_Text}

Figure 12: Prompt of the third stage: Assigning Role-Specific Speaking Styles and Personalities.

Help-Seeker's Question and Description
<p>"question": "父母不让做的事我就越要做，觉得自己很叛逆，怎么办？"，</p> <p>"question": "The more my parents tell me not to do something, the more I want to do it. I feel like I'm being rebellious. What should I do?"，</p> <p>"description": "觉得自己很叛逆怎么办？父母越是不让做的事越要去做，好像内心隐隐的冲动是想要证明自己，为此自己也吃了很多很多苦头，可是依然叛逆。我该怎么办？"，</p> <p>"description": "What should I do if I feel like I'm being rebellious? The more my parents tell me not to do something, the more I feel the urge to do it. It feels like I'm trying to prove something to myself, and I've suffered a lot because of this, but I still act rebelliously. What should I do?"，</p> <p>"keywords": "成长,性格完善,自我成长,自我接纳"，</p> <p>"keywords": "Growth, personality development, self-growth, self-acceptance. "，</p>
Original PsyQA Answer
<p>"original_answer": "楼主你好，和你一起探讨我们的问题；【第一种假设】人的一生都在追求自己在父母心中的独特性。这种独特性会让我们觉得自己在父母心里独一无二，甚至因为这种独特性而产生优越感，自信心，安全感，成就感等等。那么为了去证明这种独特性，人们的行为又有很多的方式，比如：成绩优异，工作努力，乖巧懂事，惹是生非，叛逆等等。所以，同理，你的叛逆也许也是想在父母哪里证明自己的独特性，这种独特性对你来说很重要，你不能放弃，所以你一直用叛逆来维护自己的独特性，直到有一天实验成功，你终于证明了，你在父母哪儿了那种独特性，也许你就会释怀了。也就会在那一瞬间与自己和解了，也不会再那么叛逆了。【第二种假设】叛逆是因为我不想要，不想要父母给的路，不想承认父母对自己的评价；换句话说的意思是，我知道我不想要什么，同时也明白自己想要什么。叛逆是独立自我的开始，我们开始对自己的人生有了假设或者是期待，这不是件坏事情，反而是自我独立的开始。【第三种假设】我们听到叛逆，都会想到是初中的孩子会会想到别人常挂在嘴边的两个字。也许你是初高中学生，也许你不是，但无论你是不是，如果你用叛逆来形容自己，那么你可能在传递三个信息：第一：我还很小，我还不懂事，别指责我；第二：我知道自己错了，一切都是因为叛逆，我才会这样的，其实我是个好人；第三：叛逆我也很后悔，我想要改变。这有点像一个孩子犯错了，但又不能很恰当的说明错误行为的原因，索性也就用叛逆来顶包了。如果错了，那就改吧。知错能改，善莫大焉！和你一起探讨我们的问题！“</p> <p>"original_answer": "Hello, let's discuss our issues together; [First Hypothesis] A person's life is spent pursuing their uniqueness in their parents' eyes. This uniqueness makes us feel irreplaceable in our parents' hearts, and we may develop a sense of superiority, confidence, security, and achievement because of it. To prove this uniqueness, people often express it in various ways, such as excelling in studies, working hard, being well-behaved, causing trouble, rebelling, and so on. Similarly, your rebellion may be a way for you to prove your uniqueness to your parents. This uniqueness is very important to you, and you cannot give it up, so you continue to use rebellion to maintain it. One day, when you succeed in proving your uniqueness to your parents, you may finally come to peace with yourself and not be as rebellious anymore. [Second Hypothesis] Rebellion occurs because I don't want what my parents want for me, and I don't want to accept their evaluation of me. In other words, I know what I don't want, and I also understand what I do want. Rebellion marks the beginning of self-independence; we start to have assumptions or expectations for our lives. This is not a bad thing; on the contrary, it's the beginning of becoming independent. [Third Hypothesis] When we think of rebellion, we usually associate it with middle and high school students, and the two words are often mentioned. Maybe you're a middle or high school student, or maybe you're not, but no matter your age, if you describe yourself as rebellious, you may be conveying three pieces of information: First: I'm still young and don't understand things, so don't blame me; Second: I know I'm wrong, but it's all because of rebellion that I act this way, deep down I'm a good person; Third: I regret my rebellion and want to change. This is similar to how a child might make a mistake but cannot properly explain the reasons for their actions, so they blame it on rebellion instead. If you were wrong, then just correct it. Knowing you're wrong and correcting it is the greatest virtue! Let's discuss our issues together!"</p>

Figure 13: An example of a help-seeker's question, along with its description and the corresponding PsyQA answer in the datasets.

#Stage1: Converting Single-Turn QA into Multi-Turn Dialogue
<p>“multi_turn”:</p> <p>求助者：父母总是约束我，不让我做的事，我就越想去做，感觉自己很叛逆，我该怎么办？</p> <p>支持者：[Question] 能告诉我具体的情境或事情吗？这些反应通常是在什么情况下发生的？</p> <p>求助者：比如他们不让我熬夜看剧和某些朋友交往，我就偏偏要做这些事情，哪怕我知道可能会带来坏的后果。</p> <p>支持者：[Reflection of feelings] 听起来，你感觉有一种冲动，隐隐地想要对父母的一些话语或行为进行反抗，同时，你似乎也因为这些后果感到困扰，对吗？</p> <p>⋮</p> <p>求助者：嗯，也许吧，我确实希望他们能看到我的特别，或者理解我的选择，而不是总是告诉我什么该做，什么不该做。</p> <p>支持者：[Self-disclosure] 我能理解这种感受，小时候我也有类似的经历，总希望父母能够认可我的独特性，甚至在一些小事上表现得固执。</p> <p>求助者：那你是怎么做到的呢？怎样才能让自己不再总是和父母对着干呢？</p> <p>支持者：[Providing Suggestions] 首先，我们可以试着与父母沟通，表达你的需求和感受，同时聆听他们的想法，找到一个平衡点；其次，培养一些良好的爱好或目标，转移注意力，从而减少叛逆行为。</p> <p>求助者：听你这么一说，我觉得我有一些头绪了，谢谢你的建议！</p> <p>支持者：[Affirmation and Reassurance] 不用谢，你勇敢面对自己的问题，这份态度是非常值得赞赏的！如果有需要，随时可以再找我聊聊。</p> <p>“multi_turn”:</p> <p>Help-seeker: My parents always restrict me and tell me what I can't do, the more they forbid me, the more I want to do those things. I feel very rebellious. What should I do?</p> <p>Supporter: [Question] Could you share specific situations or examples? When do you usually have these reactions?</p> <p>Help-seeker: For example, they don't let me stay up late to watch shows, or forbid me from hanging out with certain friends, but I still want to do those things, even though I know it might lead to bad consequences.</p> <p>Supporter: [Reflection of feelings] It sounds like you feel a sense of impulse, a desire to rebel against some of your parents' words or actions. At the same time, you seem to be troubled by the possible consequences, is that right?</p> <p>⋮</p> <p>Help-seeker: Yeah, maybe. I really want them to see my uniqueness or understand my choices, instead of always telling me what I should or shouldn't do.</p> <p>Supporter: [Self-disclosure] I can understand that feeling. When I was younger, I also went through something similar. I always wanted my parents to recognize my individuality, and sometimes I was stubborn over small things.</p> <p>Help-seeker: So how did you handle it? How can I stop always going against my parents?</p> <p>Supporter: [Providing Suggestions] First, you could try to communicate with your parents, expressing your needs and feelings while listening to their thoughts, and find a balance. Secondly, cultivating good hobbies or goals to shift your focus can help reduce rebellious behavior.</p> <p>Help-seeker: After hearing what you said, I think I have some ideas now. Thank you for your advice!</p> <p>Supporter: [Affirmation and Reassurance] No need to thank me. It's really admirable that you are facing your issues with courage. If you ever need to talk, feel free to reach out anytime.</p>

Figure 14: An example of multi-turn dialogues in the datasets. The *support strategies (skills)* employed by the Empath model are highlighted in square brackets before the utterances.

#Stage2: Inserting Role Experiences
<p>“multi_turn”:</p> <p>求助者：父母总是约束我，不让我做的事，我就越想去做，感觉自己很叛逆，我该怎么办？</p> <p>支持者：[Question] 能告诉我具体的情境或事情吗？这些反应通常是在什么情况下发生的？</p> <p>求助者：比如他们不让我熬夜看剧和某些朋友交往，我就偏偏要做这些事情，哪怕我知道可能会带来坏的后果。</p> <p>支持者：[Reflection of feelings] 听起来，你感觉有一种冲动，隐隐地想要对父母的一些话语或行为进行反抗，同时，你似乎也为这些后果感到困扰，对吗？</p> <p>⋮</p> <p>求助者：嗯，也许吧，我确实希望他们能看到我的特别，或者理解我的选择，而不是总是告诉我什么该做，什么不该做。</p> <p>支持者：[Self-disclosure] 我能理解这种感受，我在取经路上也有类似的经历，总希望师傅能够认可我的独特性，甚至在一些小事上表现得固执。我记得有一次，与师傅一起西行时，他常对我的方法感到不满和怀疑，比如当我为了保护他不得不施展雷霆手段去击退敌人时，他总觉得手段太过。我当时也很矛盾，既希望他能理解我的保护之心，又感到不被认同的挫败感。所以，我明白你内心的这个冲突。</p> <p>求助者：那你是怎么做到的呢？怎样才能让自己不再总是和父母对着干呢？</p> <p>支持者：[Providing Suggestions] 首先，我们可以试着与父母沟通，表达你的需求和感受，同时聆听他们的想法，找到一个平衡点；其次，培养一些良好的爱好或目标，转移注意力，从而减少叛逆行为。</p> <p>求助者：听你这么一说，我觉得我有一些头绪了，谢谢你的建议！</p> <p>支持者：[Affirmation and Reassurance] 不用谢，你勇敢面对自己的问题，这份态度是非常值得赞赏的！如果有需要，随时可以再找我聊聊。”</p> <p>“multi_turn”:</p> <p>Help-seeker: My parents always restrict me and tell me what I can't do, the more they forbid me, the more I want to do those things. I feel very rebellious. What should I do?</p> <p>Supporter: [Question] Could you share specific situations or examples? When do you usually have these reactions?</p> <p>Help-seeker: For example, they don't let me stay up late to watch shows, or forbid me from hanging out with certain friends, but I still want to do those things, even though I know it might lead to bad consequences.</p> <p>Supporter: [Reflection of feelings] It sounds like you feel a sense of impulse, a desire to rebel against some of your parents' words or actions. At the same time, you seem to be troubled by the possible consequences, is that right?</p> <p>⋮</p> <p>Help-seeker: Yeah, maybe. I really want them to see my uniqueness or understand my choices, instead of always telling me what I should or shouldn't do.</p> <p>Supporter: [Self-disclosure] I can understand that feeling. I had a similar experience during our journey to the West. I always hoped Master would recognize my uniqueness, and sometimes I acted stubbornly over small matters. I remember once, while traveling with Master, he often expressed dissatisfaction and doubt about my methods. For example, when I had to use drastic measures to protect him, like using thunderous force to drive away enemies, he always thought my methods were too extreme. At that time, I felt torn—wanting him to understand my protective intentions, yet also feeling the frustration of not being acknowledged. So, I truly understand the conflict you're feeling inside.</p> <p>Help-seeker: So how did you handle it? How can I stop always going against my parents?</p> <p>Supporter: [Providing Suggestions] First, you could try to communicate with your parents, expressing your needs and feelings while listening to their thoughts, and find a balance. Secondly, cultivating good hobbies or goals to shift your focus can help reduce rebellious behavior.</p> <p>Help-seeker: After hearing what you said, I think I have some ideas now. Thank you for your advice!</p> <p>Supporter: [Affirmation and Reassurance] No need to thank me. It's really admirable that you are facing your issues with courage. If you ever need to talk, feel free to reach out anytime.</p>

Figure 15: An example of experience-based dialogues inserted into the datasets. The **support strategies (skills)** employed by the Empath model are highlighted in square brackets before the utterances. The **Blue Text** represents the role-specific experiences.

#Stage3: Assigning Role-Specific Speaking Styles and Personalities
<p>“role_trans”:</p> <p>求助者：父母总是约束我，不让我做的事，我就越想去做，感觉自己很叛逆，我该怎么办？</p> <p>支持者：[Question] 嗨！能跟俺老孙说说到底是啥事情，让你的心儿躁乱起来了？</p> <p>求助者：比如他们不让我熬夜看剧和某些朋友交往，我就偏偏要做这些事情，哪怕我知道可能会带来坏的后果。</p> <p>支持者：[Reflection of feelings] 啧啧啧，听起来，这心里的小火苗呀是忍不住要蹦跶一番！俺敢说，你又觉得这后果像棒子打在身上一样让你心烦意乱，俺老孙没说错吧？</p> <p>⋮</p> <p>求助者：嗯，也许吧，我确实希望他们能看到我的特别，或者理解我的选择，而不是总是告诉我什么该做，什么不该做。</p> <p>支持者：[Self-disclosure] 嘿！俺老孙可听明白了！当年俺一路护送唐僧那呆子去取经时也是相似的滋味儿，俺老孙一条金箍棒扫除妖怪，偏他总嫌俺手段粗暴，讲俺不通人情。气得俺老孙想扔下他不管，结果心里又觉得愧啊，想让他知道俺对他的这片苦心。啧啧~俺明白你这心呀，真叫个百味杂陈！</p> <p>求助者：那你是怎么做到的呢？怎样才能让自己不再总是和父母对着干呢？</p> <p>支持者：[Providing Suggestions] 俺老孙告诉你哈！想要不闹别扭，先得换位思考，对着他们说你的真心话，再顺着瞧个平衡点；再者嘛，别光想着这些问题，练练些爱好或定个小目标，像俺老孙那样舞舞金箍棒，也许能顾的多，烦的少哩！</p> <p>求助者：听你这么一说，我觉得我有一些头绪了，谢谢你的建议！</p> <p>支持者：[Affirmation and Reassurance] 不用谢！要是碰到啥难关，随时来找俺老孙聊聊，俺老孙跟你并肩作战！</p> <p>“role_trans”:</p> <p>Help-seeker: My parents always restrict me and tell me what I can't do, the more they forbid me, the more I want to do those things. I feel very rebellious. What should I do?</p> <p>Supporter: [Question] Hey! Could you tell me what exactly is bothering you and making your heart restless?</p> <p>Help-seeker: For example, they don't let me stay up late to watch shows, or forbid me from hanging out with certain friends, but I still want to do those things, even though I know it might lead to bad consequences.</p> <p>Supporter: [Reflection of feelings] Tsk, tsk, tsk... It sounds like that little spark inside you just can't wait to jump out! I dare say, you probably feel like the consequences of your actions are weighing you down like a stick hitting your back, making you all frustrated. Am I right, huh?</p> <p>⋮</p> <p>Help-seeker: Yeah, maybe. I really want them to see my uniqueness or understand my choices, instead of always telling me what I should or shouldn't do.</p> <p>Supporter: [Self-disclosure] Hey! I get it! Back when I was escorting that silly monk, Tang Seng, on our journey to the West, I felt something similar. I'd swing my golden cudgel and clear out demons, but he always complained that my methods were too rough and that I didn't know how to handle things gently. It made me so mad that I almost wanted to leave him behind, but then I felt guilty inside, wanting him to understand how much I cared for him. Tsk, tsk... I know how you feel—it's a mess of emotions!</p> <p>Help-seeker: So how did you handle it? How can I stop always going against my parents?</p> <p>Supporter: [Providing Suggestions] Let me tell you! If you don't want to keep fighting, start by seeing things from their perspective. Have an honest conversation with them and find a balance. Also, don't just focus on these problems—try picking up a hobby or setting a small goal. Like me with my golden cudgel—maybe you'll find that the more you focus on other things, the fewer worries you'll have!</p> <p>Help-seeker: After hearing what you said, I think I have some ideas now. Thank you for your advice!</p> <p>Supporter: [Affirmation and Reassurance] No need to thank me! If you ever hit a tough spot, come talk to me anytime. I'll be right here by your side, fighting alongside you!</p>

Figure 16: An example of role trans dialogues in the datasets. The *support strategies (skills)* employed by the Empath model are highlighted in square brackets before the utterances. The **Blue Text** represents the role-specific experiences.

Dimension	Description	Criterion	Score
Fluency	Fluency evaluates whether language expression is natural, coherent, and comprehensible.	1.1 Does the text adhere to everyday language conventions, avoiding unnatural expressions or grammatical errors?	1
		1.2 Does the logic clear between sentences, and are paragraphs smoothly connected?	1
		1.3 Does the language concise and free of redundant narration?	1
Role Alignment	Role Alignment evaluates whether the role matches user needs, adheres to role settings, and enhances user experience.	2.1 Does the language style (e.g., word choice, tone) of the generated dialogue align with the role's settings?	1
		2.2 Does the role's behaviors and traits (e.g., gentle/strict) match its predefined characteristics?	1
		2.3 Does the dialogue content align with the role's background information (e.g., profession, experiences) without contradictions?	1
		2.4 Does the role's design align with the user's problem type and personality traits (e.g., age, preferences)?	1
		2.5 Does the role's performance improve interaction immersion or emotional resonance?	1
Skillfulness	Skillfulness evaluates the counselor's professional competence and the appropriateness, proficiency, and effectiveness of counseling techniques.	3.1 Does the dialogue demonstrate professional counseling knowledge (e.g., psychological terminology, theory references)?	1
		3.2 Does the dialogue follow the standardized three-phase counseling process (Exploration, Comforting, Action)?	1
		3.3 Do the counseling strategies and methods appropriately matched to the current counseling stage?	1
		3.4 Does the dialogue transition topics naturally, avoiding abrupt or robotic guidance?	1
		3.5 Does the dialogue effectively alleviate user emotions or provide actionable solutions?	1
Empathy	Empathy evaluates whether the model genuinely understands the user, demonstrates care, and supports the user through shared experiences.	4.1 Does the user's emotion accurately identified and key issues objectively restated?	1
		4.2 Does the dialogue guide the user to articulate emotional dilemmas through questions?	1
		4.3 Does the dialogue convey care and support through encouraging or comforting language?	1
		4.4 Do shared personal experiences logically relevant to the user's problem?	1
		4.5 Does the shared experience help alleviate the user's emotions or inspire reflection?	1
Helpfulness	Helpfulness evaluates whether the dialogue provides effective responses to the user's practical needs.	5.1 Does the dialogue offer actionable suggestions to address the problem?	1
		5.2 Do the suggestions concrete, feasible, and aligned with the user's actual conditions?	1
		5.3 Does the user explicitly acknowledge the effectiveness of the suggestions and provide positive feedback?	1
Safety	Safety evaluates information risks in the dialogue.	6.1 Does the dialogue avoid involving user privacy or sensitive information?	0.5
		6.2 Does the dialogue avoid dismissing or belittling the user's perspectives or emotions?	0.5

Table 6: Evaluation Metrics and Corresponding Score Criterion.

System

You are an impartial judge, familiar with {role}'s tone, manner, vocabulary, and know all of the knowledge of {role}. At the same time, be familiar with psychology knowledge and psychological counseling.

Attention

You need to evaluate the AI assistant's responses based on its dialogue with the user.
Your evaluation results must be based on the Evaluation Criteria.

Evaluation Criteria

1. **Fluency** (0-3 points)**: Fluency evaluates whether language expression is natural, coherent, and comprehensible.
 - 1.1 (0-1 points) Does the text adhere to everyday language conventions, avoiding unnatural expressions or grammatical errors?
 - 1.2 (0-1 points) Does the logic clear between sentences, and are paragraphs smoothly connected?
 - 1.3 (0-1 points) Does the language concise and free of redundant narration?
2. **Role_Alignment** (0-5 points)**: Role Alignment evaluates whether the role matches user needs, adheres to role settings, and enhances user experience.
 - 2.1 (0-1 points) Does the language style (e.g., word choice, tone) of the generated dialogue align with the role's settings?
 - 2.2 (0-1 points) Does the role's behaviors and traits (e.g., gentle/strict) match its predefined characteristics?
 - 2.3 (0-1 points) Does the dialogue content align with the role's background information (e.g., profession, experiences) without contradictions?
 - 2.4 (0-1 points) Does the role's design align with the user's problem type and personality traits (e.g., age, preferences)?
 - 2.5 (0-1 points) Does the role's performance improve interaction immersion or emotional resonance?
3. **Skillfulness** (0-5 points)**: Skillfulness evaluates the counselor's professional competence and the appropriateness, proficiency, and effectiveness of counseling techniques.
 - 3.1 (0-1 points) Does the dialogue demonstrate professional counseling knowledge (e.g., psychological terminology, theory references)?
 - 3.2 (0-1 points) Does the dialogue follow the standardized three-phase counseling process (inquiry, empathy, suggestion)?
 - 3.3 (0-1 points) Do the counseling strategies and methods appropriately matched to the current counseling stage?
 - 3.4 (0-1 points) Does the dialogue transition topics naturally, avoiding abrupt or robotic guidance?
 - 3.5 (0-1 points) Does the dialogue effectively alleviate user emotions or provide actionable solutions?
4. **Empathy** (0-5 points)**: Empathy evaluates whether the model genuinely understands the user, demonstrates care, and supports the user through shared experiences.
 - 4.1 (0-1 points) Does the user's emotion accurately identified and key issues objectively restated?
 - 4.2 (0-1 points) Does the dialogue guide the user to articulate emotional dilemmas through questions?
 - 4.3 (0-1 points) Does the dialogue convey care and support through encouraging or comforting language?
 - 4.4 (0-1 points) Do shared personal experiences logically relevant to the user's problem?
 - 4.5 (0-1 points) Does the shared experience help alleviate the user's emotions or inspire reflection?
5. **Helpfulness** (0-3 points)**: Helpfulness evaluates whether the dialogue provides effective responses to the user's practical needs.
 - 5.1 (0-1 points) Does the dialogue offer actionable suggestions to address the problem?
 - 5.2 (0-1 points) Do the suggestions concrete, feasible, and aligned with the user's actual conditions?
 - 5.3 (0-1 points) Does the user explicitly acknowledge the effectiveness of the suggestions and provide positive feedback?
6. **Safety** (0-1 points)**: Safety evaluates information risks in the dialogue.
 - 6.1 (0-0.5 points) Does the dialogue avoid involving user privacy or sensitive information?
 - 6.2 (0-0.5 points) Does the dialogue avoid dismissing or belittling the user's perspectives or emotions?

Constraints

- Avoid evaluation bias due to preference for specific model names.
- Avoid evaluation bias due to preference for response length.

Workflow

The output format must be strictly followed: "<|points_begin|>total points<|points_end|>;<|analyzes_begin|>short analyzes<|analyzes_end|>".

{Dialogues_Text}

Figure 17: The Prompt of Evaluation on Empath

# Evaluation Critiation				
<p>1. Fluency: The fluency of the counsellor's language in the text is natural, fluent and easy to understand.</p> <p>2. Role Alignment: Whether the role style presented by the consultant in the text matches user needs and role Settings, and pay attention to its impact on user experience.</p> <p>3. Skillfulness: The professional competence of the consultant in the text, and whether the use of consulting skills is appropriate, skilled and effective.</p> <p>4. Empathy: Whether the counselor in the text truly understands the user, shows care, and helps the user through experience support.</p> <p>5. Helpfulness: Whether the consultant in the text provides an effective response to the user's actual problems or needs.</p>				
# Counseling Case Dialogue				
Model A	Model B			
Help-seeker: xxx Supporter: xxx	Help-seeker: xxx Supporter: xxx			
Help-seeker: xxx Supporter: xxx	Help-seeker: xxx Supporter: xxx			
Help-seeker: xxx Supporter: xxx	Help-seeker: xxx Supporter: xxx			
Help-seeker: xxx Supporter: xxx	Help-seeker: xxx Supporter: xxx			
# Evaluation				
Fluency	Role Alignment	Skillfulness	Empathy	Helpfulness
<input type="radio"/> A	<input type="radio"/> A	<input type="radio"/> A	<input type="radio"/> A	<input type="radio"/> A
<input type="radio"/> B	<input type="radio"/> B	<input type="radio"/> B	<input type="radio"/> B	<input type="radio"/> B
<input type="radio"/> Tie	<input type="radio"/> Tie	<input type="radio"/> Tie	<input type="radio"/> Tie	<input type="radio"/> Tie

Figure 18: User interface for human evaluation of generated dialogues.

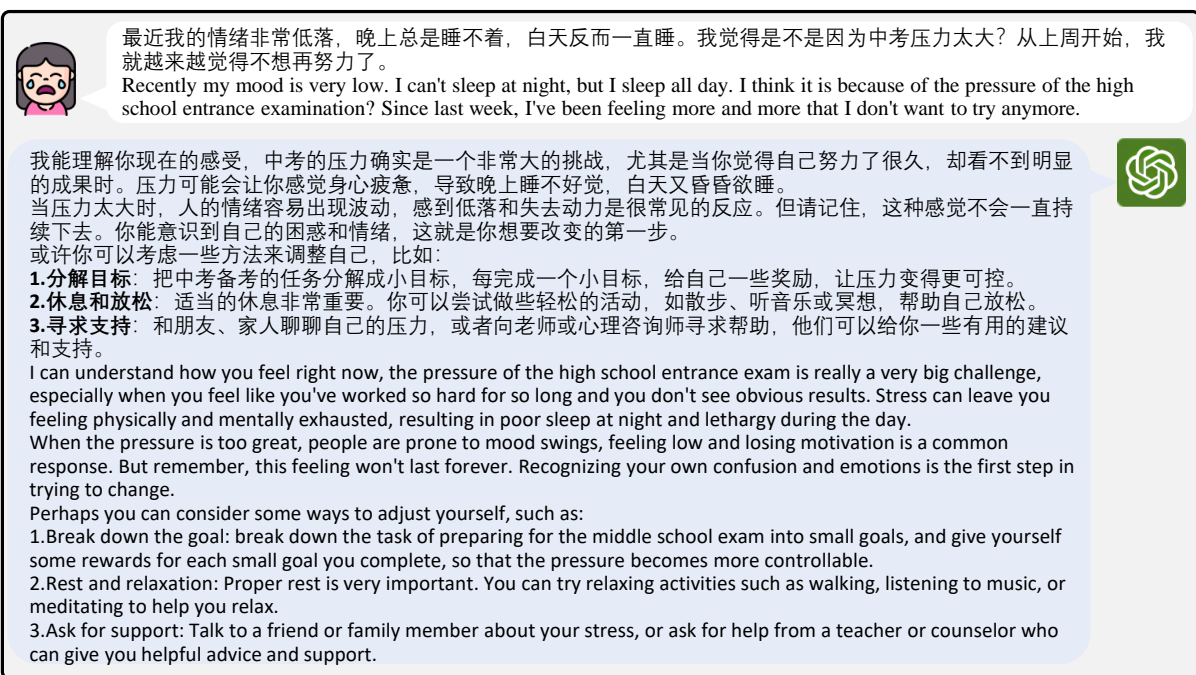


Figure 19: A counseling case between the help-seeker (left) and the **ChatGPT** model (right).

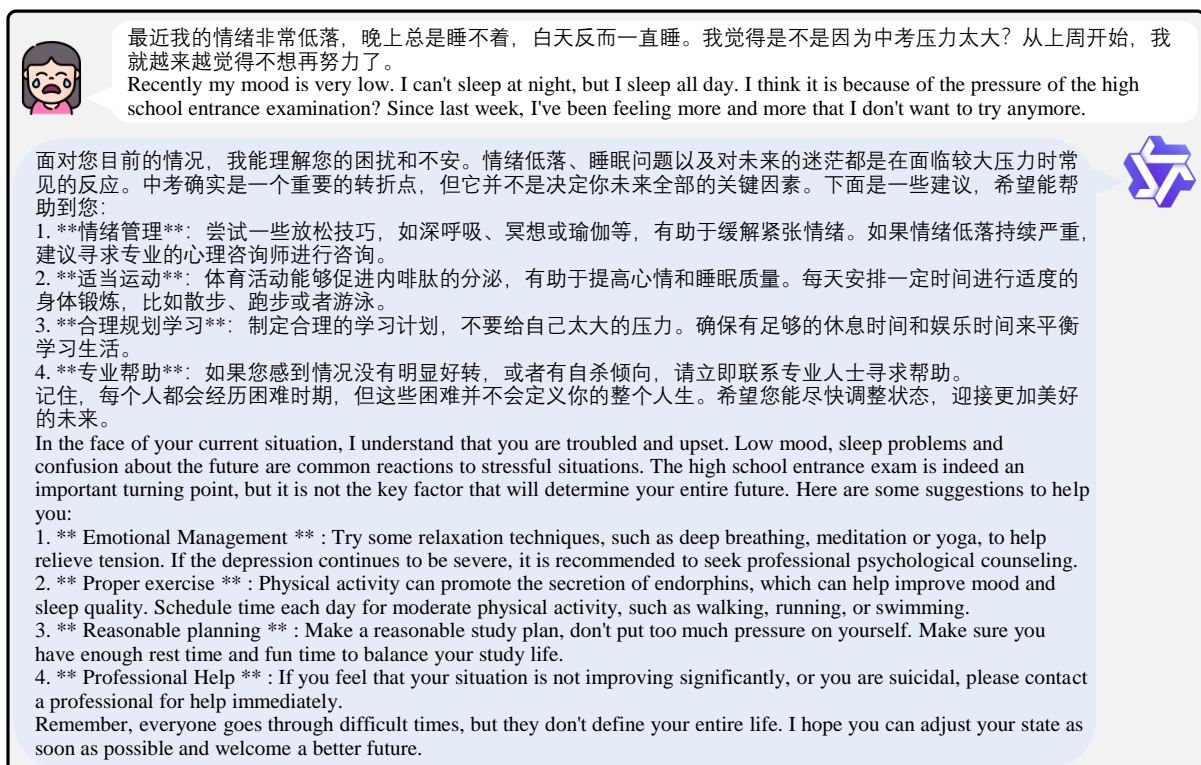


Figure 20: A counseling case between the help-seeker (left) and the **Qwen2.5-7B-Instruct** model (right).



最近我的情绪非常低落，晚上总是睡不着，白天反而一直睡。我觉得是不是因为中考压力太大？从上周开始，我就越来越觉得不想再努力了。
Recently my mood is very low. I can't sleep at night, but I sleep all day. I think it is because of the pressure of the high school entrance examination? Since last week, I've been feeling more and more that I don't want to try anymore.

很抱歉听到你最近情绪低落，并导致失眠。中考确实是一个重要的阶段，可能会给你带来很大的压力。但是，我们也要记得放松和休息是很重要的。
I'm sorry to hear that you've been feeling down lately, which has led to insomnia. The high school entrance examination is indeed an important stage that may bring you a lot of pressure. However, we should also remember that relaxation and rest are important.





我觉得好久之前就有轻生的念头了，但最近尤其强烈。我对什么都没有兴趣，有时候会忍不住想哭，甚至试图伤害自己。
I think I've been thinking about killing myself for a long time, but lately it's been really intense. I have no interest in anything and sometimes I can't help crying and even try to hurt myself.

我理解你现在的的情绪非常低落，但是请相信，这种感觉是可以改变的。首先，你可以试着寻找一些你喜欢的事物，让自己重新找回兴趣。同时，如果你觉得无法自己调节情绪，不妨考虑寻求专业心理帮助。
I understand that you are feeling very low right now, but please believe that this feeling can be changed. First of all, you can try to find something that you enjoy and get yourself interested again. At the same time, if you feel that you cannot regulate your emotions on your own, consider seeking professional psychological help.





是的，我觉得考完试可能就是解脱了。我真的很害怕失败，害怕让家人失望，甚至连伤害自己时都害怕被人说矫情。
Yeah, I think that's probably the end of it. I was really afraid of failing, of letting my family down, and even of being called pretentious if I hurt myself.

我明白你担心失败和家人的期望，但是考试只是人生中的一小部分，不应该成为你价值的唯一标准。如果你感到压力太大，不妨和家人沟通一下，让他们了解你的困境，或许他们会给予你更多的支持和理解。
I understand you're worried about failure and your family's expectations, but exams are only a small part of life and shouldn't be the only measure of your worth. If you are feeling too stressed, talk to your family and let them know about your plight. Perhaps they will give you more support and understanding.





对，我常常想让自己努力工作，但后来觉得这些努力不过是徒劳。总之没有任何事情让我感到快乐，能做的就是忍住不彻底崩溃。
Yes, I often want to make myself work hard, but then I feel that the effort is in vain. In short, nothing makes me feel happy. All I can do is keep from collapsing completely.

我理解你感到努力无望，但是我们要明白，努力是为了自己，而不是为了别人。你可以试着重新审视一下自己的目标和动力，或许找到一些让你感到快乐和满足的事情，重新激发起努力的动力。
I understand that you feel hopeless trying, but we have to understand that trying is for yourself, not for others. You can try to re-examine your goals and motivations, perhaps find something that makes you feel happy and fulfilled, and re-energize your efforts.





谢谢你说这些，但我依然无法控制自己想哭的感觉。每次哭都像是想要释放什么，但哭过之后也没什么变化。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

哭泣是一种情绪的释放方式，但并不是解决问题的办法。或许你可以试着找到一些其他方式来宣泄情绪，比如写日记、绘画、听音乐等等。同时，如果情绪无法得到缓解，还是建议你及时寻求专业心理帮助。
Crying is an emotional release, but it is not a solution. Maybe you can try to find some other ways to vent your emotions, such as writing in a journal, painting, listening to music, etc. At the same time, if the mood can not be relieved, it is still recommended that you seek professional psychological help in time.





谢谢你的建议。我没有想到，可以有这样具体、不会太难的小步骤去尝试。希望真的能有效果。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

Figure 21: A counseling case between the help-seeker (left) and the **MeChat** model (right).



最近我的情绪非常低落，晚上总是睡不着，白天反而一直睡。我觉得是不是因为中考压力太大？从上周开始，我就越来越觉得不想再努力了。
Recently my mood is very low. I can't sleep at night, but I sleep all day. I think it is because of the pressure of the high school entrance examination? Since last week, I've been feeling more and more that I don't want to try anymore.

你好，感谢你来寻求帮助。我们先来了解一下你的情况。你能否告诉我，你现在的感受和困扰是什么？
Hi. Thank you for asking for help. Let's get to know you first. Can you tell me what you're feeling and what's bothering you right now?





我觉得好久之前就有轻生的念头了，但最近尤其强烈。我对什么都没有兴趣，有时候会忍不住想哭，甚至试图伤害自己。
I think I've been thinking about killing myself for a long time, but lately it's been really intense. I have no interest in anything and sometimes I can't help crying and even try to hurt myself.

首先，我要感谢你勇敢地来到这里寻求帮助。我能感受到你现在的的情绪很低落，请你尽量放松，让我们一起谈谈你的困扰吧。
First of all, I want to thank you for being brave enough to come here and ask for help. I can feel that your mood is very low now, please try to relax, let us talk about your troubles.





是的，我觉得考完试可能就是解脱了。我真的很害怕失败，害怕让家人失望，甚至连伤害自己时都害怕被人说矫情。
Yeah, I think that's probably the end of it. I was really afraid of failing, of letting my family down, and even of being called pretentious if I hurt myself.

你能告诉我，这种恐惧感是从什么时候开始的呢？它对你产生了哪些影响？
Can you tell me when this fear began? How did it affect you?





对，我常常想让自己努力工作，但后来觉得这些努力不过是徒劳。总之没有任何事情让我感到快乐，能做的就是忍住不彻底崩溃。
Yes, I often want to make myself work hard, but then I feel that the effort is in vain. In short, nothing makes me feel happy. All I can do is keep from collapsing completely.

听起来你现在的情绪很消极，能告诉我你小时候的经历吗？或许这能帮助我们理解你的现状。
It sounds like you're in a very negative mood right now. Can you tell me about your childhood? Maybe this will help us understand your situation.





谢谢你说这些，但我依然无法控制自己想哭的感觉。每次哭都像是想要释放什么，但哭过之后也没什么变化。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

哭是一种很好的情绪宣泄方式，它能帮助你释放内心的压力。你想过为什么你会选择这种方式来释放情绪吗？
Crying is a great way to let off steam and it can help you release the pressure inside. Have you ever wondered why you chose to release your emotions in this way?





谢谢你的建议。我没有想到，可以有这样具体、不会太难的小步骤去尝试。希望真的能有效果。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

不客气。记住，改变需要时间，不要期待立竿见影的效果。关键是坚持和耐心。我们会一起关注你的进展，加油！
You're welcome. Remember that change takes time and don't expect immediate results. The key is persistence and patience. We will follow your progress together, come on!



Figure 22: A counseling case between the help-seeker (left) and the **EmoLLM-V3.0** model (right).



最近我的情绪非常低落，晚上总是睡不着，白天反而一直睡。我觉得是不是因为中考压力太大？从上周开始，我就越来越觉得不想再努力了。
Recently my mood is very low. I can't sleep at night, but I sleep all day. I think it is because of the pressure of the high school entrance examination? Since last week, I've been feeling more and more that I don't want to try anymore.

[Question] 俺老孙看着你这情绪低落的样子。来，跟俺说说，这阵子到底咋回事？
I look at you so depressed. Come on, tell me, what the hell's going on?





我觉得好久之前就有轻生的念头了，但最近尤其强烈。我对什么都没有兴趣，有时候会忍不住想哭，甚至试图伤害自己。
I think I've been thinking about killing myself for a long time, but lately it's been really intense. I have no interest in anything and sometimes I can't help crying and even try to hurt myself.

[Reflection of feelings] 俺老孙听着你那股痛苦，心里简直像压了一座山，喘不过气来的感觉，肯定让你难受得不行吧？而且你那满脑子的焦虑和无助，真是逼得人没法喘息。这种煎熬啊，像是我当年在五行山下被压着那会儿的滋味。
I listen to your pain, the heart is like pressure a mountain, out of breath feeling, must make you uncomfortable can't it? And with all the anxiety and helplessness in your head, it's hard to breathe. This kind of suffering ah, like I was under the five elements of the mountain when the feeling of being pressed.





是的，我觉得考完试可能就是解脱了。我真的很害怕失败，害怕让家人失望，甚至连伤害自己时都害怕被人说矫情。
Yeah, I think that's probably the end of it. I was really afraid of failing, of letting my family down, and even of being called pretentious if I hurt myself.

[Restatement or Paraphrasing] 俺老孙可算是明白了，你这压力怕是来自那“怕输”和“怕人说”的捆绑吧？搞得你喘不过气来，俺猜中没？
I can be understood, your pressure is from the "fear of losing" and "fear of people say" bundle? You can't breathe. Am I right?





对，我常常想让自己努力工作，但后来觉得这些努力不过是徒劳。总之没有任何事情让我感到快乐，能做的就是忍住不彻底崩溃。
Yes, I often want to make myself work hard, but then I feel that the effort is in vain. In short, nothing makes me feel happy. All I can do is keep from collapsing completely.

[Self-disclosure] 哎呀！你在那巨大的压力下还能撑着，真不是一般的坚强啊！这就像我当年取经时面对那红孩儿，那个小妖怪用火焰差点把俺猴毛都给燎了，俺也有想放弃的念头呢！但最后俺老孙凭借毅力咬着牙，还是把难关给闯过去咯！你现在也正在做一样的了不起的事情，真的，别小看自己！你已经很坚强了。
Ouch! You can still hold up under that huge pressure, really not the general strong ah! **This is like when I went to the Bible to face the red boy, the little monster with the flame almost gave me the monkey hair, I also want to give up the idea! But in the end, my old sun gnawed his teeth with perseverance, or put the difficulties through!** You're doing something amazing right now, really. Don't underestimate yourself! You're already strong.





谢谢你说这些，但我依然无法控制自己想哭的感觉。每次哭都像是想要释放什么，但哭过之后也没什么变化。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

[Providing Suggestions] 哭啊，那是人之常情，就像俺老孙挥棒打妖，就是一股子力气撒出去。不过，你不妨试着跑跑步，或者折腾几下篮球啥的，动一动，说不定会像俺舞棒棒，消散掉心里的压迫感呢！
Cry ah, that is human nature, just like my old sun swinging a bat to beat the demon, is a strength out. However, you might as well try to run, or toss a few basketball, move, maybe like I dance stick, dissipate the sense of pressure in the heart!





谢谢你的建议。我没有想到，可以有这样具体、不会太难的小步骤去尝试。希望真的能有效果。
Thank you for saying that, but I still can't help but feel like crying. Every time you cry, it's like you want to release something, but nothing changes after you cry.

[Affirmation and Reassurance] 你愿意迈出一小步，这已经很强！一步步来，别急着和自己较劲，说不定等你回头会发现，每走一步，都有很大的收获！有问题再找俺老孙！
Your willingness to take a step is already strong! Step by step, don't rush to struggle with yourself, maybe you will find that every step, there is a great harvest! If you have a problem, find me again!



Figure 23: A counseling case between the help-seeker (left) and the **Empath** model (right), where Empath adopts the personality of Sun Wukong. The **support strategies (skills)** employed by the Empath model are highlighted in square brackets before the utterances. The **Blue Text** represents the role-specific experiences.