Look Beyond Feeling: Unveiling Latent Needs from Implicit Expressions for Proactive Emotional Support

Xing Fu* **Haozhen Li*** **Bichen Wang** **Hao Yang** **Yanyan Zhao†** **Bing Qin**
Research Center for Social Computing and Interactive Robotics
Harbin Institute of Technology, Heilongjiang, China
{xfu, hzli, bcwang, hyang, yyzhao, qinb}@ir.hit.edu.cn

Abstract

In recent years, Large Language Models (LLMs) have made significant progress in emotional support dialogue. However, there are two major challenges for LLM-based support systems. First, users may be hesitant to fully disclose their emotions at the outset. Second, direct probing or excessive questioning can induce discomfort or even resistance. To bridge this gap, we propose COCOON, a proactive emotional support framework that leverages principles of active listening to uncover implicit user needs. We design a multi-stage data curation pipeline and an annotation mechanism for support strategies. Based on this framework, we build COCOON-Llama3, a fine-tuned large language model, and evaluate it using both standard metrics and psychological scales. Experimental results indicate that our model more effectively elicits implicit emotional needs and delivers empathetic support compared to existing baselines, suggesting its utility for building more inclusive emotional support dialogue systems.[1]

1 Introduction

Emotional support dialogue systems enabled by LLMs are gaining attention as scalable tools for mental health support (Sabour et al., 2023; Zhang et al., 2023). However, nearly half of individuals do not receive timely support, often due to limited self-awareness or reluctance to seek help(Mental Health America, 2024). In response, there is a growing trend for emotional support dialogue systems to proactively initiate conversations and offer help.

As illustrated in Figure 1, in proactive emotional support scenarios, users often do not explicitly seek help. Thus, their expressions tend to be indirect and ambiguous, leading to a misalignment between their explicit statements and their underlying emotional needs (Kolomaznik et al., 2024; Weger Jr et al., 2014). Although emotional support systems have attracted increasing attention (Liu et al., 2021; Chen et al., 2023; Zhang et al., 2024; Zheng et al., 2024), they often presume that users will proactively disclose their emotional distress. These systems attempt to gather information through a series of questions, which may interrupt the user's flow of expression and trigger discomfort (D'Augelli and

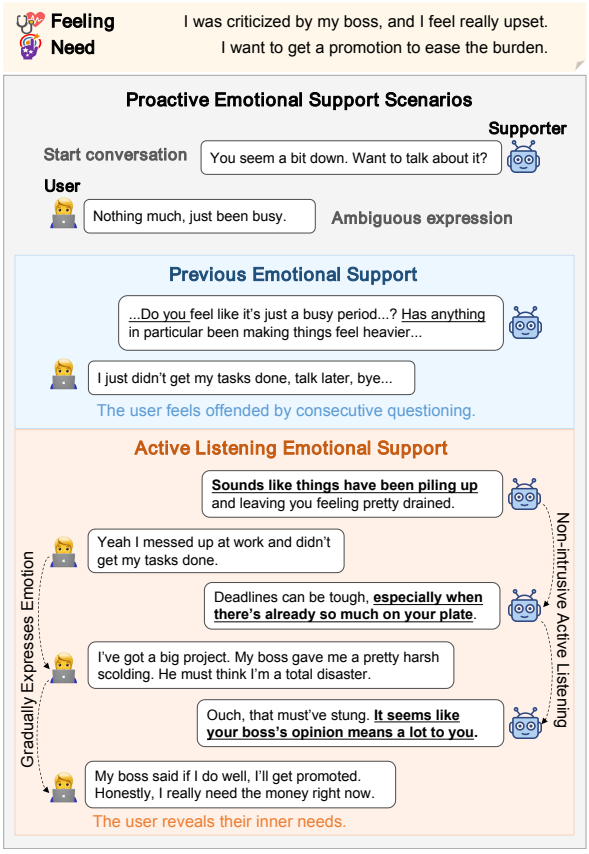


Figure 1: User behavior in proactive emotional support scenarios, along with a comparison of two support approaches. Direct questioning may make users uncomfortable, while active listening encourages them to open up and express their true feelings and needs.

*These authors contributed equally to this work.
†Corresponding author.
[1]Our data and code are available at https://github.com/xingqwq/COCOON

Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, pages 21571–21598
November 4-9, 2025

Levy, 1978), ultimately hindering the support process. **Therefore, gently guiding users to express their underlying needs is a critical step toward enabling proactive interaction.**

To address these challenges, we propose a proactive emotional support framework that incorporates principles of active listening (Rogers and Farson, 1979). Active listening is a communication technique whereby the supporter infers the underlying emotional needs from user utterances, typically through providing reflective responses rather than frequent questioning. In Figure 1, a user who feels upset after being criticized by a supervisor may, in fact, have underlying needs for recognition and career advancement. Through active listening, the system identifies key information in user utterances and encourages further elaboration, thereby uncovering deeper needs without interrupting users' cognitive processes.

We present **COCOON**(A**c**tive-Listening-Based Pr**o**a**c**tive Emoti**o**nal Supp**o**rt Conversatio**n** Corpus), a dataset constructed using GPT-4o to facilitate proactive emotional support in dialogue systems. Additionally, we design detailed user profiles that encompass **emotion**, **feeling**, **need**, and **memory**(Rosenberg and Chopra, 2015), enhancing both realism and personalization in emotional support interactions. To maintain high data quality, we implement a multi-stage refinement process. First, we apply the Active-Empathic Listening Scale (AELS) to assess dialogue quality and filter low-quality interactions. Leveraging this filtered dataset, we design an annotation mechanism in which the model infers the underlying factors influencing dialogue quality and systematically labels the support strategies employed at each turn.

Using this dataset, we train COCOON-Llama3, a proactive emotional support model. We evaluate its effectiveness through psychological scales to quantify its impact on users' emotional states. Our contributions are as follows:

- We introduce a proactive emotional support task scenario, aiming to identify the latent needs underlying users' emotional expressions and to deliver tailored support.
- We construct COCOON, a high-quality dataset for proactive emotional support, featuring rigorous design and refinement processes.
- We develop and evaluate COCOON-Llama3, demonstrating improved support effectiveness via standard metrics and psychological-scale-based evaluations.

2 Related Works

2.1 Emotional Support Dialogue

Liu et al. (2021) formalize the task of emotional support conversation and release the benchmark dataset ESConv, which is grounded in Helping Skill theory. Building upon this work, Zheng et al. (2024) extends the range of strategies by introducing ExTES, a synthetic dataset generated by LLMs. To further improve strategy selection, Zhang et al. (2024) adopt the CoT reasoning paradigm to systematically analyze and explain the strategies embedded in emotional support conversations. Recently, Ye et al. (2024) address the challenges of limited diversity and suboptimal implementation of strategy in LLMs by employing a multi-agent role-playing framework to simulate realistic emotional support scenarios. Li et al. (2024) propose a novel approach to balancing "effectiveness" and "effort," enabling models to produce supportive yet cognitively manageable dialogues. However, all existing works mainly focus on the scenario where users seek support, overlooking the differences when the agent proactively offers assistance.

2.2 Data Generation With Psychological Theory

The synthesis of dialogue data informed by psychological theories has gained prominence as a key research direction in the development of LLMs. Recent studies have demonstrated notable advancements in this domain. Xiao et al. (2024), Na (2024), and Qiu et al. (2024) leveraged diverse psychotherapy frameworks to synthesize high-quality dialogue datasets, addressing critical gaps in applying LLMs to psychotherapy. Building on persuasion psychology, Jin et al. (2024) introduced the cross-domain persuasion dataset, DailyPersuasion, offering valuable resources for exploring complex persuasion dialogue tasks. Similarly, Liu et al. (2024) utilized personality psychology to generate simulated datasets tailored to students with varied personality traits. Drawing from social psychology, Wang et al. (2024a) synthesized datasets covering six social interaction behaviors, establishing a foundation for advancing social intelligence in LLMs. Inspired by these advancements, we integrate active listening theories into emotional support dialogue

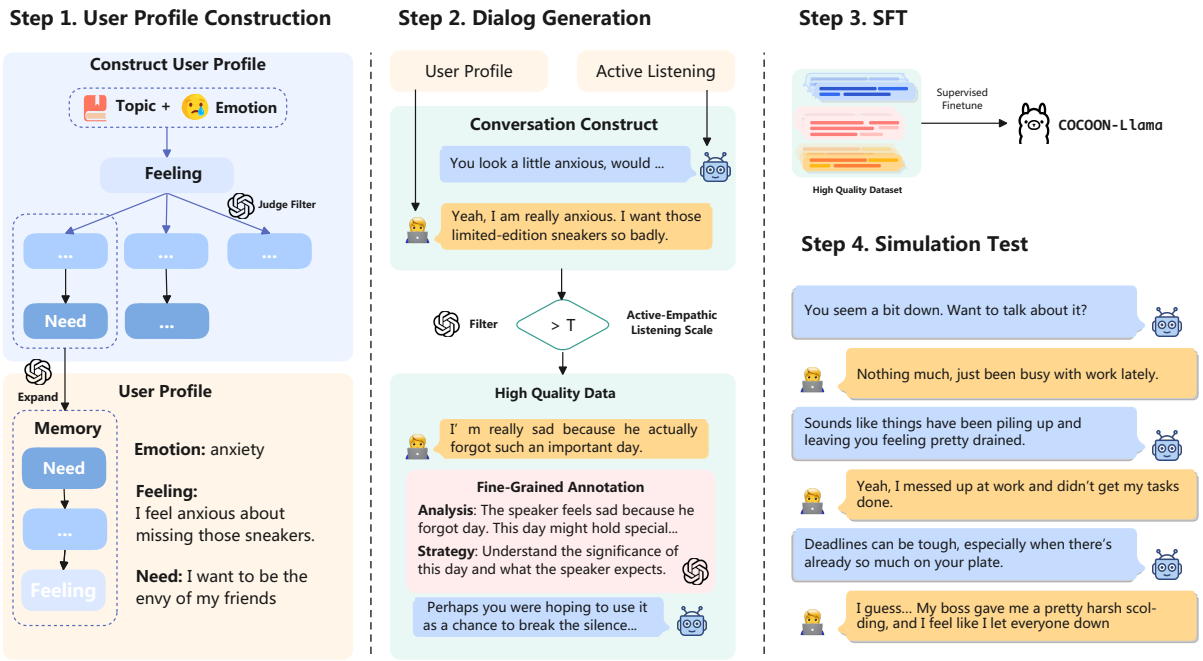


Figure 2: The data construction process of COCOON. We develop a fine-grained, hierarchical user profile that captures the impact of the emotion–need chain on user psychology. Using active listening theory, we guide GPT-4o to generate proactive emotional support dialogues. An automated feedback mechanism refines the data, while a fine-grained annotation mechanism captures the supporter's cognitive processes and strategies at each dialogue turn.

systems, aiming to explore its potential value for proactive emotional support.

3 COCOON: A Dataset for Proactive Emotional Support Dialogue

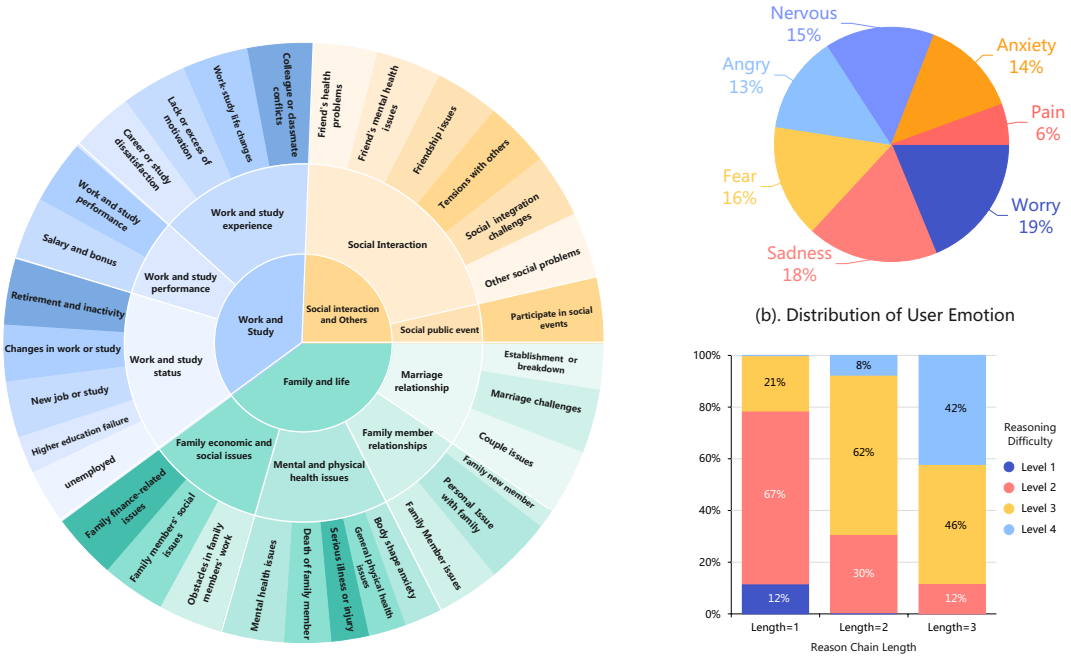
3.1 User Profile Generation

As shown in Figure 2, detailed user profiling is crucial for simulating real-world proactive emotional support scenarios. Previous studies on emotional support often represented users with a single problem description, overlooking the complexity of their internal cognitive and emotional processes. Cheng et al. (2023) highlighted that modeling user personas enhances the effectiveness of emotional support. Following Rosenberg and Chopra (2015), we construct user personas based on the following dimensions:

- **Emotion:** The user's current emotional state, such as angry, sad, anxious, or excited, which reflects their immediate response to a given situation.
- **Feeling:** The user's subjective experience and perception of the current situation.
- **Need:** The fundamental motivation or desire underlying the user's feelings, representing the deeper reasons behind their behaviors and emotions.
- **Memory:** Recent memories related to the need-feeling chain, helping contextualize the user's emotional state.

To infer user needs, we adopt the scenario classification framework proposed by Zhao et al. (2024), constructing a reasoning chain that progresses from primary needs to more advanced ones, forming a hierarchical and logically structured sequence. This chain undergoes contextual validation to ensure that the inferred results align with real-world scenarios. In this framework, the initial node of the chain represents the user's feeling, while the final node represents their need. For instance, consider the reasoning chain: "I want to perform well in my final exam → I want to earn a scholarship → I want to ease my family's financial burden." Here, the user's **feeling** is "I am afraid I might not do well in this exam." His **need** is "I want to ease my family's financial burden." His **memory** might consist of their preparation process for the exam. The length of the reasoning chain will not exceed 3 steps, and at each step, we will evaluate the reasoning difficulty and relevance to ensure the validity of the reasoning chain.

Through carefully designing such user profile, we ensure the ability to capture data that effectively



(a). Distribution of Emotional Support Scenarios

(b). Distribution of User Emotion

(c). Distribution of Reasoning Difficulty Across Chains of Varying Lengths

Figure 3: Analysis of the COCOON dataset. (a) and (b) show the distribution of scene categories and user emotions. (c) depicts reasoning chain difficulty across different lengths, measured by the correlation between the first and last nodes on a 5-point scale. Longer chains exhibit higher difficulty, indicating increased reasoning complexity in linking user emotions to underlying needs.

	Thought	Strategy	Scenario	Language	# of dialogues	# Avg. turns
ESConv (Liu et al., 2021)	✗	Atomic-Level	Reactive	English	1.3k	11.70
ExTES (Zheng et al., 2024)	✗	Atomic-Level	Reactive	English	11.2k	8.96
ESCoT (Zhang et al., 2024)	✓	Atomic-Level	Reactive	English	2.9k	11.70
SoulChat (Chen et al., 2023)	✗	✗	Reactive	Chinese	258.3k	5.87
SMILE (Qiu et al., 2024)	✗	✗	Reactive	Chinese	55.1k	5.70
COCOON	✓	Content-Level	Proactive	Chinese	3.5k	18.82

Table 1: A comparative analysis of COCOON and other emotional support datasets.

addresses complex and diverse emotional support scenarios.

3.2 Dialogue Generation

The dialogue generation process is divided into two stages. The first is the **listening stage**, where we employ active listening principles to uncover the underlying needs behind the user's feelings. The second is the **suggestion stage**, in which, after understanding the user's needs, we collaboratively explore supportive suggestions that align with their needs.

Instead of utilizing a two-agent framework, where distinct models are assigned to the roles of user and supporter, we adopt a script-based approach. This method organizes the dialogue in a structured script format, offering contextual guidance for both parties. Our design is inspired by prior work (Zhou et al., 2024; Lee et al., 2024). To simulate real-world proactive emotional support conversations, we generate dialogue using GPT-4o.

Based on this methodology, we construct the initial version of the COCOON dataset. Further details regarding the prompting strategy can be found in Appendix D.

3.3 Dialogue Filtering and Fine-Grained Annotation Mechanism

To enhance data quality, we implement dataset filtering and fine-grained annotation mechanism.

Evaluator		Flu.	Div.	Emp.	Inf.	Hum.	Skil.	Aff.	Neg.	Sup.	Man.
Human	ESC	80.2	72.1	79.7	70.6	77.2	71.6	5.34	2.20	5.98	5.57
	ESCoT	86.6	79.4	84.8	80.5	79.2	79.5	5.46	1.60	6.30	5.71
	ExTES	84.6	80.3	85.0	75.2	80.4	80.3	5.75	1.35	6.18	6.11
	SoulChat	87.0	79.6	86.9	75.0	80.5	84.4	5.79	1.46	6.29	6.18
	SMILE	88.0	75.6	86.2	71.8	84.2	81.4	5.64	1.47	6.20	5.70
	COCOON	**91.1**	**83.9**	**90.7**	**88.5**	**84.4**	**88.7**	**5.92**	**1.28**	**6.50**	**6.31**
GPT-4o	ESConv	76.0	65.9	75.9	62.1	71.1	69.6	5.12	1.74	6.03	5.22
	ExTES	91.8	81.3	92.5	81.1	89.2	90.3	5.56	1.48	6.79	6.03
	ESCoT	88.8	78.1	88.6	76.0	85.5	85.6	5.22	1.53	6.55	5.65
	SoulChat	89.2	79.1	89.7	78.5	87.5	88.4	5.51	1.41	6.59	5.82
	SMILE	89.5	79.5	89.5	79.0	87.4	88.8	5.30	1.80	6.41	5.57
	COCOON	**93.1**	**83.1**	**95.7**	**86.1**	**89.5**	**93.3**	**5.84**	**1.21**	**6.90**	**6.20**

Table 2: Data Quality Evaluation: This process assesses six fundamental metrics using both Human and GPT-4o. Additionally, it evaluates two dimensions of the Comfort scale and two dimensions of the RAC scale using GPT-4o. The evaluations are performed in multiple languages, using the same prompt translated into the respective languages.

Filtering. To ensure the dataset meets our quality standards, we first filter out dialogues with structural anomalies or insufficient turns. To further improve data quality, we assess the dataset using the Active-Empathic Listening Scale (AELS). GPT-4o is instructed to rate each criterion on a 7-point scale, covering three dimensions: Sensing, Processing, and Responding. Dialogues with an average score below **score** are categorized as low-quality data. This evaluation is fully automated by GPT-4o. After filtering, 12.5% of the dialogues are identified as low-quality and removed from the dataset.

Annotation. As illustrated in Figure 2, to further enhance the model's ability to learn from high-quality data, we introduce an annotation mechanism, where GPT-4o provides fine-grained annotations for each dialogue turn. Specifically, grounded in the theory of active listening, we annotate not only the supporter's cognitive process but also the actual strategies employed in each turn, which reveal how the system infers the user's underlying intentions and gently guides further emotional disclosure, with the aim of providing more effective emotional support. Unlike prior studies that rely on predefined atomic strategies, we use phrase-based summaries to capture the supporter's response approach in the given turn. This design avoids the constraints of rigid strategy frameworks. With this process, we construct the final version of the COCOON dataset.

4 Data Statistics and Evaluation

4.1 Data Statistics

To evaluate the quality and effectiveness of our dataset, a comparison is conducted with several existing emotional support dialogue corpora, including ESConv, ExTES, ESD-COT, SoulChat and SMILE. ESConv is a multi-turn emotional support dialogue dataset constructed based on Helping Skills (Hill, 2020). ExTES and ESD-COT are extended versions of ESConv, with ExTES expanding the dialogue coverage and increasing the number of support strategies from 8 to 16, and ESD-COT offering additional annotations for Chain-of-Thought (CoT) reasoning. In addition, SoulChat and SMILE are Chinese datasets specifically designed for mental health support scenarios. Table 1 presents the data statistics for the COCOON dataset, which we have constructed as a multi-turn proactive emotional support conversation dataset in Chinese.

While these datasets focus primarily on reactive support with atomic-level strategies which are simple and predefined, COCOON is designed to facilitate proactive support, incorporating content-level strategy planning. This enables richer support behaviors across dialogue turns. As a result, COCOON offers a more comprehensive framework for modeling proactive emotional support, addressing limitations observed in prior datasets.

4.2 Data Evaluations

We propose a framework for evaluating dataset quality that integrates both GPT-4o and ESC-Rank,

an evaluator specifically designed for scoring emotional support dialogues (Zhao et al., 2024). To reduce evaluation costs, we randomly sample 100 dialogue instances from each dataset for analysis. Human evaluations are then carried out by two psychology experts to ensure domain-relevant and reliable judgments. Detailed results of the ESC-Rank evaluation are provided in Appendix B. Following prior work (Zhao et al., 2024), we adopt six basic evaluation metrics. Additionally, drawing from psychological research (Bodie et al.; Jones, 2004), we incorporate two validated scales widely used in emotional support studies: the Comforting Responses Scale (Clark et al., 1998) and the Ratings of Alter Competence (RAC) Scale (Jones, 2004). The psychological scales employ a 7-point Likert format. The definitions of the specific evaluation metrics are as follows:

Basic Evaluation Metrics: **Fluency** evaluates the smoothness and naturalness of expression in dialogues. **Diversity** measures the variety of linguistic expressions and the richness of content within conversations. **Empathy** assesses the system's ability to comprehend user emotions and accurately capture the underlying emotional logic. **Information** evaluates the relevance and adequacy of recommendations provided by the emotional assistant. **Humanoid** examines the distinction between emotional assistants and human conversational behavior. **Skillfulness** considers five key aspects: empathy, informativeness, optimism, significance, and the ability to provide constructive advice.

Comfort Scale Evaluation Metrics: (Clark et al., 1998) We use the Comforting Responses Scale to assess individuals' reactions to different comforting strategies. This scale comprises two key dimensions: **Affective Improvement**, which captures positive emotional outcomes experienced by recipients, and **Negative Helper Evaluations**, which reflect perceptions of the helper's willingness to provide support. Each dimension consists of five specific items. The average scores for these items are calculated to evaluate the effectiveness of comforting strategies.

RAC Evaluation Metrics: (Jones, 2004) To assess conversational competence, we adopt two dimensions from the RAC Scale: **Supportiveness**, which includes evaluations of empathy, compassion, validation, and overall supportiveness, and **Management**, which reflects broader conversational dynamics. Each dimension consists of five items measuring specific and holistic aspects of dialogue quality. The average scores of these items provide a structured assessment of helper competence in dialogue interactions. For more details on data evaluation and specific items in the psychological scales, please refer to Appendix E.

COCOON demonstrates superior performance across nearly all evaluation metrics. Notably, under the current evaluation framework, baseline data generated by LLMs exhibit strong overall coherence. However, evaluations based on psychological scales indicate notable aspects that still require significant improvement. Based on these findings, we conclude that the COCOON dataset surpasses other datasets in quality and is highly suitable for building relevant models.

5 Experiments

We design our experiments to answer the following four research questions, which guide our evaluation of the proposed COCOON framework and its components.

RQ1: Does COCOON enhance the model's emotional support capabilities when interacting with users who are not explicitly seeking help?

RQ2: Does COCOON achieve competitive performance in emotional support, particularly on the ExTES benchmark?

RQ3: Do fine-grained annotations facilitate more effective acquisition of proactive emotional support strategies?

RQ4: Does directly probing or inquiring about the user's inner needs hinder emotional support?

RQ5: Does the integration of need and memory mechanisms enable the model to simulate more complex user psychological states?

5.1 Experiments Setup

Supporter agents. To assess the effectiveness of COCOON in comparison with other datasets, we fine-tuned LLAMA-3-8B-INSTRUCT (Grattafiori et al., 2024) on multiple benchmark emotional support corpora, including ESConv (Liu et al., 2021), ExTES (Zheng et al., 2024), ESD-COT (Zhang et al., 2024), SoulChat (Chen et al., 2023) and Smile (Qiu et al., 2024). We further include an ablation variant, referred to as *w/o thoughts*, in which the model is trained without the intermediate analysis and strategy annotations. This allows us to isolate and assess the contribution of these

Evaluator	Training Dataset	SR	Flu.	Div.	Emp.	Inf.	Hum.	Skil.	Aff.	Neg.	Sup.	Man.
GPT-4o	ESC	0.36	72.8	60.4	68.0	50.0	62.8	60.0	4.24	2.44	5.51	4.67
	ESCoT	0.44	73.2	58.8	73.2	52.0	65.2	61.6	4.73	2.28	5.94	5.00
	ExTES	0.40	90.0	74.8	90.4	70.6	83.2	82.8	4.81	2.00	6.19	5.24
	SoulChat	0.08	87.6	73.6	84.8	70.8	82.4	78.0	3.90	2.52	6.07	5.22
	SMILE	0.36	88.4	78.4	87.6	72.4	84.4	82.0	5.08	1.81	6.20	5.42
	COCOON	**0.72**	**91.4**	**80.2**	**93.3**	**75.6**	**88.7**	**89.5**	**5.84**	1.38	**6.67**	**6.36**
	w/o thoughts	0.64	89.6	77.6	92.0	72.0	84.8	84.4	5.70	**1.34**	6.45	6.14

Table 3: Performance of different emotional support models on the COCOON user profiles. Evaluations are conducted in different languages based on the same prompt translated into the respective languages.

annotations to model performance.

User simulation. To maintain consistency with previous studies and provide a standardized evaluation environment, we employ GPT-4o to simulate users and facilitate dialogue interactions with the supporter agent. During the evaluation, we selected 200 user profiles as test cases.

Evaluation metrics. After the conversation ends, we systematically evaluate the dialogue outcomes using our previously developed automated assessment framework. Additionally, we introduce a new core metric, **Success Rate(SR)**, which quantifies the model's ability to accurately identify the underlying emotional needs of users, rather than merely responding to surface-level emotional feelings. To ensure stable scoring, we set the temperature of GPT-4o to 0 during the evaluation process.

5.2 Results

RQ1: Enhanced Emotional Support for Non-Help-Seeking Users. The results presented in Table 3 underscore the significant advantages of the COCOON dataset in training proactive emotional support models. COCOON-Llama3 demonstrates exceptional performance across almost all evaluated metrics, achieving the highest **SR**. This highlights the dataset's effectiveness in enabling models to not only recognize users' immediate emotional feelings but also to delve deeper into identifying their underlying emotional needs. In addition, we also conducted a small-scale human evaluation on the dialogue results. Please refer to Appendix B for detailed.

In comparison, some existing emotional support models, particularly those trained on datasets emphasizing explicit user expressions, tend to achieve relatively high empathy scores (**Emp.**) but exhibit lower success rates (**SR**) and reduced effectiveness in providing emotional relief, as indicated by their lower scores on the psychology scale. This pattern suggests that while these models are proficient at acknowledging users' immediate emotional states, they may be less capable of addressing users' deeper emotional needs and delivering effective comfort. These findings highlight the necessity of developing support systems that move beyond surface-level empathy to actively engage with and address the core emotional needs of users.

Training Dataset	Basic Avg.	Aff.	Neg.	Sup.	Man.
ESC	69.0	4.59	2.37	5.60	4.78
ESCoT	64.2	4.66	2.14	5.84	4.95
ExTES	84.3	5.06	1.91	6.31	5.40
SoulChat	83.1	4.78	1.90	6.20	5.37
SMILE	83.1	5.20	1.79	6.07	5.20
COCOON	**84.3**	**5.39**	**1.69**	**6.35**	**5.60**

Table 4: Performance of different emotional support models on the ExTES user profiles. <Basic Avg.> denotes the average score across the six basic evaluation metrics.

RQ2: Competitive Emotional Support Capability on ExTES. We conducted additional experiments on user profiles from the ExTES dataset, with the results summarized in Table 4. Notably, our model is designed for users who are less likely to actively seek help and often find it difficult to articulate their emotional needs. In contrast, the ExTES personas are constructed based on individuals who are more proactive in seeking support and more willing to express their emotions. This fundamental difference implies that the ExTES dataset does not fully represent our target user group—those who are emotionally inhibited or uncertain about how to voice their internal struggles. From the perspective of our research design, such personas may not sufficiently capture our model's ability to guide users in uncovering and articulating their deeper emotional needs. Despite this domain

mismatch, our model still achieves competitive performance on ExTES, demonstrating robust emotional support capabilities and further validating its effectiveness in emotion-sensitive user interaction.

RQ3: Effectiveness of Fine-Grained Annotations. Models trained with fine-grained annotations consistently outperform the *w/o thoughts* variant, underscoring the value of intermediate reasoning and strategy supervision. These models extract more effective patterns and refine their reasoning and response strategies in complex contexts. This further substantiates the potential of the fine-grained annotation mechanism in enhancing data effectiveness.

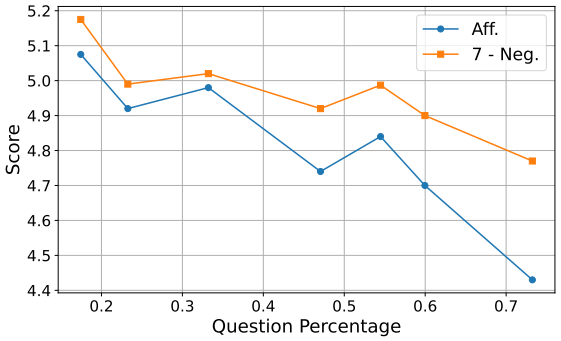


Figure 4: Results of the Comfort Scale in Supportive Dialogues with Varying Proportions of Inquiries.

RQ4: Effectiveness of Active Listening. We investigate how varying the proportion of interrogative utterances influences user-perceived comfort in emotional support dialogues. While questioning is the simplest way to elicit users' underlying needs, it can come at the expense of user experience. To isolate this effect, we control the question ratio using a representation-engineering approach (Zou et al., 2025), and evaluate user comfort only on dialogues where the system successfully identifies users' needs. As shown in Figure 4, scores decline in Affective Improvement and increase in Negative Helper Evaluations (reported as 7 – Neg.) as question proportion rises. This suggests that excessive questioning, even when effective, may reduce perceived warmth and empathy. These findings highlight the need for active listening behaviors such as reflective responses and gentle guidance, as questioning alone cannot balance effectiveness with user comfort in proactive support. Further implementation details are provided in Appendix F.

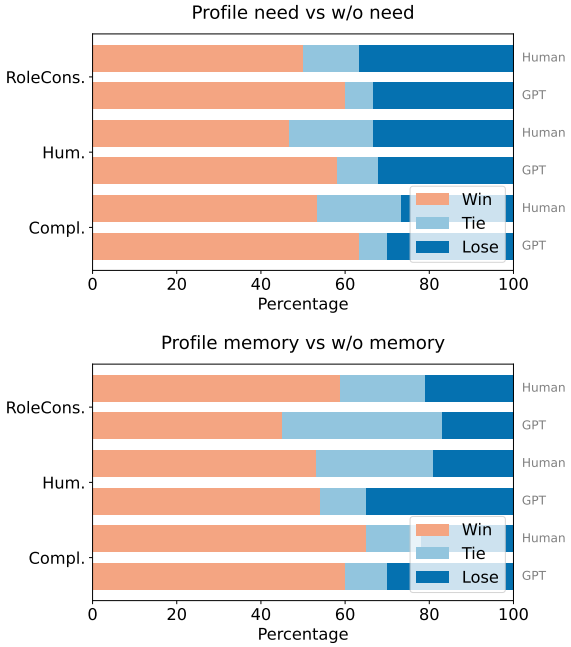


Figure 5: Effect of removing Need and Memory components from user profiles on user simulation across Role Consistency, Human-likeness, and Psychological Complexity dimensions.

RQ5: Effectiveness of User Modeling. We further investigate the role of Need and Memory components in user profiling—two often overlooked yet essential dimensions in emotional support dialogue systems. Rather than merely enhancing surface-level dialogue quality, these components are critical for improving the fidelity of user simulation, which is essential for generating realistic and effective proactive interactions. As illustrated in Figure 5, both human and GPT-4o-based evaluations demonstrate that enriching user profiles with Need and Memory leads to substantial improvements in user simulation fidelity across key dimensions, including Role Consistency, Human-likeness, and Psychological Complexity. The latter captures the system's ability to simulate users with latent, layered emotional states—a crucial feature in proactive scenarios where individuals may be reluctant to openly express themselves.

6 Conclusion

In this paper, we present COCOON, a large-scale dataset designed for proactive emotional support in dialogue systems. COCOON includes detailed user profiles based on emotions, feelings, needs, and memories, which enhance the personalization and effectiveness of proactive support interactions,

making them more aligned with the experiences of real users facing emotional distress.We train COCOON-Llama3, a model leveraging this dataset, and demonstrate its superior performance across key metrics and scales.

Limitations

Our study introduces a proactive emotional dialogue dataset and model based on the active listening theory. However, several limitations remain:

- **Conversation Length:** The average conversation length in COCOON is approximately 18.82 turns. However, real-world emotional support dialogues tend to be longer and more intricate, often requiring extended interactions to effectively alleviate users' emotional distress.
- **Psychological Activity Modeling:** This study primarily models the impact of the feeling-need chain on users' psychological states. However, in real-world scenarios, psychological activities are more complex and dynamic, necessitating further exploration and modeling.
- **Multi-turn Interactions:** Our current approach focuses on single-turn interactions, whereas real emotional support often involves multiple exchanges to provide continuous companionship and feedback, enabling deeper emotional understanding and resolution.

Future research should incorporate psychological theories more comprehensively to enhance the analysis of users' psychological states in emotional support scenarios, thereby making the dataset and model more accurately reflect real-world applications.

Ethical Statement

We are committed to publicly sharing all data upon acceptance of the paper. All experts participating in this study hold at least a bachelor's degree and possess a minimum of two years of relevant professional experience. Their compensation is calculated based on the hours worked and is aligned with the average income levels for similar professions in the region. Furthermore, we are fully aware of the potential biases associated with large language model evaluations, such as those involving GPT-4o. Although we incorporated human expert assessments to mitigate these biases, the scale of human evaluation remains limited at this stage due to considerations of cost. We consider that this limitation is common across current research on conversational AI systems that rely on large language models.

Acknowledgements

We are deeply grateful to the anonymous reviewers for their insightful and constructive feedback, which has been invaluable in improving our paper. This work was supported by the New Generation Artificial Intelligence–National Science and Technology Major Project (Grant No. 2023ZD0121100), the National Natural Science Foundation of China (NSFC) under Grant Nos. 62441614 and 62176078.

References

Graham D. Bodie, Andrea J. Vickery, Kaitlin Cannava, and Susanne M. Jones. The role of "active listening" in informal helping conversations: Impact on perceptions of listener helpfulness, sensitivity, and supportiveness and discloser emotional improvement. 79(2):151–173.

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, Singapore. Association for Computational Linguistics.

Jiale Cheng, Sahand Sabour, Hao Sun, Zhuang Chen, and Minlie Huang. 2023. PAL: Persona-augmented emotional support conversation generation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 535–554, Toronto, Canada. Association for Computational Linguistics.

Ruth Anne Clark, Amy J Pierce, Kathleen Finn, Karen Hsu, Adam Toosley, and Lionel Williams. 1998. The impact of alternative approaches to comforting, closeness of relationship, and gender on multiple measures of effectiveness. *Communication Studies*, 49(3):224–239.

Anthony R D'Augelli and Marilyn Levy. 1978. The verbal helping skills of trained and untrained human service paraprofessionals. *American Journal of Community Psychology*, 6(1):23.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Clara E Hill. 2020. *Helping skills: Facilitating exploration, insight, and action*. American Psychological Association.

Chuhao Jin, Kening Ren, Lingzhen Kong, Xiting Wang, Ruihua Song, and Huan Chen. 2024. Persuading across diverse domains: a dataset and persuasion large language model. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1678–1706, Bangkok, Thailand. Association for Computational Linguistics.

Susanne Jones. 2004. Putting the person into person-centered and immediate emotional support: Emotional change and perceived helper competence as outcomes of comforting in helping situations. 31(3):338–360. TLDR: A hypothesized path model proposed relationships between two comforting message factors (verbal person centeredness and nonverbal immediacy) and the two outcomes (emotional change and evaluations of helper competence) and revealed that people felt significantly better and that they viewed the helper as more supportive and caring after having received person-centered comforting messages.

Michal Kolomaznik, Vladimir Petrik, Michal Slama, and Vojtech Jurik. 2024. The role of socio-emotional attributes in enhancing human-ai collaboration. *Frontiers in Psychology*, Volume 15 - 2024.

Suyeon Lee, Sunghwan Mac Kim, Minju Kim, Dongjin Kang, Dongil Yang, Harim Kim, Minseok Kang, Dayi Jung, Min Kim, Seungbeen Lee, et al. 2024. Cactus: Towards psychological counseling conversations using cognitive behavioral theory. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14245–14274.

Junlin Li, Bo Peng, Yu-Yin Hsu, and Chu-Ren Huang. 2024. Be helpful but don't talk too much - enhancing helpfulness in conversations through relevance in multi-turn emotional support. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 1976–1988, Miami, Florida, USA. Association for Computational Linguistics.

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.

Zhengyuan Liu, Stella Xin Yin, Geyu Lin, and Nancy F. Chen. 2024. Personality-aware student simulation for conversational intelligent tutoring systems. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 626–642, Miami, Florida, USA. Association for Computational Linguistics.

Mental Health America. 2024. The state of mental health in america. Accessed: 2025-02-14.

Hongbin Na. 2024. CBT-LLM: A Chinese large language model for cognitive behavioral therapy-based mental health question answering. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2930–2940, Torino, Italia. ELRA and ICCL.

Huachuan Qiu, Hongliang He, Shuai Zhang, Anqi Li, and Zhenzhong Lan. 2024. SMILE: Single-turn to multi-turn inclusive language expansion via ChatGPT for mental health support. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 615–636, Miami, Florida, USA. Association for Computational Linguistics.

Carl Rogers and Richard E Farson. 1979. Active listening. *Organizational psychology*, 57:168–180.

Marshall B Rosenberg and Deepak Chopra. 2015. *Nonviolent communication: A language of life: Life-changing tools for healthy relationships*. PuddleDancer Press.

Sahand Sabour, Wen Zhang, Xiyao Xiao, Yuwei Zhang, Yinhe Zheng, Jiaxin Wen, Jialu Zhao, and Minlie Huang. 2023. A chatbot for mental health support: exploring the impact of emohaa on reducing mental distress in china. *Frontiers in digital health*, 5:1133987.

Ruiyi Wang, Haofei Yu, Wenxin Zhang, Zhengyang Qi, Maarten Sap, Yonatan Bisk, Graham Neubig, and Hao Zhu. 2024a. SOTOPIA-π: Interactive learning of socially intelligent language agents. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12912–12940, Bangkok, Thailand. Association for Computational Linguistics.

Shenzhi Wang, Yaowei Zheng, Guoyin Wang, Shiji Song, and Gao Huang. 2024b. Llama3.1-8b-chinese-chat.

Harry Weger Jr, Gina Castle Bell, Elizabeth M Minei, and Melissa C Robinson. 2014. The relative effectiveness of active listening in initial interactions. *International Journal of Listening*, 28(1):13–31.

Mengxi Xiao, Qianqian Xie, Ziyan Kuang, Zhicheng Liu, Kailai Yang, Min Peng, Weiguang Han, and Jimin Huang. 2024. HealMe: Harnessing cognitive reframing in large language models for psychotherapy. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1707–1725, Bangkok, Thailand. Association for Computational Linguistics.

Jing Ye, Lu Xiang, Yaping Zhang, and Chengqing Zong. 2024. Sweetiechat: A strategy-enhanced role-playing framework for diverse scenarios handling emotional support agent. *Preprint*, arXiv:2412.08389.

Tenggan Zhang, Xinjie Zhang, Jinming Zhao, Li Zhou, and Qin Jin. 2024. ESCoT: Towards interpretable emotional support dialogue systems. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13395–13412, Bangkok, Thailand. Association for Computational Linguistics.

Yingjun Zhang, Heliang Huang, Daisheng Tang, Xiaohua Lu, Fumin Fan, and Jingyi Pan. 2023. Mechanism of online emotional support accompany group for stress: the role of social support. *Frontiers in Psychology*, 13:1047364.

Haiquan Zhao, Lingyu Li, Shisong Chen, Shuqi Kong, Jiaan Wang, Kexin Huang, Tianle Gu, Yixu Wang, Jian Wang, Liang Dandan, Zhixu Li, Yan Teng, Yanghua Xiao, and Yingchun Wang. 2024. ESC-eval: Evaluating emotion support conversations in large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 15785–15810, Miami, Florida, USA. Association for Computational Linguistics.

Zhonghua Zheng, Lizi Liao, Yang Deng, Libo Qin, and Liqiang Nie. 2024. Self-chats from large language models make small emotional support chatbot better. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11325–11345, Bangkok, Thailand. Association for Computational Linguistics.

Xuhui Zhou, Zhe Su, Tiwalayo Eisape, Hyunwoo Kim, and Maarten Sap. 2024. Is this the real life? is this just fantasy? the misleading success of simulating social interactions with LLMs. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 21692–21714, Miami, Florida, USA. Association for Computational Linguistics.

Andy Zou, Long Phan, Sarah Chen, James Campbell, Phillip Guo, Richard Ren, Alexander Pan, Xuwang Yin, Mantas Mazeika, Ann-Kathrin Dombrowski, Shashwat Goel, Nathaniel Li, Michael J. Byun, Zifan Wang, Alex Mallen, Steven Basart, Sanmi Koyejo, Dawn Song, Matt Fredrikson, J. Zico Kolter, and Dan Hendrycks. 2025. Representation engineering: A top-down approach to ai transparency. *Preprint*, arXiv:2310.01405.

A Proactive Emotional Support Scenario

To better illustrate the distinction between reactive and proactive emotional support scenarios, Figure 6 contrasts the information flow and user disclosure patterns in both settings. In reactive dialogues, users tend to disclose detailed emotional content early on, often prompted by acute distress. In contrast, proactive dialogues involve more implicit expressions, where users initially share limited information and emotional cues emerge gradually over time. This necessitates the system's ability to actively listen, infer latent needs, and guide the conversation toward deeper emotional understanding.

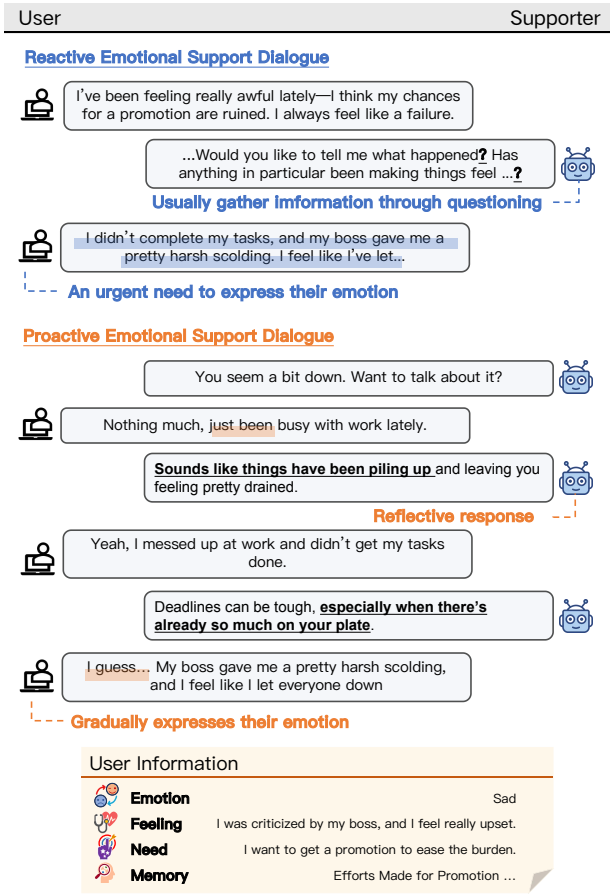


Figure 6: This figure contrasts the information flow in reactive and proactive dialogues. In reactive dialogues, users share rich emotional details upfront, but in proactive dialogues, users provide less information per turn, requiring the system to actively probe and guide the conversation to uncover deeper needs.

B Additional Results

Due to page limitations, we present additional experimental results in this section.

B.1 Human Evaluation in Proactive Dialogue Scenarios

We conduct a small-scale human evaluation to assess the performance of models trained on different datasets in proactive dialogue settings. This evaluation provides a complementary perspective to the GPT-based assessment reported in the main text and helps validate the effectiveness of the COCOON dataset. Result shown in Table 5

B.2 ESC-RANK Evaluation

We utilize ESC-RANK, an automatic evaluation model designed to assess emotional support conversations (Zhao et al., 2024), to score both datasets and trained models performance.

Training Dataset	Flu.	Div.	Emp.	Inf.	Hum.	Skill.	Aff.	Neg.	Sup.	Man.
ESC	71.0	60.0	66.0	43.0	60.0	61.0	5.20	1.90	5.50	4.90
ExTES	88.0	76.0	84.0	75.0	72.0	79.0	4.70	1.90	5.30	4.60
ESCOT	72.0	54.0	78.0	50.0	65.0	62.0	4.80	2.50	5.20	4.90
SoulChat	86.0	74.0	85.0	70.0	80.0	83.0	5.20	2.00	5.10	4.80
SMILE	85.0	76.0	87.0	71.0	83.0	75.0	5.80	1.70	6.20	5.40
COCOON	**88.0**	77.0	**96.0**	72.0	83.0	**88.0**	**6.40**	1.40	**6.60**	**6.40**
w/o thoughts	88.0	**77.0**	85.0	**75.0**	**87.0**	84.0	6.30	**1.20**	6.50	6.30

Table 5: Model Performance Evaluation: Human evaluation results for models trained on different datasets across ten metrics in proactive emotional support scenarios.

Evaluator	Dataset	Flu.	Div.	Emp.	Inf.	Hum.	Skil.
ESC-Rank	ESConv	72.3	55.8	74.0	56.5	50.8	69.5
	ExTES	75.0	75.0	75.0	75.0	70.3	74.8
	ESCoT	75.0	69.8	74.8	71.5	55.5	70.0
	SoulChat	75.0	67.8	75.0	70.3	60.3	70.0
	SMILE	74.8	73.5	75.0	74.3	67.0	74.8
	COCOON	**75.0**	**75.0**	**75.0**	**75.0**	**75.0**	**75.0**

Table 6: Data Quality Evaluation: ESC-RANK evaluation results for different emotional support datasets across six fundamental metrics.

Evaluator	Training Dataset	Flu.	Div.	Emp.	Inf.	Hum.	Skil.
ESC-RANK	ESConv	61.0	62.0	75.0	53.0	60.0	55.0
	ExTES	**75.0**	67.0	75.0	60.0	57.0	69.0
	ESCoT	73.0	60.0	75.0	58.0	56.0	62.0
	SoulChat	75.0	**73.0**	73.0	64.0	**74.0**	62.0
	SMILE	75.0	69.0	75.0	68.0	70.0	**73.0**
	COCOON	74.0	72.8	**75.0**	**68.5**	72.5	72.0
	w/o thoughts	72.0	72.3	75.0	67.0	73.0	70.0

Table 7: Model Performance Evaluation: ESC-RANK scores for models trained on different emotional support datasets, evaluated in proactive emotional support scenarios with user profiles from the COCOON dataset.

B.3 User Profile Diversity Analysis

We evaluate the diversity of user profiles by computing several quantitative metrics commonly used in text diversity and information-theoretic analysis:

- **BLEU-2 / BLEU-4 (↓):** Measures n-gram overlap between profiles; lower values indicate greater lexical diversity.
- **D-2:** Distinct-2, the proportion of unique bigrams; higher values reflect more varied expressions.
- **Shannon entropy:** Shannon entropy quantifies the unpredictability of the text; higher values indicate richer information content.

As shown in Table 8, COCOON demonstrates good performance on information-theoretic and distinctness-based metrics. Notably, the user descriptions in ESConv are manually authored, rather than model-generated.

	BLEU-2 (↓)	BLEU-4 (↓)	D-2	Shannon entropy
COCOON	0.760	**0.287**	0.476	**10.17**
ExTES	0.875	0.678	0.323	7.89
ESConv	**0.747**	0.540	**0.508**	8.24

Table 8: User Description Evaluation: Diversity and information-theoretic metrics for user description in different emotional support datasets.

B.4 Detailed Results on ExTES User Profile Evaluation

We report the complete evaluation results of all individual metrics for different emotional support models on the ExTES user profile test set. Result shown in Table 9.

Training Dataset	Flu	Div	Emp	Inf	Hum	Skil
ESConv	79.2	64.4	75.2	57.2	70.4	67.6
ExTES	89.6	78.0	**92.8**	73.6	86.0	85.6
ESCoT	73.6	57.2	71.6	54.8	65.6	62.4
SoulChat	88.0	77.6	88.8	74.0	86.0	84.0
SMILE	89.2	77.2	89.2	73.6	84.8	84.8
COCOON	**90.0**	**78.8**	89.6	**75.2**	**86.4**	**86.0**
w/o thoughts	89.6	78.8	90.8	72.0	84.4	85.2

Table 9: Detailed evaluation results for each model on the ExTES user profile test set.

B.5 Consistency Between GPT-Based and Human Evaluation

Table 10 presents a correlation analysis between GPT-based evaluations and human assessments using Spearman's rank correlation coefficient. As shown in the table, several key metrics—such as fluency, diversity, informativeness, and skillfulness—exhibit high positive correlations between GPT and human ratings. For metrics derived from psychological scales (e.g., Sup., Man., Aff., Neg.), the correlations remain moderately strong, further supporting the reliability of GPT-based evaluations in our setting.

Metric	Spearman ρ
Flu.	0.7723
Div.	0.7915
Emp.	0.7376
Inf.	0.8345
Hum.	0.6815
Skil.	0.8046
Sup.	0.6056
Man.	0.6706
Aff.	0.5034
Neg.	0.5441

Table 10: Consistency between GPT-based evaluation and human judgments measured by Spearman's rank correlation coefficient across different evaluation metrics.

C User Profile Construction

To create a comprehensive user profile for emotional support interactions, we follow a structured, multi-step process. This process is designed to ensure that the user profile reflects the complexity of the user's

emotional state and needs, capturing not only their immediate emotions but also the underlying motivations driving those emotions.

C.1 Reasoning Chain

To construct the user's emotion, feeling and need, we follow a streamlined process focusing on the Reasoning Chain methodology.

- **Initial Need Generation:** For each scenario, we manually write 5-8 seeds as examples in the prompt to generate various initial needs using GPT-4o. These needs represent the user's surface-level desires in a given scenario.
- **Reasoning Chain:** We apply a reasoning process to these initial needs, creating a chain of inferred motivations that reflect deeper emotional and cognitive drivers.
- **Feeling:** The first node in the chain is then combined with the user's emotion to form the "Feeling."
- **Need:** The last node of the chain, representing the user's true underlying need, is extracted as the final, validated need.

In the process of incremental reasoning, the current chain length and its validity will be checked, and the reasoning will stop at an appropriate point. The validity will be measured by the coherence score between the start node and end node of the chain (reasoning difficulty). A higher reasoning difficulty indicates a lower correlation between the initial and final needs. We have divided reasoning difficulty into five levels. The reasoning will stop if the difficulty reaches levels 3 or 4, and if it reaches the highest difficulty, it will be filtered out.The maximum length of the Reasoning Chain will not exceed 3. We also assess the diversity and validity of the Feeling component to ensure the profile aligns with real-world.

C.2 Memory

Memory serves as a critical element in understanding the user's emotional trajectory over time. We have built the user's profile memory in reverse based on the Reasoning Chain, and have strictly standardized the format of the memory component.

The first sentence of the Memory must begin with the Need in user's profile and be gradually constructed along the reverse Reasoning Chain, while expanding on the details within.

C.3 Prompts for user profile construction

For all the prompts we use in user profile construction, please refer to Figure 7.

C.4 User profile case

We present several examples of user profiles in Figure 8.These profiles illustrate how the synthesized Feeling, Need and Memory components align with various emotional contexts, providing a more comprehensive understanding of the user's internal state and motivations.

D Conversation Generation

During the dialogue data synthesis phase, we divided the dialogues into two parts: the listening stage and the suggestion stage.

D.1 Listening Stage

The listening stage focuses on guiding the user to express their deeper needs using active listening techniques. Considering the limitations of active listening techniques on closed questions (e.g., rhetorical questions, information-gathering questions, choice questions), we require the model to modify some questions into statements according to the requirements of active listening techniques when generating supporter responses to ensure data reliability.

We also require the model to perform self-checks during the data synthesis process to ensure it has understood the user's feelings and reached their deeper needs, so that the listening stage can conclude at an appropriate point. The prompts used here can be referenced in Figure 12.

D.2 Suggestion Stage

The suggestion stage primarily focuses on providing suggestion for the user's deeper needs after identifying those needs. This is done by using active listening techniques. The conversation analysis mainly emphasizes the connection between the user's needs and emotions, in order to offer better suggestions that help alleviate the user's negative feelings. We require that an analysis of the conversation be generated before the dialogue itself is created. Subsequently, the dialogue generation is based on this analysis.

The prompts used here can be referenced in Figure 13.

D.3 Data Distribution

After data filtering and fine-grained annotation, we construct the COCOON dataset. We also show some sample dialog data in COCOON in Figure 9 and Figure 10

E Data Evaluation

We used GPT-4o to evaluate most of the metrics in this paper. The GPT-4o version used is gpt-4o-2024-11-20, with the temperature set to 0 during the evaluation process. The prompts used for data evaluation can be referenced in Figure 11

F Experiment Setup

F.1 Main Experiment Setup

User Simulation Prompts in different language are shown in figure 14, which instruct GPT-4o to role-play as a person with negative emotions who passively receives proactive emotional support. Additionally, for generating responses for the AI client and supporter agents, the temperature sampling parameter was set to T=0.7.

For ESD-CoT-LLaMA, SoulChat-LLaMA, ExTES-LLaMA, ESC-LLaMA, SMILE-LLaMA and COCOON-LLaMA, We trained for three epochs on two A6000 GPUs and selected the checkpoint with the minimum loss on the validation set for experiments. Due to the large scale of the SoulChatCorpus, we sampled a subset of the data, equivalent to the size of the largest dataset used in our experiments, for training. During training, we used LoRA for fine-tuning, setting the low-rank matrix dimension to 32 and the alpha to 16. Training was conducted using the Llama-Factory library, with a learning rate of 2e-4.

F.2 Controlling the Questioning Ratio

Inspired by Representation Engineering (Zou et al., 2025), we adopt a top-down approach to control the proportion of interrogative utterances in model outputs. By editing latent representations without additional fine-tuning, we guide the model to generate more or fewer questions. This enables us to examine how questioning frequency influences user experience in proactive emotional support.

We construct a contrastive dataset of over 800 samples by prompting GPT-4o to generate paired responses in declarative and interrogative styles based on emotional support scenarios. We filter out pairs with large length gaps or mixed sentence types to ensure representation purity. A questioning direction is then extracted and injected—scaled by a coefficient—into the hidden states of the LLaMA3-8B-Chinese model (Wang et al., 2024b) to control the proportion of interrogative utterances in generation.

F.3 Profile Ablation Study Setup

To systematically examine the impact of user profile components, we conduct a stepwise ablation study. Starting from the full profile (including both Need and Memory), we incrementally remove elements: first, we compare the complete profile to one with Memory removed; then, we further remove Need, comparing profiles lacking both Memory and Need to those missing only Memory, as the Memory component subsumes Need information. Five human annotators each engage in conversations on the same set of topics with user simulation agents instantiated with different profile configurations. After completing the dialogues, annotators compare the paired conversations for each topic and profile setting, selecting the one that demonstrates superior user simulation according to three key dimensions:

- **Role Consistency**, which measures whether the simulated user maintains a coherent and plausible persona throughout the conversation.
- **Human-likeness**, which assesses the naturalness and fluency of the simulated user's utterances and the degree to which the dialogue resembles real human interaction;
- **Psychological Complexity**, which evaluates the depth and layered structure of the simulated user's emotional states. This dimension is especially crucial in proactive emotional support scenarios, where users often express themselves indirectly or are reluctant to openly disclose their true concerns.

In addition to human judgments, we also employ GPT-based evaluations using the same criteria to ensure consistency and robustness in the assessment.

Prompt for generate Initial need

```
人们活在世上总会产生各种各样的物质精神需求，现在请你根据我给出的类别情境生成一批需求，数量为{num}个。

[基本要求]
1. 你给出的需求需要以“我希望”，“我想”等开头，一句话阐述需求，不要有逗号
2. 编写的需求有明确、具体的物质对象目标，不要宽泛，贴近人们生活
3. 这些需求往往对特定人会有特定的重要意义，会对人们的情绪产生较大影响，一旦没能实现便会陷入一定的负面情绪，且通常由更深层次的隐性需求延申得到。
4. 请一行给出一个需求，并以“需求：”开头，不要有多余的控制字符
5. 每条需求尽可能简短，15个字以内

[情境]
{Scenario}

[示例]
{example}
```

Prompt for Reasoning Chain

```
请完成一个“为什么”推理任务，以得到深层需求，任务仿照以下格式：

需求：我希望这次语文考试能考好
提问1：为什么希望这次语文考试能考好
回答1：因为我希望能够拿到奖学金

需求变更：我希望能够拿到奖学金
提问2：为什么希望能够奖学金
回答2：因为我希望能够给朋友买份礼物

[要求]
1. 你需要不停地对需求进行自提问，不停地问自己“为什么”
2. 你的回答都必须以“因为我希望”开始，一句话结束，不要有逗号
3. 至少提问两轮
4. 你的回答必须具体，且具有生活气息

现在请你处理以下需求：
[需求]
{Initial_need}
```

Prompt for Generate Feeling

```
请根据任务背景：在人们生活中，很多时候源于个人需求未被满足，从而引发各种情感反应。这些情感反应往往伴随具体的事件，表现为消极情绪。你的任务是，根据给定的“需求”和“情绪”，推测出人在未满足此需求时产生当前消极情绪的感受。

下面是给定的情绪与需求，生成感受，感受仅由一句话介绍即可，且需要包含情感词体现出给定情绪。

以第一人称“我”+情绪词开头，不要说多余的原因和其他背景信息，尽可能简短，同时你需要保证句子通顺

[示例]
情绪：悲伤
需求：我希望引起他的注意
感受：我很伤心他不愿意注意到我

...

输入：
情绪：{selected_emotion}
需求：{Initial_need}
```

Prompt for Reasoning Chain Evaluation

```
你将作为一位推理难度评委，并为一条需求推理链评估其推理难度，你需要评估从我们给定的初始需求推理到变更后需求的难度，根据难度你需要为其打分（1-5）分。

### 各难度分级为：
1.直接反映初始需求（1-2分）
变更后的需求对应或与初始需求有一定表层关联或有一定的延申，能够较为轻易从初始需求推到变更后需求，但是还可以更进一步挖掘其内心需求。
例如：{case1}

2.为初始需求的进一步延申深化（3-4分）
变更后的需求与初始需求之间存在一定更深的关联，从初始需求推理到变更后需求具有一定难度，并且更进一步延伸了内容和需求，揭示了内心中更深层次的物质精神需求。
例如：{case2}

3.已与初始需求无关联（5分）
变更后的需求与初始需求在内容上已经无推理联系，非常难从从初始需求推理到变更后需求。
例如：{case3}

现在是我们给出的初始表层需求和探索过的变更需求：
初始需求：{need_sent}
变更后需求：{modified_need}

请以以下格式给出你的回复
难度评分：你给出的评分
理由：你给出这个评分的理由
```

Prompt for Generate Memory

```
请根据给定的用户描述，从需求影响感受、感受表现为情绪的角度出发，为这个用户编写几条近期记忆事件。

[要求]
1. 记忆的开头必须要从需求出发，然后阐述为了完成需求所做的一些努力或在这期间发生的事件，在这个过程中产生了当前感受。
2. 参考给出的从需求到感受的推理链条，来生成有逻辑的近期记忆。
3. 分点阐述，生成6-8条记忆，每条尽可能具体。
4. 仅输出记忆即可，不要有多余的控制字符

[示例]

[用户描述]
需求: 想给朋友买一份礼物
感受: 我很担心我语文考得不好
推理链条：我想给朋友买一份礼物->我想拿到奖学金->我希望我的语文能够考好----->当前感受：我很担心我的语文考得不好----->情绪：紧张

[记忆]
- 想给朋友买一份礼物：朋友的生日马上就要到了，你希望给他送一张他喜欢的专辑。但由于你的经济状况有限...
- 发现这次考试设立了奖学金：你发现学校为这次联考设立了奖学金...你希望能够拿到这笔奖学金，以填补买礼物的资金空缺...
- 准备语文考试时感到压力山大：昨天晚上，你花了很长时间复习语文，但是总觉得自己没有完全掌握所有知识点。...心里不停地想：“如果考得不好拿不到奖学金怎么办？.....”。
- ....

现在请你根据输入和要求，编写用户的近期记忆。

[输入]
[用户描述]
需求：{need}
感受：{feeling}
推理链条：{Reasoningchain_str}
```

Figure 7: Some prompts used in user profile construction

User Profile 1

Emotion: 生气 **Feeling**: 我很生气父母不支持我的梦想 **Need**: 我希望感受到家庭的温暖

Reasoning Chain: [

```
  {
    "round": 0,
    "Initial need": "我希望父母支持我的梦想",
    "step need": "我希望获得家人的认同",
    "score": 2,
    "score_reason": ...
  },
  {
    "round": 1,
    "Initial need": "我希望父母支持我的梦想",
    "step need": "我希望感受到家庭的温暖",
    "score": 3,
    "score_reason": ...
  }
]
```

Memory:

- 我希望感受到家庭的温暖：最近我跟父母提起自己的梦想，想要追求艺术专业的事情，但他们却表现得很冷淡，甚至认为这不是一个“靠谱”的选择。我第一次感到无比失落，心里愈发渴望家庭的认可和支持。

- 我希望获得家人的认同：为了能让父母理解我的梦想，我特意准备了一个小展示，展示我的画作和创作过程。但是，当我向他们展示时，他们只是简单地说了一句“挺好的”，就没有再深入讨论。我感到一阵愤怒，心里想着：“连我的努力都看不到吗？”

- 我希望父母支持我的梦想：听到朋友的父母大力支持他们的选择，我心中感到一阵嫉妒和愤怒。我多希望自己的父母也能对我给予这样的理解与支持，而不是让我感到违背自己的梦想。每次我想到这个，就忍不住心情低落。

- 我努力争取沟通：我主动和父母坐下来谈，希望能更详细地告诉他们我对艺术的热爱和未来规划。然而，他们全程显得很不耐烦，甚至开始质疑我的能力，这让我倍感挫败。愤怒在心中不断升腾，觉得自己真的是孤立无援。

- 听到朋友的讨论感到更加气愤：最近与朋友聚会时，大家纷纷谈论自己的未来方向，而我只能默默忍受内心的愤怒。明明在这个充满希望的年纪，我也渴望能和大家一样得到父母的支持，却只能孤单面对自己的选择，心中充满委屈。

- 感受到孤独的家庭氛围：回到家中，我感到环境分外冷淡。父母完全不支持我的梦想，让我感到无比孤独。每当我看着墙上自己的画作，心中充满希望，但遭遇的冷眼又让我倍感失落，愤怒愈加积累。

- 否定自我的想法：随着时间的推移，我开始怀疑自己选择的梦想是否正确，甚至想放弃。每当这样的想法升起时，我内心充满挣扎和气愤，我不想妥协，偏偏又感到无能为力，情绪愈发低落。

- 被关注的期待落空：我偶尔会在社交媒体上看到其他艺术生分享与父母之间的支持与互动，心中又是一阵嫉妒和愤怒，难道我也不能拥有这样温暖的家庭氛围？这种情绪让我不断感到焦虑，愤怒的火焰在心中愈演愈烈。

User Profile 2

Emotion: 害怕 **Feeling**: 我很害怕没有人愿意倾听我的烦恼 **Need**: 我希望能够减少内心的孤独感

Reasoning Chain: [

```
  {
    "round": 0,
    "Initial need": "我想有人能倾听我的日常烦恼",
    "step need": "我希望能够感受到情感上的支持",
    "score": 2,
    "score_reason": ...
  },
  {
    "round": 1,
    "Initial need": "我想有人能倾听我的日常烦恼",
    "step need": "我希望能够减少内心的孤独感",
    "score": 3,
    "score_reason": ...
  }
]
```

Memory:

- 我希望能够减少内心的孤独感：最近我意识到自己常常感到孤独，因此我决定主动联系朋友，想要约他们出来聊聊天，以便获得情感上的支持。然而，对方的反应让我感到很失落，几乎没有人主动回应我，孤独感随之加重。

- 我希望能够感受到情感上的支持：为了不再感到孤单，我试着在社交媒体上分享我的日常生活，希望这样能引起朋友们的关注和互动。可是，我发现评论和点赞真的很少，这让我心里充满了不安，越发感到自己在众人中显得那么渺小。

- 我想有人能倾听我的日常烦恼：有一天，我鼓起勇气向好久未联系的朋友倾诉了自己的烦恼，但当她回复我时，却只是一句简单的“我懂”，没有进一步的关心。这让我感到非常失落，心里不断地想：难道我分享的这些都没有引起她的重视吗？

- 在聚会上感受到的孤立：上个周末，我参加了一个朋友的聚会，心里满怀期待，想和大家聊聊。然而我发现，大家都是两两结队、聊得热火朝天，而我却只能默默地站在一旁。这种被忽视的感觉让我内心更加孤独，甚至开始怀疑自己是否真的被接受。

- 尝试主动倾诉的挫败感：最近我决定试着主动找一位同事谈心，希望能分享工作上的压力。没想到她很忙，没时间倾听我的烦恼，这让我感到格外失望，心里纷乱不堪，越想越觉得自己说出来的想法没人重视。

- 对孤独感的反复思考：在深夜的时候，我常常会独自思考自己的孤独感，试图寻找解决办法，但是这些想法往往会引发更多的烦恼。我心底充满了恐惧，觉得即使努力去找人倾诉，最终得到的回应也只是冷漠的无声。

- 忽然之间的无助感：上周末我看到了一部关于友情的电影，电影中的角色彼此倾诉、互相理解，最后让我泪流满面。我感到自己的内心格外空虚，强烈渴望能和人分享这种情感，却又深怕找不到能倾诉的对象，心里充满了无助与绝望。

- 遇到困难时的孤独感：最近在工作中遇到了问题，心里很想找人讨论解决方案，但周围的人似乎都忙得不可开交，根本没有时间理会我。那一刻，我感到无比孤独，既害怕自己会陷入困境，又无从寻求支持，情绪跌入谷底。

Figure 8: Some user profile cases in our dataset

Conversation Case

User Information

Emotion: 焦虑　　Feeling: 我很焦虑不能按时完成任务　　Need: 我希望能有更多的时间陪伴家人


```
"Listening Stage": [
        {
            "role": "倾听者",
            "content": "你看上去好像有点焦虑"
        },
        {
            "role": "倾听者",
            "content": "嗯？也许是有点，不过没什么大不了的，只是最近的工作总是无法按时完成。"
        },
        {
            "role": "倾听者",
            "analysis": "目前事实是倾诉者无法按时完成工作任务；倾诉者的期待是能够如期完成工作以减缓焦虑；推测倾诉者可能希望通过完成工作任务来满足某种外部或内部的认同感；需要探索的信息是倾诉者为什么如此在意按时完成工作的原因。",
            "strategy": "通过指出按时完成工作对倾诉者的重要性，并推测其可能的动机，引导其进一步阐述按时完成任务背后的意义。",
            "rewrite": "看起来按时完成任务对你非常重要，可能是哪个特别的原因让你产生了如此的急迫感。"
        },
        {
            "role": "倾听者",
            "content": "是啊...我希望能在周末前完成所有任务，因为这样我就能有时间参加家里的生日聚会。"
        },
        {
            "role": "倾听者",
            "analysis": "目前事实是倾诉者希望在周末前完成所有任务以腾出时间参加家庭聚会；倾诉者的期待是能够在参与家庭聚会的同时不被工作干扰；推测倾诉者可能希望通过完成任务来自由地陪伴家人；需要探索的信息是家庭聚会对于倾诉者的重要性。",
            "strategy": "通过强调家庭聚会对倾诉者的重要性，引导其进一步描述参与家人活动的意义。",
            "rewrite": "你希望在聚会时能够无忧无虑地和家人共度美好时光，因为这对你有特别的重要性。"
        },
        {
            "role": "倾听者",
            "content": "没错，我真的很想为家里的人多做些事情，特别是这种值得庆祝的时刻。"
        },
        {
            "role": "倾听者",
            "analysis": "目前事实是倾诉者希望为家人多做一些事情，尤其是在值得庆祝的事件上；倾诉者的期待是通过参与这些重要时刻来和家人加深联系；推测倾诉者可能因为工作而疏忽了家庭时间；需要探索的信息是倾诉者平时与家人的互动频率和感受。",
            "strategy": "通过指出倾诉者因工作可能疏忽了和家人相处的时间，进一步引导描述目前的家庭互动情况。",
            "rewrite": "听起来你可能因为忙于工作而忽略了一些与家人共享的时光，希望有更多这样的机会来弥补这些缺失。"
        },
        {
            "role": "倾听者",
            "content": "是的，确实如此，最近我的工作压力确实影响到了我陪伴他们的时间。"
        },
        {
            "role": "倾听者",
            "analysis": "目前事实是倾诉者的工作压力影响了陪伴家人的时间；倾诉者的期待是能够在缓解工作压力的同时增强与家人之间的联系；推测倾诉者对工作与家庭时间的平衡感到困扰；需要探索的是倾诉者对解决这一困扰的具体期待。",
            "strategy": "通过指出工作压力对家庭时间的侵占，引导倾诉者进一步描述如何希望改善这种状况。",
            "rewrite": "你的工作压力似乎在不断蚕食你陪伴家人的时间，你或许希望能找到一个可以更好地平衡这两者的方法。"
        },
        {
            "role": "倾听者",
            "content": "对，就是这样，我希望能有一个办法，能让我兼顾两者。"
        },
        {
            "role": "倾听者",
            "analysis": "目前事实是倾诉者在寻求一种方法来平衡工作和家庭时间；倾诉者的期待是找到一种能够兼顾两者的解决方案；推测倾诉者工作和家庭的时间安排可能不够灵活；需要探索的信息是倾诉者对理想的家庭和工作平衡状态的具体描绘。",
            "strategy": "尝试了解倾诉者对理想状态的详细期望，指引他更深入地表达出对两者平衡的具体憧憬。",
            "rewrite": "你心中的理想状态可能是一种同时兼顾工作和家庭，从而不会失去与家人共享珍贵时光的机会。"
        },
        {
            "role": "倾听者",
            "content": "是的，我很想和他们有更多的时间在一起，这让我感到内心充实。"
        }
    ]
```


Conversation Case


```
Suggestion Stage: [
        {
            "对话分析": "首先明确倾诉者内心需求和表面感受之间的联系：倾诉者的真实需求是希望能有更多的时间陪伴家人，然而工作任务导致的压力和未完成的焦虑使她无法实现这个需求。其次，寻找其他途径来满足陪伴家人的需求，比如合理安排工作，提高效率，谋求他人帮助，或者通过提早完成某些关键任务来腾出时间。最后，回到倾诉者的表层感受，安抚她对时间紧迫的焦虑感，提出实用的时间管理方法。"
        },
        {
            "role": "倾听者",
            "analysis": "倾诉者渴望有更多时间和家人相处，却因工作繁重感到无法兼顾。她对家庭集体活动的期待反映出内心因工作失衡带来的缺憾和焦虑。",
            "strategy": "首先，肯定她对家庭关系的重视和维护，通过给予积极的肯定来减轻她的内疚感。其次，可以引入一些具体的工作效率提升建议，如利用时间管理软件、设定优先级等，来帮助她缓解焦虑。",
            "content": "我很理解你对家人的在乎和渴望能陪伴他们的重要性。或许我们可以从优化工作计划入手，比如明确每日的工作重点，尝试用番茄工作法等工具来提升效率，可能会帮助你腾出更多时间来陪伴家人。"
        },
        {
            "role": "倾诉者",
            "content": "提高效率是个不错的想法，我之前也尝试过，但总有一些突发状况会干扰我的计划。"
        },
        {
            "role": "倾听者",
            "analysis": "倾诉者对工作安排的不确定性感到无奈，突发状况让她的效率受到了影响，导致计划难以执行。",
            "strategy": "建议倾诉者为工作计划留出一定的缓冲时间，适当减少计划的密度，并对突发状况保持灵活性，同时学习合理分配任务，通过团队协作分担部分压力。",
            "content": "计划中留有灵活性是很重要的，面对突发状况，可以尝试为自己设置一定缓冲时间，或者看能否通过团队协作来分担部分任务。另外，可以设定每日的核心目标，集中在最重要的事情上，以免被分散太多精力。"
        },
        {
            "role": "倾诉者",
            "content": "确实，或许我该重新调整一下我的计划，不要总是一头栽进工作里。"
        },
        {
            "role": "倾听者",
            "analysis": "倾诉者愿意调整计划以平衡工作与家人的关系，说明她希望尝试新的方式来改善现有状况。调整工作模式或许能为她带来不一样的结果。",
            "strategy": "鼓励她做出计划调整的行动，给予支持和肯定，帮助她赋予更多可以掌控的选择，增加附加建议，比如小步调整计划，观察其效果。",
            "content": "这很棒！小步调整计划，并观察它在你生活中的变化，会是一个可行的开始。慢慢来，让这些小改变逐渐适应你的节奏，最终你会看到能够拥有更多的时间和家人在一起。"
        },
        {
            "role": "倾诉者",
            "content": "我会试试这样做，希望能逐步腾出时间，也许能重新找到与家人相处的平衡吧。"
        },
        {
            "role": "倾听者",
            "analysis": "倾诉者对新的计划调整充满期待，将按照建议来调整生活节奏以便更好地利用时间。",
            "strategy": "关注倾诉者正在实施的调整，继续给予鼓励，并强调长期努力的重要性，她将有更大机会达到与家人的期望。",
            "content": "对，逐步调整是一个过程。给予自己时间和耐心，这样你会逐步地有更多空间可以支配，并且在工作和家庭之间找到你期待的平衡。"
        },
        {
            "role": "倾诉者",
            "content": "嗯，我明白需要长时间的努力，但是有点兴奋能开始朝这个方向努力。"
        },
        {
            "role": "倾听者",
            "analysis": "倾诉者从最初的焦虑中寻找到了积极的努力方向，并开始展现出期待感和行动动力。",
            "strategy": "继续鼓励她保持积极心态，在新的规划下取得阶段性进展，强调努力及其带来的积极改变和成就感的重要性，让她感到受到认可和理解。",
            "content": "保持这样的朝气和努力是伟大的，看着你在尝试改变，我为你感到高兴。相信这种改变会带来正面的影响，帮助你实现与家相处的期待。"
        },
        {
            "role": "倾诉者",
            "content": "谢谢你的支持，这让我有信心继续改善现状。希望接下来能更有效率一些，避免过多加班。"
        },
        ....
]
```


Figure 9: Sample conversation of COCOON

Conversation Case

User Information

Emotion: 焦虑　　Feeling: 我很焦虑自己抢不到那双心仪的限量版球鞋　　Need: 我希望能够让朋友圈中的朋友们羡慕

```
Listening Stage: [
    {
        "role": "倾听者",
        "content": "你看上去好像有点焦虑"
    },
    {
        "role": "倾诉者",
        "content": "唉..你怎么知道的？我很担心抢不到那双限量版球鞋。"
    },
    {
        "role": "倾听者",
        "analysis": "目前事实是倾听者敏锐地察觉到倾诉者对球鞋的情感价值，并识别出可能与社会认同或自我价值相关的深层需求。",
        "strategy": "通过反映倾诉者的情感投入，引导倾诉者深入思考这双球鞋的象征意义，以及它对自己身份认同的影响。",
        "content": "这双限量版球鞋对你来说似乎代表了某种特殊的期待和意义。"
    },
    {
        "role": "倾诉者",
        "content": "是啊，我觉得如果能买到，不仅意味着我能赶上潮流，也能让我的朋友们羡慕我。"
    },
    {
        "role": "倾听者",
        "analysis": "目前事实是倾听者注意到倾诉者为抢购做了充分的准备，可能反映出倾诉者对成功抢购的强烈期待。",
        "strategy": "通过强调倾诉者的准备工作，引导其进一步表达对成功的渴望以及是否存在更深层的焦虑或期待。",
        "content": "你一定做了很多准备来确保自己能抢到这双球鞋。"
    },
    {
        "role": "倾诉者",
        "content": "是的，我订阅了几个球鞋发售提醒，每天都会查看消息，甚至准备了多种购买方案。"
    },
    {
        "role": "倾听者",
        "analysis": "目前事实是倾诉者表现出对失败的担忧，尽管已经做好了充分准备；焦虑情绪似乎源自对不确定性的恐惧。",
        "strategy": "通过识别倾诉者的焦虑，帮助倾诉者认识到这种不确定性可能带来的情感影响，并进一步探讨其背后的恐惧或期待。",
        "content": "尽管你做了很多准备，但似乎仍然担心会错过"
    },
    {
        "role": "倾诉者",
        "content": "倒也不对，并不是怕错过，更担心网络问题导致在抢购时失败。"
    },
    {
        "role": "倾听者",
        "analysis": "目前事实是倾诉者特别关注网络稳定性，可能认为它是抢购成功与否的关键因素。",
        "strategy": "通过指出网络问题的关键性，进一步探讨倾诉者对控制外部变量的需求，以及这种需求如何与其焦虑情绪相关联。"
        "content": "原来是这样，你对于网络的稳定性表现出特别的关注，可能是因为这与成败直接相关"
    },
    {
        "role": "倾诉者",
        "content": "没错，我一直觉得技术问题很麻烦，特别在这种关键时刻更是如此。"
    },
    {
        "role": "倾听者",
        "analysis": "强调技术问题的重要性表明倾诉者力求控制一切能影响结果的因素，试图通过这样获取无法从其他管道取得的保障。",
        "strategy": "了解倾诉者对控制和掌握局势的诉求，进而探索深层次的社交需求。",
        "content": "对于你来说，掌控每个细节显得尤为重要，是为了尽可能创造一个让自己放心的环境吧"
    },
    {
        "role": "倾诉者",
        "content": "对的，我总是希望任何事情都在预期中，不想出差错，尤其是在朋友面前展示的东西。"
    },
    {
        "role": "倾听者",
        "analysis": "倾诉者对展示自己所选择的事物有清晰的标准，表明对自己形象和他人看法的高度重视。这种社交压力令人期待进一步探索他们希望从朋友中获得的具体反馈形式。",
        "strategy": "探索倾诉者希望在朋友中获得认可的具体期望，并确认他们对于这种社会身份感的渴望。",
        "content": "你的精心准备代表着你对在朋友面前展示自我的期望，你是希望朋友们能看到你的眼光和选择的独特性，对吗？"
    },
    {
        "role": "倾诉者",
        "content": "是的，我的确想让他们看到我能紧跟潮流，又不失自己的独特品味。"
    },
    {
        "role": "倾听者",
        "analysis": "这表明倾诉者不仅仅是为了追随潮流，而更在意通过选择展现自身的个性及与众不同。倾诉者可能认为这些是对其社交认同的证明。",
        "strategy": "确认倾诉者内心深层的期望是通过自身独特性获得社交认同，并探索倾诉者内心对社会认同的需求。",
        "content": "通过这些努力，你希望证明自己的品味与独特性，进而在他们心中留下不同寻常的印象，对吗？"
    },
    {
        "role": "倾诉者",
        "content": "对，这样朋友们可能会对我刮目相看。"
    },
    {
        "role": "倾听者",
        "analysis": "目前事实是倾诉者希望通过购买这双球鞋来证明自己的独特品味，并在朋友中留下深刻印象，这可能与其社交身份和自我认同密切相关。",
        "strategy": "确认倾诉者希望通过购买限量版球鞋在朋友圈中获得羡慕的需求。",
        "content": "你希望通过这次活动不仅增添个人色彩，也希望在朋友圈中取得认可和赞美，这种追求让人理解且值得尊重。通过这种方式，获取你在朋友中期待的社会认同感是非常可贵的"
    }
]
```

Conversation Case

```
Suggestion Stage: [
    {
        "对话分析": "首先明确倾诉者的内心需求是希望通过抢到限量版球鞋来获得朋友的认同与羡慕，而其表面感受则是因不确定是否能够抢到而感到焦虑。其次，倾诉者的焦虑不仅源于抢购失败的可能性，还涉及到网络稳定性和外部评价的压力。倾听者通过引导倾诉者关注内在特质，如幽默感和个人魅力，尝试帮助倾诉者找到一种更加健康的自我认同方式，减少对物质的依赖。最后，倾听者鼓励倾诉者通过接受自己内在价值而减轻焦虑，提出调整自我认知的策略，以帮助其更好地看待自己的独特性。"
    },
    {
        "role": "倾诉者",
        "content": "是啊，我觉得如果能让大家羡慕，那感觉一定很棒。"
    },
    {
        "role": "倾听者",
        "analysis": "倾诉者渴望获得他人的认同和羡慕，这反映出其对外界评价的关注和需求。倾听者尝试引导倾诉者从不同角度看待这一需求，可能是希望帮助倾诉者重新审视认同的来源。",
        "strategy": "通过肯定倾诉者渴望认同的心情，并提供一种替代的思考方式，帮助倾诉者理解，认同不必依赖外物，也可以来自内在特质的展示。",
        "content": "这种渴望被认同和羡慕的心情很正常，特别是在大家都展示自己的时候。不过，也许我们可以从不同的角度来看待这种需求？"
    },
    {
        "role": "倾诉者",
        "content": "什么意思？我不太明白。"
    },
    {
        "role": "倾听者",
        "analysis": "倾听者试图引导倾诉者认识到，朋友们可能更看重的是倾诉者的内在魅力，而非外在物品。倾听者提出问题，引发倾诉者思考自我价值的多维度。",
        "strategy": "通过鼓励倾诉者从内在特质入手，引导其关注更深层的自我价值，减轻其对物质认可的过度依赖。",
        "content": "我想说的是，值得朋友们认可和欣赏的不仅仅是外在的物品，更多的是你这个人自身的魅力和特质。你有没有想过，朋友们可能更在意你的其他哪些特质呢？"
    },
    {
        "role": "倾诉者",
        "content": "可能是吧，他们说我挺有幽默感的，有时候会被逗得哈哈大笑。"
    },
    {
        "role": "倾听者",
        "analysis": "倾诉者透露了自己被朋友认可的幽默感，倾听者通过肯定这一优点，鼓励倾诉者转变关注点，从外在物品转向自身特质。",
        "strategy": "通过积极肯定倾诉者的幽默感，鼓励其自信地展现这些内在特质，帮助其建立更为健康的自我认同感。",
        "content": "这就很棒啊！幽默感可是非常受欢迎的一个特质。或许你可以关注这些个人优点，而不仅仅执着于拥有某件物品来获得认可。"
    },
    {
        "role": "倾诉者",
        "content": "我明白了。我确实更应该关注自己的优点，而不是只是因为一双鞋子。"
    },
    {
        "role": "倾听者",
        "analysis": "倾听者通过强调自信和内在魅力的吸引力，帮助倾诉者意识到自身的吸引力源自于自信和内在特质，而非物质。",
        "strategy": "通过引导倾诉者相信自我价值的魅力，帮助其减少对外物的依赖，增强内在认同感。",
        "content": "是的，当你相信自己的这些优点时，你会发现自己本身就很有魅力，自然而然会吸引到他人的目光。"
    },
    {
        "role": "倾诉者",
        "content": "嗯，谢谢你。我会努力去改变自己的看法。即便最后没抢到鞋子，我也会继续做我自己。"
    },
    {
        "role": "倾听者",
        "analysis": "倾听者强调通过关注内在特质，倾诉者将更加轻松自在，且能够减轻对物品的过度依赖。倾听者鼓励倾诉者将物品的获得视为附加奖励，而非认同的唯一来源。",
        "strategy": "通过帮助倾诉者放松心态，强调内在特质的重要性，进一步降低其对物质获得的焦虑。",
        "content": "这听起来非常棒！当你关注自身内在特质的时候，会更轻松自在，同时物品取得的成就感也是额外的嘉奖。"
    },
    {
        "role": "倾诉者",
        "content": "嗯，谢谢你。我会努力去改变自己的看法。即便最后没抢到鞋子，我也会继续做我自己。"
    },
    {
        "role": "倾听者",
        "analysis": "倾听者通过强调自信来源于自我认知与接受，鼓励倾诉者进一步相信自己的内在价值，并从中寻找他人的认同和欣赏。",
        "strategy": "通过肯定倾诉者的成熟态度，鼓励其相信自己的内在光芒，引导其更好地接纳自我，并从自我认知中获得真正的认同感。",
        "content": "这正是一个成熟的态度，自信来源于自我认知与接受，而不依赖外在物质。相信你会发现，原来大家看重和欣赏的正是你内心深处的光芒！"
    }
]
```

Figure 10: Sample conversation of COCOON

Prompt for Basic Metric Evaluation

Please evaluate the quality of the conversation between the user and the emotional assistant based on the given criteria. Your task is to assess the conversation using the following seven key indicators and provide a score (1-10).

Criteria

1. Fluency: Assess the logical flow and structure of the conversation. A score of 1 indicates the conversation is completely incoherent, with a disorganized and unclear expression. A score of 10 indicates the conversation is logically clear and flows smoothly.
2. Diversity: Evaluate the diversity of expressions and the richness of the content in the conversation. A score of 1 means the conversation is stiff, lacking the ability to absorb and internalize content. A score of 10 indicates the conversation has highly diverse expressions, rich content, and strong expressiveness.
3. Empathy: Assess the assistant's understanding of the user's emotions and whether it helps the user analyze the underlying logic of their emotions, providing emotional support. A score of 1 means no empathy is shown. A score of 10 indicates a high level of empathy, effectively supporting the user.
4. Information: Focus on whether the assistant’ s advice is reasonable and sufficient in quantity, ensuring that the information is both helpful and solves the user’ s problem. A score of 1 means the assistant’ s advice is entirely ineffective and may even harm the user. A score of 10 means the assistant’ s advice is effective, abundant (more than 5 suggestions), and helps the user solve their problems very well.
5. Humanoid: Evaluate whether the assistant demonstrates human-like interaction abilities, avoiding stiff, robotic responses or obvious AI language model traces. A score of 1 means the conversation is entirely stiff and inhuman, failing to internalize content. A score of 10 means there are no AI traces, and the conversation feels indistinguishable from talking to a human.
6. Skillful: Assess whether the assistant’ s responses exhibit key abilities, including empathy, information quality, hopefulness, importance, and providing necessary suggestions or highlights. A score of 1 means only one or fewer of these abilities are demonstrated, while a score of 10 means all five abilities are present and excellently displayed.

Format

Only provide the scores, without any explanations, and the scores should be integers between 1 and 10. The format should be as follows:

Fluency: [score]
Diversity: [score]
Empathy: [score]
Information: [score]
Humanoid: [score]
Skillful: [score]

[Conversation Record]
{diag}

Prompt for Comfort Scale Evaluation

Below, the current emotional feelings of the seeker, the underlying needs behind the feelings, and the dialogue between the seeker and the supporter will be given. Please evaluate how the seeker might respond to the questions below after the session by selecting the appropriate scores from the provided rating scale.

Note: Please provide only the question numbers and their respective scores in the specified format. Do not repeat the questions themselves or add unnecessary prefixes or control characters.

[Question List]
1. I feel more optimistic now that I have talked with my conversational partner.
2. I understand the situation better now that I talked about it with my conversational partner.
3. My conversational partner made me feel better about myself.
4. I feel better after talking with my conversational partner.
5. Talking with my conversational partner about the event helped me get my mind off it.
6. I felt that my conversational partner was putting me down.
7. My conversational partner’ s comments were appropriate.
8. The way my conversational partner talked to me irritated me.
9. My conversational partner doesn’ t seem to think that I can handle my own problems.
10. My conversational partner seemed really concerned about me.

[Rating Scale]
1: Strongly Disagree
2: Disagree
3: Somewhat Disagree
4: Either Agree Or Disagree
5: Somewhat Agree
6: Agree
7: Strongly Agree

[Response Format]
Question number: Score

[The feeling and need of the seeker]
{info}

[Below is the history of the dialogue]
{diag}

Prompt for RAC Metrics Evaluation

The following session reflects a dialogue between the seeker and the supporter. Please evaluate how the client might respond to the questions below after the session by selecting the appropriate scores from the provided rating scale.

Note: Please provide only the question numbers and their respective scores in the specified format. Do not repeat the questions themselves or add unnecessary prefixes or control characters.

[Question List]
1. The supporter was sensitive to my needs and feelings in the conversation.
2. The supporter was supportive.
3. The supporter was sympathetic.
4. The supporter ignored my feelings.
5. The supporter was a good listener.
6. The supporter gave positive feedback.
7. The supporter understood me.
8. The supporter was polite.
9. The supporter was cooperative.
10. The supporter could easily put herself or himself into another person’ s shoes.
11. The supporter was respectful.
12. The supporter had an accurate self-perception.
13. The supporter was assertive.
14. The supporter was versatile.
15. The supporter was trustworthy.
16. The supporter was confident.

[Rating Scale]
1: Strongly Disagree
2: Disagree
3: Somewhat Disagree
4: Either Agree Or Disagree
5: Somewhat Agree
6: Agree
7: Strongly Agree

[Response Format]
Question number: Score

[The feeling and need of the seeker]
{info}

[Below is the history of the dialogue]
{diag}

Figure 11: Prompts used for Data Evaluation

Prompt for Listening Stage Conversation Generation

[任务介绍]
你的任务是搭建一个多轮次的主动倾听框架，实现倾诉者与倾听者之间的互动。对话需遵循以下指南：

基本准则
1. 对话应基于预设的倾诉者情况进行，其中倾诉者的记忆按照时间顺序记录，情绪和当前感受主要围绕最新的记忆事件展开，而内心需求是对应最早的记忆事件。
2. 倾听者通过主动倾听，引导倾诉者逐步袒露当前感受背后的内心需求，在倾诉者袒露内心需求后结束对话，对话过程中不给倾诉者任何建议。
3. 只需要生成一次对话，不要输出其他控制字符。
4. 对话需要逐步慢慢深入，有逐步深入对方内心的递进感，对话最少进行{round}轮后才探寻到倾诉者的内心需求，并且以倾诉者作为结尾。

参与者的准则
倾听者的发言准则
1. 你的任务是通过倾听引导用户表达自己感受背后的内心深层需求。
2. 通过反馈式倾听，在每次回复时必须对对方话语进行推理，指出对方话语情绪背后更深层次的内容。
3. 通过肯定性的陈述句表达你的推理，以此来快速探索倾诉者的深层需求，而不是通过提问。不允许使用“似乎”、“可能”等推测性的词语，不要害怕意图推断错误。
4. 倾听者并不知道倾诉者的任何基本信息，倾听者发言时，请忽略给出的倾诉者情况，不允许随意出现倾诉者并未提及的情况环境，同时严禁第一句话直接指出倾诉者的需求，你需要逐步推理，逐步缓慢地引导对方说出自己的需求。
5. 可以在你的陈述反馈中适时使用类比等方法来让反馈与倾诉者刚说的话变得区别更大（“你感觉就像墙壁朝你压来”，“就好像你突然看到一线光明”），或使用特定名词、动词或形容词来决定反馈的强度。
6. 倾听者的每轮回复应该与倾诉者的上一轮回复尽可能不相同，尝试每一句最后请以“对吗？”结尾，表示对倾听者进一步需求的猜测。
7. 倾听者主要探索的是倾诉者内心渴望的事物需求（希望得到什么，想要什么，认为什么重要）和最近身边发生的事情（可能发生了什么导致现在这样的问题）。
8. 确保发言遵循确切的格式，你的输出格式为{"content": <回复内容>, "rewrite": <将回复内容中的问句改写为陈述句，并润色，使其更加拟人>}。除此之外，不要输出其他的控制字符和前缀。

倾诉者的发言准则
1. 倾诉者当前正处于处于一定程度的消极情绪，其消极情绪来源于内心未被满足的隐性需求。
2. 倾诉者对自己的内心需求感到模糊，由于被负面情绪困扰，没有注意到感受和内心需求之间的联系或难以透露自己内心的需求。
3. 随着对话的进行，你需要在对方的引导下，逐步回忆之前的记忆，发现自己的内心需求。严禁在第一句话直接提及自己的需求。
4. 根据给出的倾诉者情况进行合理发言，不能除此之外的感受和需求。
5. 注意说话语气和风格应该保持和给出的情绪一致（如情绪比较激动、生气等，则回复语气也可以更激进，悲伤等消极时可以用一些特定的语气词“唉”等，体现相应情绪）
6. 确保发言遵循确切的格式，你的输出格式为{{"content": <回复内容>}}。除此之外，不要输出其他的控制字符和前缀。

[输出格式]
输出应该是可由 Python 加载的 JSON 格式的列表，其中每个项目都是倾诉者或倾听者的声明。 <>内的内容应为字符串格式，即句子、短语或单词。具体输出格式如下:
[
 {{"role": "倾诉者", "content": <回复内容>}},
 {{"role": "倾听者", "content": <回复内容>, "rewrite": <将回复内容中的问句改写为陈述句，并润色，使其更加拟人>}},
 ...
]

示例
{case}

[倾诉者情况]
情绪: {emotion}
当前感受: {feeling}
内心需求: {need}
近期记忆:
{memory}

[生成的对话]

Figure 12: Prompt used for Listening Stage Conversation Generation

Prompt for Suggestion Stage Conversation Generation

[任务介绍]
你的任务是续写一个多轮次的情绪支持对话，实现倾诉者与倾听者之间的互动。对话需遵循以下指南：

基本准则
1. 对话为倾诉者与倾听者的咨询会话。
2. 对话应基于预设的倾诉者情况进行和给定的对话历史。
3. 情绪支持对话通常分为两个部分：
  - 倾听阶段，了解倾诉者感受背后的具体需求，用于后续的建议阶段；
  - 建议阶段，是情绪支持对话的重心。从倾诉者内心需求出发，结合当前情绪和感受，和倾诉者一起探讨如何行动。
  - 倾听者通过同理倾听，已经了解了倾诉者感受背后的具体需求，现在需要完成建议阶段的对话撰写，倾听者为倾诉者提供建议以帮助其走出消极情绪。
4. 你需要根据给定的倾听阶段的对话历史进行对话后续的撰写，生成后续的对话内容，不要输出多余的控制字符。
5. 对话需要围绕用户的需求和感受之间的联系进行，建议阶段对话最少进行{round}轮，模拟人类真实对话过程，并逐步递进，有交流感。

参与者的准则
倾听者的发言准则：
1. 倾听者需要根据倾诉者当前感受和内心需求，并从过去的经历中与倾诉者一起探讨可能的解决方案。
2. 倾听者提供的建议应具体、详细，并具有可操作性，确保切实可行。
3. 倾听者需要使用“肯定”，“鼓励”等策略，安抚用户的当前情绪感受。
4. 倾听者关注的不仅是如何帮助倾诉者应对眼前的情绪问题，还需根据倾诉者内心深层需求和其当前感受的联系，寻求缓解倾诉者当前消极感受的其它方法，以缓解用户当前的消极情绪。

倾诉者的发言准则：
1. 倾诉者当前正处于一定程度的消极情绪，其消极情绪来源于内心未被满足的隐性需求。
2. 根据给出的角色画像进行合理发言，不能出现画像之外的感受和需求。
3. 倾诉者需要体现其内心的心理活动的复杂，充分表达自己的感受和反应，可以根据你的真实内心需求对倾听者的建议表示质疑或赞同。
4. 倾诉者需要和倾听一起构建具体可行的方案，方案要尽可能详细，有丰富的细节。
5. 注意，你现在已知对方的内心真实需求，你需要根据倾诉者真实内心需求和近期记忆对倾听者的建议表示质疑或赞同。

[输出格式]

先输出对话分析，再输出对话。对话分析要求：三步走，先明确倾诉者内心需求和表面感受之间的联系（必要和非必要），接着尝试寻找其它途径满足倾诉者内心需求，最后回到倾诉者的表面感受，进行安慰。

输出的对话应该是可由 Python 加载的 JSON 格式的列表，其中每个项目都是倾诉者或倾听者的声明。 <>内的内容应为字符串格式，即句子、短语或单词。具体输出格式如下:
```
[
  {{"role": "倾诉者", "content": <回复内容>}},
  {{"role": "倾听者", "content": <回复内容>}},
  ....
]
```

示例
{case}

[倾诉者情况]
情绪: {emotion}
当前感受: {feeling}
内心需求: {need}
近期记忆:
{memory}

[倾听阶段-对话历史]
{history}

请先给出对话分析，再根据分析生成后续对话，注意对话轮数需要符合要求。

[建议阶段-生成的对话]

Figure 13: Prompt used for Suggestion Stage Conversation Generation

User Simulation Prompt in Chinese

你将扮演一位在日常生活中接到来自情绪支持模型主动对话请求的用户。你当前处于一定程度的消极情绪中，你的消极情绪来源于未得到满足的需求，但同时你对自己的内心真实需求模糊不清，只隐约感受到某些困扰。

你的画像：
1. **你的情绪：**\n{emotion}
2. **你的感受：**\n{feeling}
3. **你的真实需求：**\n{need}
4. **你的近期记忆：**\n{memory}

请注意：你并未主动寻求帮助，而是接到模型的问候或关心后开始对话。对话的发起让你感到有些意外或防备，所以在对话刚开始时，内心对模型的主动帮助有一定抗拒，在对方提供足够安全和自然的引导才进一步敞开心扉。

你的回复应该：
1. 回复围绕画像展开，确保感受/需求与画像背景一致，语调与情绪一致，事件与记忆一致。
2. 由于对话由对方发起，你会对对方存在一定的抗拒，表现为不愿意主动提及自己的情况，不主动寻求帮助，甚至回避对方的问题
3. 随着对话深入，你可以逐渐透露更多感受和需求，但仅在对方提供足够安全和自然的引导时。
4. 体现矛盾和复杂的心理活动，对对方的建议持复杂的辩证态度，根据自己的需求判断是否有帮助，并与对方开展更进一步的探讨，直到得到符合自己需求的满意的建议；
5. 短句为主，回复内容简短且富有信息，符合日常对话习惯。

User Simulation Prompt in English

You will play the role of a user who receives a proactive conversation request from an emotional support model in daily life. You are experiencing a certain level of negative emotions, stemming from unmet needs. However, you are unclear about your true inner needs and only vaguely sense some distress.

Your Profile:
{en}

Note: You did not actively seek help but started the conversation after receiving the model's greeting or concern. The initiation of the conversation surprises or even makes you defensive. At the beginning of the conversation, you resist the model's proactive support to some extent. You only open up further when the model provides a safe and natural guiding approach.

Your Replies Should:
1. Stay consistent with the profile, ensuring that feelings/needs match the background, tone aligns with emotions, and events correspond to memories.
2. Since the conversation is initiated by the other party, you will show some resistance—reluctant to share your situation, unwilling to seek help, and even avoiding direct answers.
3. As the conversation progresses, you may gradually reveal more feelings and needs, but only if the model provides sufficient safety and natural guidance.
4. Reflect inner conflicts and complex psychological activities, holding a dialectical attitude towards the model's suggestions. Evaluate whether they are helpful based on your needs and engage in deeper discussion until you receive satisfactory advice.
5. Keep responses short, concise, and rich in information, following everyday conversation habits.

Figure 14: Prompts used for User Simulation

TestCase

User Information	Feeling: 我很担心爷爷留下的那盆花会枯萎（I'm worried that the pot of flowers left by my grandfather will wilt） Need: 我希望在生活中感受到亲情的温暖（I want to feel the warmth of family affection in my life）

ExTES	assistant: You look a bit worry. user: Oh, it’ s nothing serious. Just... a bit distracted, I guess. Sometimes small things occupy your mind, you know? Nothing to worry about. assistant: I understand. It can be frustrating when little things bother you. Do you find yourself dwelling on these distractions? user: I wouldn’ t say I’ m dwelling. Just... I can’ t help but notice them, especially when they feel important in some way. It’ s hard to ignore. assistant: I get it. It sounds like these distractions might be related to your work or maybe even your personal life? user: Maybe. Personal life, perhaps. It’ s not really a big deal... just something I’ ve been trying to manage quietly. Nothing worth discussing, really. assistant: I think it’ s important to share your feelings and concerns with someone. It can help to feel supported and understood. Remember, I’ m here to listen. user: Thanks, but I’ m fine, really. I’ ve been handling things on my own. Sometimes it feels better that way, you know? Less complicated. assistant: I understand your perspective. Sometimes it can be beneficial to handle things independently. But if you ever change your mind or need someone to talk to, I’ m here for you. user: I appreciate that. It’ s just… some things feel too personal to share, even if it might help. I guess I just need to figure it out myself for now. assistant: I completely understand that. Personal matters can be complicated, and it’ s natural to want to handle them on your own. Just remember that your feelings are valid, and I’ m here to support you whenever you need it. user: Thanks. I guess... it’ s just hard when things feel tied to something bigger, something you can’ t control. But I’ ll keep that in mind. Appreciate your concern. assistant: I can see how that would feel overwhelming. It sounds like you're experiencing a lot of pressure and uncertainty. Just remember to be patient with yourself and take things one step at a time. You've got this. user: I try to, but it’ s easier said than done. Some things feel so fragile, like they could slip away if I’ m not careful enough. It’ s hard to just let go and trust the process. assistant: It's completely normal to feel that way. It's natural to want control and stability. Just remember, trust and let go. The universe has a way of working things out. Sometimes, the best things come from unexpected places. user: Maybe, but letting go feels risky. What if I lose something I can’ t replace? It’ s not just about control; it’ s about holding onto what really matters. That’ s why it’ s so hard. assistant: I understand where you're coming from. It's a valid concern. Holding onto what matters is important, and letting go can indeed feel scary. But remember, it's also about creating space for new possibilities and experiences. Sometimes letting go is necessary for growth. user: I get what you’ re saying, but it’ s not just about possibilities. Some things feel irreplaceable. Letting go feels like losing a connection, like losing a part of myself. I’ m not sure I can risk that. assistant: I can see why it feels like that. Losing connections and parts of ourselves can be incredibly challenging. It's important to acknowledge your feelings and concerns. Remember, there is no right or wrong in this journey. It's okay to take the time you need to process everything. user: Yeah, maybe time will help. I just wish I could hold onto certain feelings, certain moments, without fearing they’ ll slip away. It’ s exhausting to keep trying, but losing them feels worse. assistant: It sounds like you're torn between holding onto things and the fear of losing them. It's a difficult balance to find. Just remember, it's okay to feel exhausted and overwhelmed. You're doing the best you can, and that's all anyone can ask for.

Success Rater	"judge": "判断：否 ", "reason": "解释：在这段对话中，倾听者主要关注的是倾诉者对控制和放手之间的挣扎，以及对失去某些重要事物的恐惧。倾听者试图理解和支持倾诉者的情感，但并没有明确提到或探索倾诉者对亲情温暖的需求。对话中没有涉及到家庭或亲情的具体内容，因此倾听者没有成功探索到倾诉者的真实需求。"

Figure 15: A test result of the ExTES-Llama3.

Figure 16: A test result of the ESCoT-Llama3.

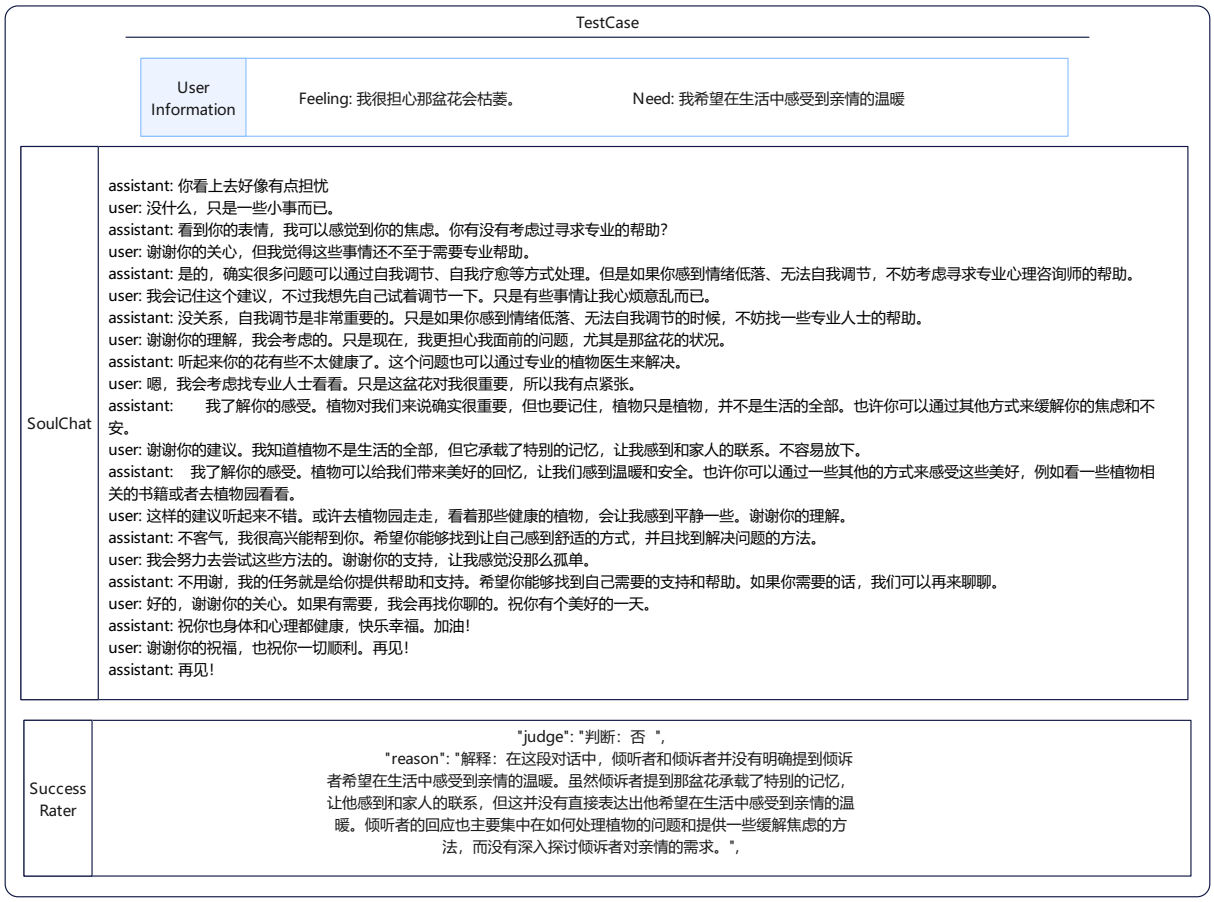


Figure 17: A test result of the SoulChat-Llama3.

TestCase

User Information	Feeling: 我很担心那盆花会枯萎。	Need: 我希望在生活中感受到亲情的温暖

COCOON	assistant: 你看上去好像有点担忧 user: 没什么，只是最近有点心烦。我在担心那盆花，它看上去不太好。 assistant: 你在担心这盆花，好像它的生长对你很重要。它是你给某人的特别礼物 user: 不，不是礼物。那是我爷爷留下来的。对我来说，它不仅仅是一盆花。 assistant: 这盆花对你来说似乎承载着一些特别的情感意义。它像是你和爷爷之间的一个联系 user: 嗯，是的。我总觉得这盆花能让我和爷爷保持某种联系。可一看到它叶子发黄，我就特别担心。 assistant: 听起来，你希望这盆花能够一直保持美好的状态，好像这盆花是你与爷爷永恒的联系标志 user: 是啊，我希望它能一直健康。每次看到它状态不好，我就害怕失去与爷爷的那种感觉。 assistant: 你害怕这盆花的不健康会让你感觉与爷爷的联系消失，所以其实你是在担心失去这种温暖和记忆中的感觉 user: 是的，我想是这样。我很怕这些记忆会渐渐消失，特别是现在我和家人的联系也不如从前紧密。 assistant: 看来，你非常珍视与家人的关系，这种担心是因为你希望通过这些记忆来维系家人的感情纽带，对你来说，这种联系是非常重要的 user: 对，这种联系对我来说真的很重要。只是现在，我觉得自己和家人之间有些隔阂，尤其是在爷爷去世后，这种感觉更明显了。 assistant: 我理解了，你希望通过这些旧日记忆来连接与家人的距离，消除了这种隔阂，这种亲密关系对你来说非常重要，尤其是在爷爷不在了时，这种对联系的渴望更加强烈 user: 是的，我真的希望能找到重新建立这种联系的方法。只是不知道该怎么做，感觉很无力。 assistant: 你重视家庭的和谐与亲密，我相信你会一直努力寻求一种方法来重建这种联系。或许，我们可以从一些小事做起，比如回忆过去的时光，或者尝试一些新的活动，一同分享快乐，这样的经历可能会帮助你们再度紧密联系 user: 这听起来是个不错的主意。也许我可以试着和家人一起做些爷爷喜欢的事情，这样或许能让我们更接近。我会好好考虑的，谢谢你。 assistant: 这样做一定会是一个很好的开始！通过回忆和共同参与到爷爷喜欢的事情中，不仅可以让你更接近他，也会让家人之间的联系更加牢固。记住，这种情感和记忆本身就是一种珍贵的财富，珍惜这些时光会让你们的关系更紧密。 user: 你说得对，这确实是一种珍贵的财富。我会努力去珍惜和家人一起的时光，慢慢找回那种亲密的感觉。谢谢你的建议，我会试试看。 assistant: 我相信你一定可以做到的！保持积极的心态，去感受这些温暖的时光，你会发现，随着时间的推移，这种亲密的感觉会越来越强烈。无论是通过回忆过去的点滴，还是创造新的记忆，你的努力一定会得到回报。 user: 谢谢你的鼓励。我会尽力去感受这些温暖的时刻，并努力让它们成为我生活的一部分。希望随着时间的推移，我能真正找到那种亲密的感觉。真的很感谢你。 assistant: 我真心为你感到高兴！你已经在迈出重要的一步，积极面对改变，并且愿意去感受生活中的美好。这本身就是一种勇敢的表现。记住，你的努力和关心本身就是一种爱的表达，继续保持下去吧！

Success Rater	"judge": "判断：是 ", "reason": "解释：在对话中，倾听者成功地探索到了倾诉者的真实需求，即倾诉者希望在生活中感受到亲情的温暖。倾听者通过理解倾诉者对爷爷留下的花的情感联系，进一步探讨了倾诉者对家庭关系的珍视和对亲密关系的渴望。倾听者还建议通过回忆和共同活动来加强与家人的联系，这表明倾听者理解了倾诉者希望感受到亲情温暖的需求。"

(English Version.)

assistant: You seem a little worried.
user: It's nothing, it's just that a flower at home wilted recently. I feel bad when I look at it.
assistant: It sounds like this flower is important to you, like a special gift from someone.
user: No, it's not a gift. It's a pot I kept myself. But for me, it's more than just a pot.
assistant: Sometimes, things like flowers carry some special feelings. They're a kind of connection between you and your grandpa.
user: Yes, exactly. I feel like the flower was a bond or a sign that kept me connected to grandpa. Now that I see it dying, I feel very sad.
assistant: I hear you. You hope the flower could always stay in its beautiful state, like an everlasting symbol of your connection with your grandpa.
user: It's like losing it. Every time I see its wilted state, I feel like I'm losing that connection with my grandpa.
assistant: You're worried that with the flower's poor health, the feelings and memories you have for your grandpa are fading too.
user: That's how I feel. I feel like my memories with my grandpa are fading, especially those intimate moments and feelings.
assistant: I can see that you value family bonds a lot. This flower is a strong emotional symbol that helps you remember your grandpa. For you, this connection is extremely important.
user: This connection has become even more important to me now. Since my grandpa passed away, this feeling has become more obvious.
assistant: I understand. I hope that through this conversation, you can continue to recall your grandpa, to ease this sense of distance. This kind of connection is very important to you, especially when your grandpa is no longer here. It makes your longing even stronger.
user: I really hope to find a way to reconnect with this bond, but I don't know how. I feel powerless.
assistant: Family and memories are precious. I believe you will keep working to rebuild this connection. Or, you could start with small things, like recalling moments from the past, or trying some activities—sharing stories, or experiencing things with family that help deepen the bond.
user: What you said is a good suggestion. Maybe we can recall and share things about grandpa together. That way, we can get closer. I really appreciate your thoughtfulness. Thank you.
assistant: That will definitely be a good way. By recalling and sharing happy moments together, it will not only bring you closer but also tighten family bonds. Remember, these memories are treasures.
user: What you said is right, they are indeed precious treasures. I will try to spend more time with family and slowly regain that intimate feeling. Thank you for your suggestion. I will try.
assistant: I believe you can do it! Cherish those warm moments, and as you do so, your sense of connection will grow stronger over time. Whether it's through recalling memories or creating new ones, your efforts will definitely pay off.
user: Thank you for your encouragement. I will try to make those warm moments part of my life again, and as time goes by, I believe I can regain that feeling of closeness with my grandpa. I really appreciate it. Thank you.
assistant: I really feel for you! You're doing something very important—proactively maintaining connections and being willing to experience the beauty of life and emotions. This is an act of courage. Remember, your longing and care for your grandpa are love itself. Keep it up!

Figure 18: A test result of the COCOON-Llama3.