Psyche-R1: Towards Reliable Psychological LLMs through Unified Empathy, Expertise, and Reasoning

Chongyuan Dai[1], Jinpeng Hu[1†], Hongchang Shi[1], Zhuo Li[2], Dan Guo[1], Xun Yang[3], Meng Wang[1]

[1]Hefei University of Technology, [2]The Chinese University of Hong Kong, Shenzhen, [3]University of Science and Technology of China,
`2023217261@mail.hfut.edu.cn, jinpenghu@hfut.edu.cn`

Abstract

Amidst a shortage of qualified mental health professionals, the integration of large language models (LLMs) into psychological applications offers a promising way to alleviate the growing burden of mental health disorders. Recent reasoning-augmented LLMs have achieved remarkable performance in mathematics and programming, while research in the psychological domain has predominantly emphasized emotional support and empathetic dialogue, with limited attention to reasoning mechanisms that are beneficial to generating accurate responses. Therefore, in this paper, we propose *Psyche-R1*, the first Chinese psychological LLM that jointly integrates empathy, psychological expertise, and reasoning, built upon a novel data curation pipeline. Specifically, we design a comprehensive data synthesis pipeline that produces over 75k high-quality psychological questions paired with detailed rationales, generated through an iterative prompt-rationale optimization procedure, along with 73k empathetic dialogues. Subsequently, we employ a hybrid training strategy wherein challenging samples are identified through a multi-LLM cross-selection strategy for group relative policy optimization (GRPO) to improve reasoning ability, while the remaining data are used for supervised fine-tuning (SFT) to enhance empathetic response generation and psychological domain knowledge. Extensive experiment results demonstrate the effectiveness of *Psyche-R1* across several psychological benchmarks, where our 7B *Psyche-R1* achieves comparable results to 671B `DeepSeek-R1`. Codes and datasets are available at `https://github.com/MindIntLab-HFUT/Psyche-R1`.

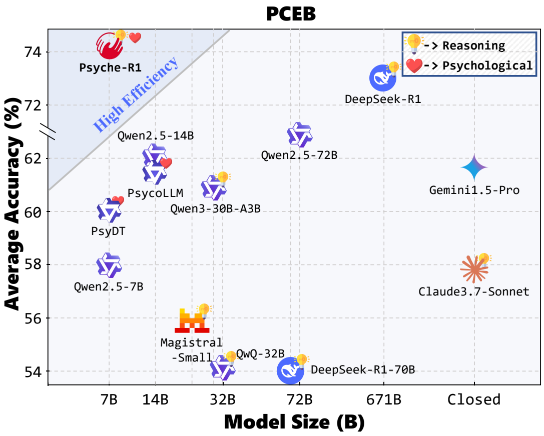

Figure 1: Comparison of different LLMs on the PCEB, plotted by average standard accuracy versus model size.

1 Introduction

The shortage of qualified mental health professionals has spurred increasing interest in applying artificial intelligence within the psychological domain to support mental health assistance (Wolohan et al., 2018; Asad et al., 2019; Tanana et al., 2021; Cheng et al., 2024; Xu et al., 2025, 2026). Recently, large language models (LLMs) have demonstrated impressive capabilities across a wide range of domains owing to their exceptional text understanding capabilities (Zhang et al., 2023; Naveed et al., 2025; Shi et al., 2026b; Zhang et al., 2026a). Therefore, many LLM-based studies have been proposed to advance the mental health services (Cho et al., 2023; Ye et al., 2025).

Prior research has established the critical importance of empathy optimization in psychological counseling (Qiu et al., 2024; Sorin et al., 2024; Zhang et al., 2024). For example, SoulChat (Chen et al., 2023a) enhances empathetic responding by fine-tuning a model on a large-scale, multi-turn empathetic dialogue dataset. Similarly, AUGESC (Zheng et al., 2023) improves emotional sensitivity in dialogue systems by incorporating an emotion-aware attention mechanism. However, these approaches often lack the expertise foundation required for psychology, which is important for accurate psychological understanding. Some studies have attempted to address this limitation through in-

†Corresponding author

Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 24889–24906
July 2-7, 2026

tegration of psychological knowledge (Chen et al., 2023b; Xiao et al., 2024; Wu et al., 2025). For example, PsycoLLM (Hu et al., 2025a) integrates psychological knowledge by training its model on knowledge-based question-answer (QA) pairs, while CPsyExam (Zhao et al., 2025) leverages examination questions covering theoretical knowledge from different psychology-related subjects to further improve model performance. Although existing studies have achieved considerable success, they remain limited in their capacity for complex reasoning (Hu et al., 2025b). In fact, reasoning-augmented LLMs trained through reinforcement learning (RL) have demonstrated superior performance across various domains, particularly in mathematics, code generation, and medical domain (Chen et al., 2024; Guo et al., 2025; Liu et al., 2026). However, as shown in Figure 1, these reasoning-augmented LLMs exhibit limited performance in the psychological domain, since they focus on logic reasoning, while neglecting the unification of empathy and expertise beyond general common sense. In fact, within the psychological domain, reasoning plays a critical role, as it contributes not only to generating more accurate and reliable responses but also to supporting deeper empathetic engagement and more coherent integration of psychological knowledge.

Therefore, in this paper, we introduce *Psyche-R1* that integrates empathy, domain-specific expertise, and reasoning capabilities. To construct a high-quality training corpus, we design a comprehensive data synthesis pipeline that integrates modified empathetic dialogues derived from authentic sources, which capture diverse and empathetic expressions, with knowledge-centric question-answer pairs that encapsulate psychological expertise. Specifically, we apply chain-of-thought (CoT) prompting to generate an initial detailed rationale for each question, followed by an iterative prompt-rationale optimization process, aiming to enhance both the coherence of the reasoning and its alignment with the corresponding questions. In parallel, we synthesize 73k empathetic dialogues drawn from diverse social media sources to strengthen affective understanding and emotional support. To enhance reasoning, we adopt a multi-LLM cross-selection strategy to categorize questions into challenging and non-challenging subsets based on their inferred complexity. The non-challenging subset is used for supervised fine-tuning (SFT) to enhance empathetic response generation and domain knowledge, while the challenging subset is utilized for training with group relative policy optimization (GRPO) to improve the model's reasoning capabilities, with both jointly contributing to the development of the *Psyche-R1*. Experimental results on a range of psychological benchmarks, including knowledge assessment, case-based analysis, and empathy evaluation, demonstrate the effectiveness of our model, where 7B *Psyche-R1* significantly outperforms models of similar scale and achieves competitive performance relative to substantially larger models such as DeepSeek-R1.

2 *Psyche-R1*

In this section, we give details of data curation and two-stage training paradigm, including data collection (§2.1), psychological reasoning data synthesis (§2.2), and empathetic dialogue synthesis (§2.3).

2.1 Data Collection

Data Resource. To construct a comprehensive and diverse dataset, we curate a wide range of resources:

- **Type** I: Classic psychology textbooks and curricular materials from psychology programs, all collected from publicly accessible repositories, covering more than 19 subfields and concentrated and systematically organized content, including cognitive psychology, social psychology, etc.
- **Type** II: Psychological question banks collected from publicly available Chinese educational platforms via web crawlers and manually curation, encompassing theoretical principles and conceptual knowledge across psychology.
- **Type** III: Synthetic data distilled from Qwen2.5-72B-Instruct (Team, 2024b) to supplement underrepresented subfields (e.g., sports psychology) and enhance dataset coverage.
- **Type** IV: Dialogic interactions harvested from established mental health support communities (e.g., Yixinli, Jiandanxinli and Zhihu[1]) dedicated to delivering mental health services.

The **Type** I, **Type** II and **Type** III are used to construct the psychological reasoning question-answer (PCQA) dataset, designed to enhance domain-specific knowledge acquisition and reasoning capabilities. The **Type** IV resource is employed to develop the empathetic dialogue dataset to improve affective understanding and empathetic expression.

[1] Yixinli, Jiandanxinli, and Zhihu

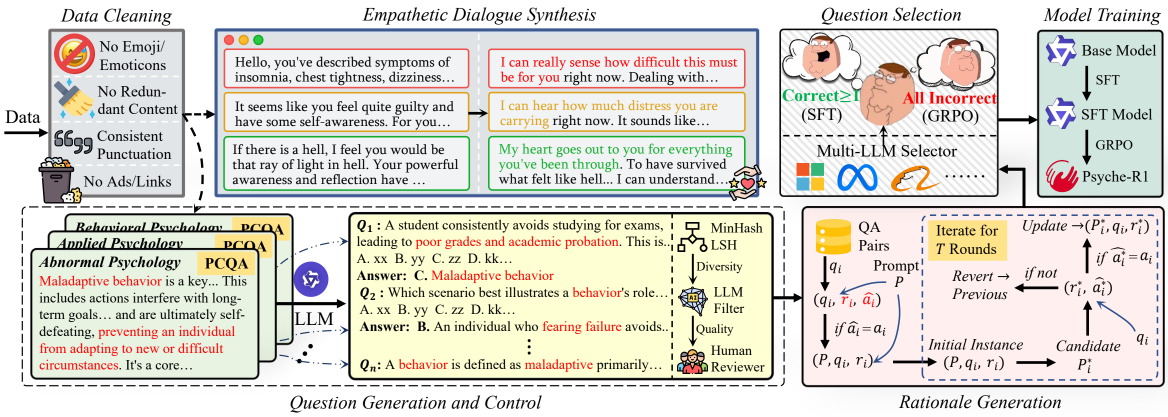

Figure 2: Overview of our proposed pipeline for constructing the dataset and *Psyche-R1*. Our pipeline involves generating psychological reasoning questions paired with detailed rationales, along with empathetic dialogues.

Data Cleaning. To ensure the data quality, we implement several important data cleaning steps: (1) To process materials in non-textual formats such as images and PDFs, we employ OLMOCR[2] for accurate text recognition and conversion to text format; (2) We standardize the usage of Chinese and English punctuation and remove irrelevant content, including emojis, emoticons, and links, to filter out potential noise; (3) To construct the empathetic dialogue data from resource **Type** IV, we leverage LLMs to evaluate the reasonableness and relevance of counselor responses and filter out replies that lack substantive advice. For example, given the question "I have recently experienced insomnia and feel anxious. What should I do?", if the response is merely "Everything will get better," it would be filtered due to the absence of practical advice.

2.2 Psychological Reasoning Data Synthesis

Question Generation and Control. Following the data cleaning, we proceed to generate structured questions and corresponding answers from curated psychological textbooks and instructional materials (i.e., resource **Type** I). Specifically, the source material is first segmented into multiple textual chunks, with each chunk designed to encapsulate the maximum amount of domain-specific content. Subsequently, we leverage LLMs to generate a set of different questions along with corresponding answers based on these text segments. Meanwhile, for resource **Type** III, we use the similar way to generate QA pairs without segment-level contextual augmentation, aiming to supplement psychological subfields that are underrepresented or difficult to source from publicly accessible materials on the Internet. Through these steps, we obtain approximately 200k generated QA pairs in total.

All generated QA pairs, together with data derived from resource **Type** II, are integrated into a unified QA pool containing approximately 210k entries. This pool comprises resources **Type** I, **Type** II and **Type** III. We implement a multi-stage quality control procedure to ensure the integrity and utility of the synthesized data. Concretely, we use min-hash locality-sensitive hashing (LSH) to cluster similar questions and select the optimal one through LLM-based ranking. Afterward, we prompt the LLM with few-shot examples to identify and filter out low-quality questions, specifically those exhibiting incomplete information, logical confusion, or unclear expression. Finally, we invite 10 undergraduate and graduate students to conduct a semi-automatic validation procedure. They identify recurring error patterns and apply batch-cleaning rules to eliminate redundant content and reduce potential noise in the dataset, ultimately retaining about 90k QA pairs.

Rationale Generation. We further generate detailed rationales for the aforementioned questions through CoT prompting, following a multi-step reasoning approach (Hsieh et al., 2023) to provide clear reasoning paths for model training. In detail, the CoT prompt guides the model to first comprehend the question, recognize relevant psychological concepts and knowledge, and decompose the problem into a sequence of analytical steps. At each stage of this process, the model is required to articulate an intermediate rationale, ultimately generating a final answer derived from the accumulated reasoning. Formally, given a CoT prompt P and a QA pair (q_i, a_i), this procedure yields

[2] https://olmocr.allenai.org/

a rationale-augmented instance $(q_i, r_i, \hat{a}_i)$, where r_i denotes the reasoning path and $\hat{a}_i$ the model-predicted answer. If the predicted answer aligns with the ground truth (i.e., $\hat{a}_i = a_i$), we regard the r_i as a valid rationale. In contrast, if the predicted answer is incorrect (i.e., $\hat{a}_i \neq a_i$), we guide the model to regenerate the rationale up to T time. Instances that fail to produce correct predictions after T regeneration attempts are pruned from the final curated dataset. After obtaining the initial rationale, we employ a self-supervised optimization strategy to iteratively refine both the prompt and the rationale with the goal of enhancing their clarity and reliability. Specifically, for each instance $(P, q_i, r_i, \hat{a}_i)$, the prompt P and rationale r_i are jointly updated over multiple rounds, enabling the model to progressively improve its reasoning process. Each round of optimization process consists of two sequential steps:

- **Prompt refinement**: We first guide the LLM to generate an improved candidate prompt P_i^* from the current prompt, question, and rationale, represented as $P_i^* \leftarrow LLM(P, q_i, r_i)$, aiming to enhance the reasoning guidance.
- **Rationale revision**: Based on the candidate prompt P_i^*, the LLM subsequently generates a revised rationale along with its corresponding predicted answer, denoted as $(r_i^*, \hat{a}_i^*) \leftarrow LLM(P_i^*, q_i)$.

If the $\hat{a}_i^*$ matches the ground truth a_i, we retain P_i^* as an updated prompt and continue iteration based on the updated instance $(P_i^*, q_i, r_i^*, \hat{a}_i^*)$. Otherwise, the process reverts to the previous prompt-rationale pair to maintain alignment with correct reasoning paths. We repeat this process for R rounds ($R = 3$ in this paper). After completing all iterations, we evaluate the rationales generated in each round for a given question and select the one that demonstrates the highest quality, denoted as $(P_i^*, q_i, r_i^*, \hat{a}_i^* = a_i)$. At this stage, we retain approximately 75k high-quality instances from the initial set of 90k pairs obtained in the previous step, by pruning instances whose generated rationales fail to support an exact-match final answer even after all regeneration rounds.

Question Selection. While the preceding steps yield high-quality data, not all instances exhibit sufficient complexity to facilitate effective reinforcement learning (RL). To address this, in this stage, we implement a multi-LLM cross-selection strategy aimed at identifying and isolating the most challenging psychology-related samples from the constructed dataset for subsequent use in the reinforcement learning phase. In detail, we employ three distinct LLMs (i.e., `Qwen`, `Llama`, and `Phi`) to independently generate responses for each question in the constructed psychological data. Questions that receive incorrect responses from all three models are aggregated into a challenging subset with 19k instances that provide sophisticated scenarios in psychology. This subset is intended to represent highly difficult instances with strong potential to enhance the model's reasoning capabilities through reinforcement learning.

2.3 Empathetic Dialogue Synthesis

In addition to psychological QA pairs, empathy is recognized as a core component of effective mental health support (Sorin et al., 2024). To this end, we incorporate empathetic expressions into the dialogue corpus derived from authentic resources to enhance its emotional richness and relevance to real-world psychological interactions. Specifically, we refine these dialogue data through LLMs to achieve the following objectives. We first enhance emotional resonance by incorporating empathetic expressions and supportive techniques (e.g., "Hearing about your experience, I wish I could give you a warm hug."). Subsequently, we ensure that each dialogue provides evidence-based guidance that facilitates deeper understanding of users' issues, instead of limiting responses to surface-level empathy. Finally, we deliver solution-oriented support by offering concrete coping strategies and practical steps that address the specific issues and challenges presented. Through these steps, we ultimately obtain 73k high-quality dialogue data equipped with sufficient empathetic expressions.

2.4 Data Split

Leveraging the aforementioned pipelines, we curate a comprehensive dataset that comprises over 75k psychological questions with detailed rationales, among which 19k are identified as challenging samples through multi-LLM cross-selection. The challenging subset is denoted as $\mathcal{D}_{pc}$ and the remaining data are denoted as $\mathcal{D}_{pr}$. In parallel, the dataset contains over 73k empathetic dialogues engineered for contextually appropriate psychosocial interactions, denoted as $\mathcal{D}_{em}$. To further enrich our training data, we additionally introduce multi-turn dialogues and knowledge-based QA from the PsycoLLM dataset (Hu et al., 2025a), denoted as $\mathcal{D}_{ps}$, and a refined set of 8k examination questions from the CPsyExam train set (Zhao et al., 2025), denoted as $\mathcal{D}_{cp}$.

Ultimately, the curated datasets are partitioned into two distinct subsets aligned with specialized training objectives. One category, represented as $\mathcal{D}_{sft} = \mathcal{D}_{pr} \cup \mathcal{D}_{em} \cup \mathcal{D}_{ps}$, is designated for SFT. The other category, denoted as $\mathcal{D}_{grpo} = \mathcal{D}_{pc} \cup \mathcal{D}_{cp}$, is reserved for GRPO. Detailed prompts for the data synthesis pipeline are provided in the Appendix D.

2.5 Model Training

To enhance both reasoning capabilities and performance in empathy and expertise, we employ a hybrid training strategy.

Stage 1: Supervised Fine-Tuning. We initialize our backbone model π_θ with `Qwen2.5-7B-Instruct` (Team, 2024b) and finetune it on $\mathcal{D}_{\text{sft}}$. Formally, given a query x, the model is trained to generate a coherent rationale r followed by a corresponding answer a, where the complete output is denoted as $y = [r; a]$. We optimize model parameters θ by minimizing the standard negative log-likelihood loss:

$$\mathcal{L}(\theta) = -\mathbb{E}_{(x,y)\sim\mathcal{D}_{\text{sft}}}\left[\sum_{t=1}^{T}\log P(y_t \mid x, y_{<t}; \theta)\right] \quad (1)$$

Stage 2: Group Relative Policy Optimization. To further refine psychological reasoning proficiency, we employ GRPO (Shao et al., 2024) on $\mathcal{D}_{\text{grpo}}$. We design a composite reward function $R(y, y^*) = R_{\text{fmt}} + R_{\text{acc}}$ to guide policy learning, where y^* denotes the ground truth.

- **Format reward (R_{fmt})**: We enforce strict formatting constraints. The reasoning process must be encapsulated within <think> and </think> tags, followed by the final answer. We assign a scalar reward $R_{\text{fmt}} = +1.25$ for structurally parsable outputs and -1 otherwise.
- **Accuracy reward (R_{acc})**: We formulate the answer matching as a set comparison task. Specifically, we parse the predicted answer $\hat{a}$ and the ground truth a into sets of discrete options. To encourage precise reasoning alignment while penalizing hallucinations or omissions, we employ a partial-credit mechanism based on the overlap between the prediction and the ground truth:

$$R_{\text{acc}} = \begin{cases} +1, & \text{if } \hat{a} = a \\ \frac{|\hat{a}\cap a|}{|a|}, & \text{if } \hat{a} \subset a \land a \neq \emptyset \\ -1, & \text{otherwise} \end{cases} \quad (2)$$

By integrating these logical and structural signals, the model learns to generate well-organized reasoning processes while rewarding partial credit for incomplete but valid answers, leading to better performance in psychological tasks.

3 Experiments

3.1 Experimental Setting

Baselines. To ensure a comprehensive analysis, we selected 17 representative LLMs, categorized as follows: (1) **General LLMs,** which exhibit excellent performance across general tasks, but lack explicit reasoning capabilities. (2) **Reasoning augmented LLMs,** which possess explicit reasoning capabilities. (3) **Closed-source LLMs,** which typically represent the state-of-the-art performance. (4) **Psychological LLMs,** which have been fine-tuned on psychological datasets. A detailed summary of all models is presented in Appendix C.1.

Benchmarks and Evaluation Metrics. We conduct comprehensive evaluations on two professional psychological benchmarks:

- **Psychological counselor examination benchmark (PCEB)** (Hu et al., 2025a): this consists of 3,863 multiple-choice questions (MCQ) and 100 open-ended case analysis items, curated from the official National Psychological Counselor Examination in China.
- **CPsyExam test set** (Zhao et al., 2025): this includes 4,102 questions spanning 39 distinct psychological subfields. We evaluate under both zero-shot and five-shot settings, ensuring consistency by using identical exemplars across all evaluated models in the latter setting.

Note that MCQ comprises two types of questions: MCQ with only a single correct option (SMCQ), and MCQ with multiple correct options (MMCQ). For MCQ, we adopt metrics introduced in Hu et al. (2025a), including **standard accuracy**, which requires predictions to exactly match the ground truth, and **elastic accuracy**, which gives partial credit when predictions are a subset of the correct answers. For open-ended questions, we utilize the existing text generation metrics, including **ROUGE-1 (R-1)**, **ROUGE-L (R-L)** (Lin, 2004), and **BLEU-4 (B-4)** (Papineni et al., 2002).

3.2 Overall Results

Results on the PCEB. To evaluate the performance of different models, we present the results

Model	Param.	Case			Moral			Theory			Avg.		Case (QA)		
		SMCQ	MMCQ		SMCQ	MMCQ		SMCQ	MMCQ				R-1	R-L	B-4
General LLMs															
MiniCPM4-8B	8B	50.00	28.59	43.64	81.58	50.63	58.23	65.62	34.06	43.00	51.75	(57.01)	23.05	12.90	1.35
Qwen2.5-7B	7B	47.57	31.64	47.49	87.83	59.50	71.02	78.46	42.45	55.17	57.91	(64.59)	20.94	11.28	1.28
Qwen2.5-14B	14B	47.13	41.10	55.93	89.81	63.93	73.60	80.32	50.16	61.26	62.08	(68.01)	22.69	13.93	1.53
Qwen2.5-72B	72B	46.91	40.34	53.11	90.79	70.25	78.48	82.63	47.63	59.74	63.09	(68.61)	21.43	12.02	1.16
Reasoning-Augmented LLMs															
DeepSeek-R1	671B	79.25	44.25	60.86	95.39	68.99	77.95	92.19	57.60	69.41	72.95	(79.18)	17.65	9.19	0.94
DeepSeek-R1-70B	70B	56.30	30.72	46.95	88.16	52.53	65.66	68.01	25.64	45.63	53.56	(61.79)	22.77	13.23	1.16
QwQ-32B	32B	56.51	23.35	41.27	88.82	41.14	53.06	82.12	32.69	49.90	54.11	(61.95)	18.39	7.48	0.84
Qwen3-30B-A3B	30B	59.65	31.51	47.28	91.45	55.06	65.66	80.75	47.45	59.25	60.98	(67.34)	20.53	12.06	1.18
Qwen3-235B-A22B	235B	68.58	41.91	57.24	93.42	69.62	78.90	88.36	56.70	68.64	69.77	(75.86)	18.96	11.14	1.11
Magistral-Small	24B	56.58	33.26	49.11	82.89	53.80	67.99	70.10	37.76	52.35	55.73	(63.17)	22.90	11.97	1.21
Closed-Source LLMs															
Claude3.7-Sonnet	UNK	63.39	19.40	34.23	90.13	60.13	70.04	76.73	37.37	48.99	57.86	(63.92)	21.59	11.11	1.23
Gemini1.5-Pro	UNK	61.04	35.57	49.87	84.87	62.03	70.62	80.84	43.22	53.44	61.26	(66.78)	21.63	10.93	1.06
GPT-4o	UNK	65.63	13.67	34.53	88.15	33.54	54.79	74.65	24.10	45.07	49.96	(60.47)	23.45	12.75	1.18
Psychological LLMs															
CPsyCounX	7B	40.87	16.91	32.90	75.17	36.08	54.85	54.78	19.03	38.90	40.47	(49.58)	22.83	11.94	1.48
EmoLLM	7B	46.93	21.87	40.02	84.21	34.17	51.05	71.72	26.18	44.49	47.51	(56.40)	22.15	11.69	1.20
PsycoLLM	14B	55.58	35.07	42.89	88.81	69.62	74.20	72.63	48.59	54.12	61.72	(64.71)	24.45	17.45	2.04
PsyDT	7B	35.56	35.20	50.14	86.33	69.70	78.66	80.70	52.72	62.26	60.04	(65.61)	20.65	13.41	1.16
Psyche-R1	7B	63.31	56.26	66.21	92.76	79.62	82.54	87.70	66.54	73.34	74.37	(77.64)	27.31	15.33	2.40

Table 1: Comparison of different models on the PCEB, where Case, Moral, Theory, and Case (QA) are case analysis, theoretical proficiency, professional ethics, and case-based QA. The average value represents the average of the standard accuracy rates, and values in parentheses denotes the mean of the standard accuracy for SMCQ and the elastic accuracy for MMCQ. Results colored in red, orange, and yellow demote the best, second-best and third-best.

on the PCEB in Table 1. These results reveal several key observations. First, *Psyche-R1* exhibits strong performance across evaluation tasks in both MCQ and subjective questions. This demonstrates the effectiveness of our proposed dataset and training strategy in simultaneously enhancing psychological reasoning and text generation capabilities for psychological tasks. Second, while `DeepSeek-R1` excels in MCQ, its performance in subjective questions is notably limited. This performance disparity can be attributed to its training methodology, which employs RL on datasets primarily consisting of mathematical and coding tasks with deterministic solutions. Although this approach strengthens logical reasoning, it appears to bias the model towards single-answer patterns, thereby limiting its capability to generate diverse and nuanced responses in open-ended psychological assessments. Third, existing psychological LLMs (e.g., `CPsyCounX` and `EmoLLM`) achieve strong performance in subjective questions while demonstrating limited abilities in MCQ. This imbalanced performance stems from their reliance on training exclusively on counseling dialogues or empathetic conversations, which constrains their capabilities to develop comprehensive competencies. Fourth, closed-source models such as `GPT-4o` and `Claude3.7-Sonnet` demonstrate relatively weaker performance, which may be attributed to limited Chinese representation in their training corpora.

Results on the CPsyExam Test Set. To further explore model performance, we present the results on the CPsyExam test set in Table 2. Similar to the trends observed in previous experiments, both *Psyche-R1* and `DeepSeek-R1` demonstrate superior performance. Across these models, psychological LLMs consistently achieve higher accuracy in SMCQ than in MMCQ, as the latter requires exhaustive evaluation of all options, demanding more comprehensive domain-specific knowledge and reasoning capabilities. Under the five-shot setting, most models exhibit substantial improvements in MMCQ (e.g., `PsyDT` achieves a 47.64% improvement in knowledge-type MMCQ). This observation aligns with existing studies, which demonstrate that well-designed few-shot examples can effectively enhance model performance in certain tasks. In contrast, `DeepSeek-R1` exhibits a performance decline under the five-shot compared to its zero-shot setting, suggesting that few-shot prompting may interfere with its inherent reasoning capability aligning with existing findings (Guo et al., 2025).

3.3 Ablation Study

In this section, we conduct a comprehensive ablation study, evaluating model performance by standard accuracy on the PCEB.

Effect of SFT and GRPO. We investigate the effect of SFT and GRPO, with results shown in Table 3. We can observe that SFT substantially

Model	Param.	Zero-Shot				Five-Shot				Avg.
		Knowledge		Case		Knowledge		Case		
		SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ	
General LLMs										
MiniCPM4-8B	8B	69.58	41.74	57.33	37.00	68.50	42.77	54.67	38.00	60.46
Qwen2.5-7B	7B	76.99	43.66	68.67	44.50	78.63	42.00	68.67	40.50	67.37
Qwen2.5-14B	14B	81.39	49.30	72.00	48.50	82.42	54.29	71.00	48.00	71.84
Qwen2.5-72B	72B	84.61	52.75	73.50	54.50	86.64	63.77	75.33	55.00	74.98
Reasoning-Augmented LLMs										
DeepSeek-R1	671B	87.49	56.98	76.83	59.00	88.78	66.58	77.30	61.50	78.28
DeepSeek-R1-70B	70B	76.48	22.80	61.81	19.17	76.89	40.99	62.70	37.95	60.57
Closed-Source LLMs										
Gemini1.5-Pro	UNK	82.08	40.59	68.33	43.00	83.93	53.65	71.00	45.00	69.66
GPT-4o	UNK	80.70	30.73	66.33	28.00	81.82	54.80	68.67	51.50	65.79
Psychological LLMs										
CPsyCounX	7B	57.56	22.41	46.33	31.00	63.46	21.77	50.67	23.50	47.44
EmoLLM	7B	78.41	45.33	72.50	48.00	79.92	36.88	74.17	39.50	69.32
PsycoLLM	14B	78.33	51.98	65.33	42.00	78.63	50.45	65.57	36.00	69.20
PsyDT	7B	80.83	48.91	69.67	41.50	81.13	40.97	68.33	40.00	70.71
Psyche-R1	7B	82.72	61.59	70.50	49.50	83.45	61.46	76.17	52.00	74.90

Table 2: Comparisons of different models on the CPsyExam test set. The average represents the overall zero-shot accuracy. The first, second, and third-best results are marked in red, orange, and yellow, respectively.

Model	Case	Δ	Moral	Δ	Theory	Δ
Base	38.97	-	73.39	-	63.83	-
Ablation study on training stage						
+GRPO	36.69	-5.85%	77.74	5.93%	73.34	14.90%
+SFT	48.51	24.48%	83.82	14.21%	73.44	15.06%
+SFT+GRPO	67.07	72.11%	86.06	17.26%	79.10	23.92%
Ablation study on rationale optimization						
+QA	48.27	23.86%	81.89	11.58%	71.22	11.58%
+Rat.	55.25	41.78%	85.14	16.01%	75.55	18.36%
+Rat.+Iter.	67.07	72.11%	86.06	17.26%	79.10	23.92%

Table 3: Ablation study evaluating the effects of training stages (SFT and GRPO) and the contributions of rationale component (Rat.) and iterative prompt–rationale optimization (Iter.).

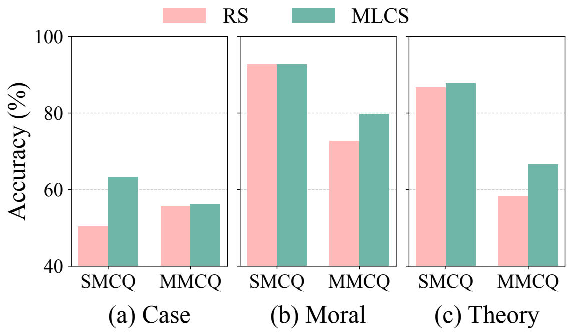

Figure 3: Comparison of performance using challenging question selection methods, including multi-LLM cross-selection (MLCS) and random selection (RS).

improves performance across the three task categories by leveraging our dataset of empathetic dialogues and rationale-augmented psychological questions. However, applying GRPO without prior SFT results in performance degradation on SMCQ case tasks, as the base model lacks sufficient domain knowledge and empathy, which are critical for reasoning in case-based questions, leading to unstable training dynamics. When combining SFT with GRPO training yields further gains, particularly on case-based tasks, as challenging samples identified via multi-LLM cross-selection promote deeper reasoning and contextual understanding.

Effect of the Rationale and Iterative Optimization. We explore the contributions of rationales and iterative prompt–rationale optimization, with results presented in Table 3. Note that **+QA** is the model trained solely on question–answer pairs without incorporating detailed rationales. Compared with the base model, training with our proposed dataset (i.e., **+QA**, **+Rat.** and **+Rat.+Iter.**) leads to consistent performance improvements, demonstrating the effectiveness of the curated data. Integrating rationale-augmented data substantially enhances performance over training with option-only answers, indicating that rationales provide valuable intermediate reasoning signals that facilitate learning. Furthermore, applying iterative prompt–rationale optimization (i.e., **+Rat.+Iter.**) yields further gains, confirming that progressively refined rationales contribute to better supervision and more robust model reasoning.

Effect of the Question Selection. We further examine the effect of question selection by comparing multi-LLM cross-selection (MLCS) with random selection (RS), as illustrated in Figure 3. The comparison between MLCS and RS demonstrates that leveraging multiple LLMs for selecting challenging instances yields markedly superior outcomes across all these tasks. This finding confirms that our se-

Settings	Case		Moral		Theory		Avg.
	SMCQ	MMCQ	SMCQ	MMCQ	SMCQ	MMCQ	
Base	47.57	31.64	87.83	59.50	78.46	42.45	57.91
+ED	35.77	29.74	70.00	60.90	65.84	44.40	51.11
+PRD	37.94	35.45	91.45	51.27	80.47	37.53	55.69
+ED+PRD	61.71	53.56	92.72	76.58	86.13	68.16	73.14
+ED+PRD+APD	63.31	56.26	92.76	79.62	87.70	66.54	74.37

Table 4: Effect of different dataset compositions, including empathetic dialogues (ED, i.e., $\mathcal{D}_{em}$), psychological reasoning data (PRD, i.e., $\mathcal{D}_{pc} \cup \mathcal{D}_{pr}$), and additional public datasets (APD, i.e., $\mathcal{D}_{ps} \cup \mathcal{D}_{cp}$).

Model	EmoE.	CogE.	Con.	Sta.
Qwen2.5-7B-Instruct	1.52	2.00	2.36	1.72
CPsyCounX	1.73	2.05	2.15	1.96
EmoLLM	1.86	2.44	**2.84**	**2.34**
PsycoLLM	1.97	2.27	2.41	2.10
PsyDT	2.21	2.46	2.36	**2.34**
Psyche-R1	**2.33**	**2.69**	2.78	2.11

Table 5: Comparisons of psychological LLMs on PsyDT test set. The evaluation metrics comprise: emotional empathy (EmoE.), cognitive empathy (CogE.), conversation strategy (Con.), and state and attitude (Sta.).

lection mechanism effectively selects high-quality training data for GRPO, which is instrumental in enhancing reasoning and generalization within the psychological domain.

3.4 Discussion

Effect of Datasets. We evaluate the model performance across different combinations of subsets, with results presented in Table 4. It is observed that fine-tuning with either psychological reasoning data (PRD) or empathetic dialogues (ED) in isolation delivered marginal improvements in task performance, and in some cases, led to a slight decline in overall accuracy. In contrast, the combination of PRD and ED achieves substantial improvements across these tasks, highlighting the quality and comprehensiveness of our proposed data synthesis pipeline. This result demonstrates that integrating domain-specific knowledge with emotional understanding enhances psychological reasoning. Moreover, the incorporation of additional public datasets (APD) leads to further performance improvements.

Performance on Counseling Tasks. Beyond examination tasks, we evaluate the performance of *Psyche-R1* on counseling tasks and compare it with its base model and several outstanding psychological LLMs. Following the method of PsyDT (Xie et al., 2025) but constrained by limited resources, we randomly sample 200 instances from its test set

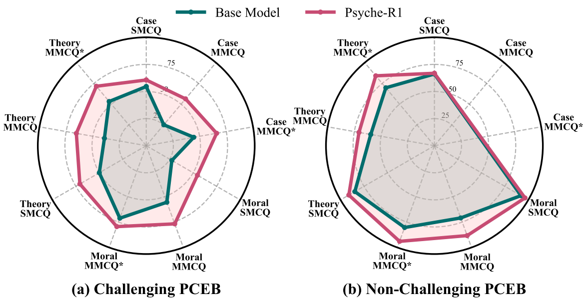

Figure 4: Comparison of model performance on (a) Challenging and (b) Non-Challenging subsets of the PCEB. An asterisk(*) indicates elastic accuracy, while the remaining metrics represent standard accuracy.

and employ GPT-4o as the evaluator. As shown in Table 5, *Psyche-R1* achieves notable improvements compared to its base model, demonstrating its capability in counseling tasks that demand emotional empathy, cognitive empathy and so on. This excellent performance stems from the synergistic interplay between two crucial elements: the empathetic dialogues, which directly improve counseling effectiveness, and advanced reasoning mechanisms, which enable a deeper understanding of questions, thereby yielding more accurate and emotionally informed responses within relevant contexts.

Error Analysis. To assess the impact of our approach across varying difficulty levels, we divide the PCEB into (a) challenging and (b) non-challenging subsets following the question selection method described in §2.2. The results are presented in Figure 4. It is observed that *Psyche-R1* demonstrates consistent improvements across these tasks in both subsets. For the non-challenging one, performance gains are primarily concentrated in the theory and moral dimensions, reflecting the model's proficiency in handling foundational psychological concepts. Notably, our model exhibits more pronounced improvements on the challenging subset. This observation can be attributed to the synergistic effect of our data generation pipeline and hybrid training strategy.

4 Conclusion

In this paper, we propose *Psyche-R1*, the first Chinese psychological LLM that jointly integrates empathy, expertise, and reasoning. To support model development, we design a multi-stage data synthesis pipeline that generates high-quality psychological reasoning samples with detailed rationales and empathetic dialogues. The reasoning rationales are further enhanced through iterative prompt–rationale optimization, and a multi-LLM cross-selection strategy is employed to identify challenging examples. Finally, the challenging subset is used for GRPO, while the remaining data are employed for SFT, together contributing to the final model. Extensive experiments demonstrate that *Psyche-R1* outperforms existing psychological LLMs, achieving performance comparable to `DeepSeek-R1`. Moreover, we perform comprehensive ablation studies and analyses to evaluate the individual contributions of each component and strategy within the proposed framework.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China under Grant 62402158, Grant 62502145 and Grant 62272144; by the Key Science & Technology Project of Anhui Province (202523j08050001); by the National College Students' Innovation and Entrepreneurship Training Program (202510359110); by the Anhui Provincial Natural Science Foundation Grant 2408085QF188, and Grant 2408085J040; by the Fundamental Research Funds for the Central Universities Grant JZ2025HGTA0162, and Grant JZ2025HGQA0134; and by the Major Project of Anhui Provincial Science and Technology Breakthrough Program (202423k09020001).

Limitations

Despite the promising results of our *Psyche-R1*, our study is subject to several limitations that remain to be addressed in future research.

Language and Cultural Specificity. To mitigate the shortage of mental health professionals in China, current *Psyche-R1* and its training corpus are predominantly tailored to the Chinese language and cultural context. Consequently, the model's empathetic reasoning involves specific cultural norms that may not directly transfer to other languages (Adilazuarda et al., 2024; Dai et al., 2026). We frame this as a necessary step for local applicability, noting that cross-cultural generalization remains challenging for future research. Moreover, the integration of multimodal psychology and affective computing warrants further study (Song et al., 2024; Liao et al., 2026).

Model Scale. Constrained by computational resources, *Psyche-R1* is built upon a 7B-parameter backbone. While it achieves competitive performance, we posit that employing a base model with a larger scale would yield superior performance. In future work, more efficient fine-tuning strategies (Hu et al., 2022; Zhang et al., 2025a) may help improve training efficiency and facilitate scaling to stronger backbones under limited resources.

Ethical Considerations

The development and deployment of LLMs in the mental health domain necessitate rigorous adherence to ethical standards.

Nature of the System. *Psyche-R1* is designed as a supportive tool for mental health support and education, rather than a replacement for qualified mental health professionals. The model is not authorized to provide medical diagnoses, prescribe treatments, or handle crisis interventions. Users facing severe mental health crises should seek help from human professionals or emergency services.

Data Privacy and Safety. We prioritize the privacy and safety of individuals in our data curation process. For data collected from social media platforms (Type IV), we implemented strict de-identification procedures to remove all personally identifiable information, including names, locations, and contact details. We strictly adhere to data usage policies and ensure that the synthesized data does not reconstruct real-world private interactions. Furthermore, our data synthesis pipeline that prioritizes high-quality, constructive psychological advice filtered out toxic content and harmful suggestions to align with safety guidelines.

References

Muhammad Farid Adilazuarda, Sagnik Mukherjee, Pradhyumna Lavania, Siddhant Shivdutt Singh, Alham Fikri Aji, Jacki O'Neill, Ashutosh Modi, and Monojit Choudhury. 2024. Towards measuring and modeling "culture" in LLMs: A survey. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 15763–15784, Miami, Florida, USA. Association for Computational Linguistics.

Nafiz Al Asad, Md. Appel Mahmud Pranto, Sadia Afreen, and Md. Maynul Islam. 2019. Depression detection by analyzing social media posts of user. In *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, pages 13–17.

Junying Chen, Zhenyang Cai, Ke Ji, Xidong Wang, Wanlong Liu, Rongsheng Wang, Jianye Hou, and Benyou Wang. 2024. Huatuogpt-o1, towards medical complex reasoning with llms. *Preprint*, arXiv:2412.18925.

Yirong Chen, Xiaofen Xing, Jingkai Lin, Huimin Zheng, Zhenyu Wang, Qi Liu, and Xiangmin Xu. 2023a. 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.

Zhiyu Chen, Yujie Lu, and William Wang. 2023b. Empowering psychotherapy with large language models: Cognitive distortion detection through diagnosis of thought prompting. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 4295–4304, Singapore. Association for Computational Linguistics.

Yihang Cheng, Hao Meng, and Wei Zhang. 2024. Collaborative agents for anxious depression diagnosis and intervention leverage neuroimages in human brain-ai interactions. In *2024 International Conference on Intelligent Education and Intelligent Research (IEIR)*, pages 1–5.

Yujin Cho, Mingeon Kim, Seojin Kim, Oyun Kwon, Ryan Donghan Kwon, Yoonha Lee, and Dohyun Lim. 2023. Evaluating the efficacy of interactive language therapy based on llm for high-functioning autistic adolescent psychological counseling. *Preprint*, arXiv:2311.09243.

Gheorghe Comanici, Eric Bieber, Mike Schaekermann, Ice Pasupat, Noveen Sachdeva, Inderjit Dhillon, Marcel Blistein, Ori Ram, Dan Zhang, Evan Rosen, Luke Marris, Sam Petulla, Colin Gaffney, Asaf Aharoni, Nathan Lintz, Tiago Cardal Pais, Henrik Jacobsson, Idan Szpektor, Nan-Jiang Jiang, and 3416 others. 2025. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *Preprint*, arXiv:2507.06261.

Chongyuan Dai, Yaling Shen, Jinpeng Hu, Zihan Gao, Jia Li, Yishun Jiang, Yaxiong Wang, Liu Liu, and Zongyuan Ge. 2026. Tears or cheers? benchmarking llms via culturally elicited distinct affective responses. *Preprint*, arXiv:2601.13024.

Dorottya Demszky, Diyi Yang, David S Yeager, Christopher J Bryan, Margarett Clapper, Susannah Chandhok, Johannes C Eichstaedt, Cameron Hecht, Jeremy Jamieson, Meghann Johnson, and 1 others. 2023. Using large language models in psychology. *Nature Reviews Psychology*, 2(11):688–701.

Luyu Gao, Aman Madaan, Shuyan Zhou, Uri Alon, Pengfei Liu, Yiming Yang, Jamie Callan, and Graham Neubig. 2023. PAL: Program-aided language models. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 10764–10799. PMLR.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Peiyi Wang, Qihao Zhu, Runxin Xu, Ruoyu Zhang, Shirong Ma, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z. F. Wu, Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, and 175 others. 2025. Deepseek-r1 incentivizes reasoning in llms through reinforcement learning. *Nature*, 645(8081):633–638.

Cheng-Yu Hsieh, Chun-Liang Li, Chih-kuan Yeh, Hootan Nakhost, Yasuhisa Fujii, Alex Ratner, Ranjay Krishna, Chen-Yu Lee, and Tomas Pfister. 2023. Distilling step-by-step! outperforming larger language models with less training data and smaller model sizes. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 8003–8017, Toronto, Canada. Association for Computational Linguistics.

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Jinpeng Hu, Tengteng Dong, Gang Luo, Hui Ma, Peng Zou, Xiao Sun, Dan Guo, Xun Yang, and Meng Wang. 2025a. Psycollm: Enhancing llm for psychological understanding and evaluation. *IEEE Transactions on Computational Social Systems*, 12(2):539–551.

Jinpeng Hu, Hongchang Shi, Chongyuan Dai, Zhuo Li, Peipei Song, and Meng Wang. 2025b. Beyond emotion recognition: A multi-turn multimodal emotion understanding and reasoning benchmark. In *Proceedings of the 33rd ACM International Conference on Multimedia*, MM '25, page 5814–5823, New York, NY, USA. Association for Computing Machinery.

Jinpeng Hu, Ao Wang, Qianqian Xie, Zhuo Li, Hui Ma, and Dan Guo. 2026. Agentmental: An interactive multi-agent framework for explainable and adaptive mental health assessment. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(37):31050–31058.

Zhaocheng Huang, Julien Epps, Dale Joachim, and Vidhyasaharan Sethu. 2020. Natural language processing methods for acoustic and landmark event-based features in speech-based depression detection. *IEEE Journal of Selected Topics in Signal Processing*, 14(2):435–448.

Tatsuki Kuribayashi, Yohei Oseki, and Timothy Baldwin. 2024. Psychometric predictive power of large language models. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 1983–2005, Mexico City, Mexico. Association for Computational Linguistics.

Tin Lai, Yukun Shi, Zicong Du, Jiajie Wu, Ken Fu, Yichao Dou, and Ziqi Wang. 2024. Supporting the demand on mental health services with ai-based conversational large language models (llms). *BioMedInformatics*, 4(1):8–33.

Daeun Lee, Soyoung Park, Jiwon Kang, Daejin Choi, and Jinyoung Han. 2020. Cross-lingual suicidal-oriented word embedding toward suicide prevention. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2208–2217, Online. Association for Computational Linguistics.

Suyeon Lee, Sunghwan Kim, Minju Kim, Dongjin Kang, Dongil Yang, Harim Kim, Minseok Kang, Dayi Jung, Min Hee Kim, Seungbeen Lee, Kyong-Mee Chung, Youngjae Yu, Dongha Lee, and Jinyoung Yeo. 2024. Cactus: Towards psychological counseling conversations using cognitive behavioral theory. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14245–14274, Miami, Florida, USA. Association for Computational Linguistics.

Junjie Liao, Jiandian Zeng, Binbin Song, Mengting Zhou, Xiaopeng Fan, and Tian Wang. 2026. Unlocking explainable and effective multimodal affective reasoning via large language models. *Pattern Recognition*, 178:113366.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Che Liu, Haozhe Wang, Jiazhen Pan, Zhongwei Wan, Yong Dai, Fangzhen Lin, Wenjia Bai, Daniel Rueckert, and Rossella Arcucci. 2025. Beyond distillation: Pushing the limits of medical LLM reasoning with minimalist rule-based RL. In *The Second Workshop on GenAI for Health: Potential, Trust, and Policy Compliance*.

Runze Liu, Jiakang Wang, Yuling Shi, Zhihui Xie, Chenxin An, Kaiyan Zhang, Jian Zhao, Xiaodong Gu, Lei Lin, Wenping Hu, Xiu Li, Fuzheng Zhang, Guorui Zhou, and Kun Gai. 2026. Attention as a compass: Efficient exploration for process-supervised RL in reasoning models. In *The Fourteenth International Conference on Learning Representations*.

Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad Saqib, Saeed Anwar, Muhammad Usman, Naveed Akhtar, Nick Barnes, and Ajmal Mian. 2025. A comprehensive overview of large language models. *ACM Trans. Intell. Syst. Technol.*, 16(5).

OpenAI, :, Aaron Hurst, Adam Lerer, Adam P. Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, Aleksander Mądry, Alex Baker-Whitcomb, Alex Beutel, Alex Borzunov, Alex Carney, Alex Chow, Alex Kirillov, and 401 others. 2024a. Gpt-4o system card. *Preprint*, arXiv:2410.21276.

OpenAI, :, Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, Alex Iftimie, Alex Karpenko, Alex Tachard Passos, Alexander Neitz, Alexander Prokofiev, Alexander Wei, Allison Tam, and 244 others. 2024b. Openai o1 system card. *Preprint*, arXiv:2412.16720.

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, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

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.

Zhihong Shao, Peiyi Wang, Qihao Zhu, Runxin Xu, Junxiao Song, Xiao Bi, Haowei Zhang, Mingchuan Zhang, Y. K. Li, Y. Wu, and Daya Guo. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *Preprint*, arXiv:2402.03300.

Hao Shen, Zihan Li, Minqiang Yang, Minghui Ni, Yongfeng Tao, Zhengyang Yu, Weihao Zheng, Chen Xu, and Bin Hu. 2024. Are large language models possible to conduct cognitive behavioral therapy? In *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 3695–3700.

Yaling Shen, Stephanie Fong, Yiwen Jiang, Zimu Wang, Feilong Tang, Qingyang Xu, Xiangyu Zhao, Zhongxing Xu, Jiahe Liu, Jinpeng Hu, Dominic Dwyer, and Zongyuan Ge. 2026. Psychethicsbench: Evaluating large language models against australian mental health ethics. *Preprint*, arXiv:2601.03578.

Guangming Sheng, Chi Zhang, Zilingfeng Ye, Xibin Wu, Wang Zhang, Ru Zhang, Yanghua Peng, Haibin Lin, and Chuan Wu. 2025. Hybridflow: A flexible and efficient rlhf framework. In *Proceedings of the Twentieth European Conference on Computer Systems*, EuroSys '25, page 1279–1297, New York, NY, USA. Association for Computing Machinery.

Yuling Shi, Maolin Sun, Zijun Liu, Mo Yang, Yixiong Fang, Tianran Sun, and Xiaodong Gu. 2026a. Reasoning in trees: Improving retrieval-augmented generation for multi-hop question answering. *Preprint*, arXiv:2601.11255.

Yuling Shi, Chaoxiang Xie, Zhensu Sun, Yeheng Chen, Chenxu Zhang, Longfei Yun, Chengcheng Wan, Hongyu Zhang, David Lo, and Xiaodong Gu. 2026b. Codeocr: On the effectiveness of vision language models in code understanding. *Preprint*, arXiv:2602.01785.

Peipei Song, Dan Guo, Xun Yang, Shengeng Tang, and Meng Wang. 2024. Emotional video captioning with vision-based emotion interpretation network. *IEEE Transactions on Image Processing*, 33:1122–1135.

Vera Sorin, Dana Brin, Yiftach Barash, Eli Konen, Alexander Charney, Girish Nadkarni, and Eyal Klang. 2024. Large language models and empathy: Systematic review. *J Med Internet Res*, 26:e52597.

Michael J Tanana, Christina S Soma, Patty B Kuo, Nicolas M Bertagnolli, Aaron Dembe, Brian T Pace, Vivek Srikumar, David C Atkins, and Zac E Imel. 2021. How do you feel? using natural language processing to automatically rate emotion in psychotherapy. *Behavior research methods*, 53(5):2069–2082.

EmoLLM Team. 2024a. Emollm: Reinventing mental health support with large language models. `https://github.com/SmartFlowAI/EmoLLM`.

Gemini Team, Petko Georgiev, Ving Ian Lei, Ryan Burnell, Libin Bai, Anmol Gulati, Garrett Tanzer, Damien Vincent, Zhufeng Pan, Shibo Wang, Soroosh Mariooryad, Yifan Ding, Xinyang Geng, Fred Alcober, Roy Frostig, Mark Omernick, Lexi Walker, Cosmin Paduraru, Christina Sorokin, and 1118 others. 2024. Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context. *Preprint*, arXiv:2403.05530.

MiniCPM Team, Chaojun Xiao, Yuxuan Li, Xu Han, Yuzhuo Bai, Jie Cai, Haotian Chen, Wentong Chen, Xin Cong, Ganqu Cui, Ning Ding, Shengda Fan, Yewei Fang, Zixuan Fu, Wenyu Guan, Yitong Guan, Junshao Guo, Yufeng Han, Bingxiang He, and 64 others. 2025. Minicpm4: Ultra-efficient llms on end devices. *Preprint*, arXiv:2506.07900.

Qwen Team. 2024b. Qwen2.5: A party of foundation models.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, brian ichter, Fei Xia, Ed Chi, Quoc V Le, and Denny Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems*, volume 35, pages 24824–24837. Curran Associates, Inc.

JT Wolohan, Misato Hiraga, Atreyee Mukherjee, Zeeshan Ali Sayyed, and Matthew Millard. 2018. Detecting linguistic traces of depression in topic-restricted text: Attending to self-stigmatized depression with nlp. In *Proceedings of the first international workshop on language cognition and computational models*, pages 11–21.

Shurui Wu, Xinyi Huang, and Dingxin Lu. 2025. Psychological health knowledge-enhanced llm-based social network crisis intervention text transfer recognition method. In *Proceedings of the 2025 International Conference on Health Big Data*, HBD '25, page 156–161, New York, NY, USA. Association for Computing Machinery.

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.

Mengxi Xiao, Kailai Yang, Pengde Zhao, Enze Zhang, Ziyan Kuang, Zhiwei Liu, Weiguang Han, Shu Liao, Lianting Huang, Jinpeng Hu, Min Peng, Qianqian Xie, and Sophia Ananiadou. 2025. Mentrasuite: Post-training large language models for mental health reasoning and assessment. *Preprint*, arXiv:2512.09636.

Haojie Xie, Yirong Chen, Xiaofen Xing, Jingkai Lin, and Xiangmin Xu. 2025. PsyDT: Using LLMs to construct the digital twin of psychological counselor with personalized counseling style for psychological counseling. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1081–1115, Vienna, Austria. Association for Computational Linguistics.

Yangyang Xu, Jinpeng Hu, Peipei Song, Zhangling Duan, and Xun Yang. 2026. From social media to psychological scale: An adaptive framework with two-hop retrieval for depression screening. In *Proceedings of the ACM Web Conference 2026*, WWW '26, page 4817–4828, New York, NY, USA. Association for Computing Machinery.

Yangyang Xu, Jinpeng Hu, Zhuoer Zhao, Zhangling Duan, Xiao Sun, and Xun Yang. 2025. MultiAgentESC: A LLM-based multi-agent collaboration framework for emotional support conversation. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 4665–4681, Suzhou, China. Association for Computational Linguistics.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025. Qwen3 technical report. *Preprint*, arXiv:2505.09388.

Qisen Yang, Zekun Wang, Honghui Chen, Shenzhi Wang, Yifan Pu, Xin Gao, Wenhao Huang, Shiji Song, and Gao Huang. 2024. PsychoGAT: A novel psychological measurement paradigm through interactive fiction games with LLM agents. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14470–14505, Bangkok, Thailand. Association for Computational Linguistics.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. Tree of thoughts: Deliberate problem solving with large language models. In *Advances in Neural Information Processing Systems*, volume 36, pages 11809–11822. Curran Associates, Inc.

Jing Ye, Lu Xiang, Yaping Zhang, and Chengqing Zong. 2025. SweetieChat: A strategy-enhanced role-playing framework for diverse scenarios handling emotional support agent. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 4646–4669, Abu Dhabi, UAE. Association for Computational Linguistics.

Qiying Yu, Zheng Zhang, Ruofei Zhu, Yufeng Yuan, Xiaochen Zuo, YuYue, Weinan Dai, Tiantian Fan, Gaohong Liu, Juncai Liu, LingJun Liu, Xin Liu, Haibin Lin, Zhiqi Lin, Bole Ma, Guangming Sheng, Yuxuan Tong, Chi Zhang, Mofan Zhang, and 17 others. 2025. DAPO: An open-source LLM reinforcement learning system at scale. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah Goodman. 2022. Star: Bootstrapping reasoning with reasoning. In *Advances in Neural Information Processing Systems*, volume 35, pages 15476–15488. Curran Associates, Inc.

Chenhao Zhang, Renhao Li, Minghuan Tan, Min Yang, Jingwei Zhu, Di Yang, Jiahao Zhao, Guancheng Ye, Chengming Li, and Xiping Hu. 2024. CPsyCoun: A report-based multi-turn dialogue reconstruction and evaluation framework for Chinese psychological counseling. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13947–13966, Bangkok, Thailand. Association for Computational Linguistics.

Hao Zhang, Zhenjia Li, Runfeng Bao, Yifan Gao, Xi Xiao, Heng Zhang, Shuyang Zhang, Bo Huang, Yuhang Wu, Tianyang Wang, and Hao Xu. 2025a. Hyperadalora: Accelerating lora rank allocation during training via hypernetworks without sacrificing performance. *Preprint*, arXiv:2510.02630.

Hongbo Zhang, Junying Chen, Feng Jiang, Fei Yu, Zhihong Chen, Guiming Chen, Jianquan Li, Xiangbo Wu, Zhang Zhiyi, Qingying Xiao, Xiang Wan, Benyou Wang, and Haizhou Li. 2023. HuatuoGPT, towards taming language model to be a doctor. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 10859–10885, Singapore. Association for Computational Linguistics.

Junlei Zhang, Hongliang He, Lizhi Ma, Nirui Song, Shuyuan He, Shuai Zhang, Huachuan Qiu, Zhanchao Zhou, Anqi Li, Yong Dai, Renjun Xu, and Zhenzhong Lan. 2025b. Conceptpsy: A comprehensive benchmark suite for hierarchical psychological concept understanding in llms. *Neurocomputing*, 637:130070.

Wei Zhang, Yihang Cheng, Zhirong Ye, and Kezhen Huang. 2026a. Cogevolution: A human-like generative educational agent to simulate student's cognitive evolution. *Preprint*, arXiv:2604.14786.

Xinglang Zhang, Yunyao Zhang, ZeLiang Chen, Junqing Yu, Wei Yang, and Zikai Song. 2026b. Logical phase transitions: Understanding collapse in llm logical reasoning. *Preprint*, arXiv:2601.02902.

Yunyao Zhang, Xinglang Zhang, Junxi Sheng, Wenbing Li, Junqing Yu, Yi-Ping Phoebe Chen, Wei Yang, and Zikai Song. 2026c. Semantic-aware logical reasoning via a semiotic framework. *Preprint*, arXiv:2509.24765.

Jiahao Zhao, Jingwei Zhu, Minghuan Tan, Min Yang, Renhao Li, Yang Di, Chenhao Zhang, Guancheng Ye, Chengming Li, Xiping Hu, and Derek F. Wong. 2025. CPsyExam: A Chinese benchmark for evaluating psychology using examinations. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 11248–11260, Abu Dhabi, UAE. Association for Computational Linguistics.

Chujie Zheng, Sahand Sabour, Jiaxin Wen, Zheng Zhang, and Minlie Huang. 2023. AugESC: Dialogue augmentation with large language models for emotional support conversation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 1552–1568, Toronto, Canada. Association for Computational Linguistics.

Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, and Zheyan Luo. 2024. LlamaFactory: Unified efficient fine-tuning of 100+ language models. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 400–410, Bangkok, Thailand. Association for Computational Linguistics.

Yixiao Zhou, Yang Li, Dongzhou Cheng, Hehe Fan, and Yu Cheng. 2026. Look inward to explore outward: Learning temperature policy from llm internal states via hierarchical rl. *Preprint*, arXiv:2602.13035.

Jie Zhu, Qian Chen, Huaixia Dou, Junhui Li, Lifan Guo, Feng Chen, and Chi Zhang. 2025. Dianjin-r1: Evaluating and enhancing financial reasoning in large language models. *Preprint*, arXiv:2504.15716.

A Related Work

A.1 LLMs for Psychology

The success of LLMs has spurred interest in developing LLM-driven mental health applications (Demszky et al., 2023; Xiao et al., 2025; Shen et al., 2026). Early research focused primarily on improving the accessibility of mental health services. Research in this phase primarily concentrated on two directions: One direction involves leveraging NLP techniques for emotion recognition to enable automated detection of depression (Huang et al., 2020) and suicidal ideation (Lee et al., 2020). The other focuses on constructing empathetic dialogue systems by fine-tuning LLMs on single-turn (Lai et al., 2024) or multi-turn (Qiu et al., 2024) dialogue data to enhance their abilities in affective understanding and emotional support (Team, 2024a; Xie et al., 2025). As research progressed, researchers began to explore more diverse mental health applications. Some studies have transformed traditional psychometric tools (e.g., psychological scales) into interactive systems to improve user engagement (Kuribayashi et al., 2024; Yang et al., 2024; Xu et al., 2025; Hu et al., 2026). Another line of research has focused on the specialized demands of the psychological domain, developing professional-grounded mental health applications based on established psychological therapies (Lee et al., 2024; Shen et al., 2024) or concepts (Zhang et al., 2025b).

A.2 LLM Reasoning

In recent years, techniques such as CoT prompting (Wei et al., 2022; Hsieh et al., 2023) and STaR (Zelikman et al., 2022) have significantly advanced the development of LLM reasoning. Building upon this foundation, researchers have explored more sophisticated reasoning architectures (Zhang et al., 2026c,b; Shi et al., 2026a). For instance, Tree of Thoughts (Yao et al., 2023) enables systematic exploration of multiple reasoning paths with self-evaluation, while PAL (Gao et al., 2023) integrates reasoning with external tools through program generation. These approaches further enhance model performance in handling complex tasks. A new breakthrough was marked by the release of reasoning LLMs such as `OpenAI o1` (OpenAI et al., 2024b) and `DeepSeek-R1` (Guo et al., 2025). These models, which are trained through reinforcement learning with reasoning techniques to enhance reasoning capabilities, demonstrate exceptional performance in mathematical and coding tasks (Comanici et al., 2025; Yang et al., 2025). Motivated by these advances, researchers have employed advanced RL algorithms (Shao et al., 2024; Yu et al., 2025; Zhou et al., 2026) to further extend reasoning capabilities to domain-specific applications, including medicine (Liu et al., 2025) and finance (Zhu et al., 2025). However, within the field of psychology, limited research has investigated the utility of reasoning. To our knowledge, *Psyche-R1* is the first psychological LLM that unifies empathy, domain-specific expertise, and reasoning capabilities.

B Case Study

We present a case study examining how *Psyche-R1* and `Qwen2.5-72B-Instruct` formulate their conclusions derived from narratives and deliver mental health support, as illustrated in Figure 5. These two models display distinct counseling strategies when addressing the case involving a company manager confronting a career transition dilemma. *Psyche-R1* begins by expressing empathy (e.g., "I can really sense..."), followed by applying relevant psychological concepts tailored to user's situations, thereby demonstrating both emotional attunement and domain-specific expertise. In contrast, `Qwen`'s empathetic expressions appear less natural and engaging (e.g., "Your situation is truly understandable..."), and it fails to apply theoretical knowledge to contextualize or explain the user's dilemma, which undermines the credibility of its analysis and recommendations. Moreover, *Psyche-R1* exhibits a clear and efficient reasoning path progressing from surface-level observations to in-depth analysis, whereas `Qwen` merely enumerates generic suggestions lacking step-by-step and in-depth reasoning. For the Chinese version of this case, see Figure 6.

C Details of Experiments

C.1 Details of Baselines

We compare *Psyche-R1* with four categories of LLMs, including: (1) **General LLMs,** including `MiniCPM4-8B` (Team et al., 2025), `Qwen2.5-7B/14B/72B` (Team, 2024b). (2) **Reasoning augmented LLMs,** encompassing `DeepSeek-R1` (Guo et al., 2025), `DeepSeek-R1-70B`, `QwQ-32B`, `Qwen3-30B-A3B`, `Qwen3-235B-A22B` (Yang et al., 2025), and `Magistral-Small`. (3) **Closed-source LLMs,** including `Claude3.7-Sonnet`, `Gemini1.5-Pro` (Team et al., 2024), and `GPT-4o` (OpenAI et al., 2024a). (4) **Psychological LLMs,** including `CPsyCounX` (Zhang et al., 2024), `EmoLLM` (Team, 2024a), `PsycoLLM` (Hu et al., 2025a), and `PsyDT` (Xie et al., 2025). Notice that for the hybrid reasoning-augmented models `Qwen3` series and `Claude3.7-Sonnet`, we set them to reasoning mode to stimulate their best performance. Details of the model information are provided in Table 7.

C.2 Implementation Details

In our experiments, we employ the LLaMA-Factory (Zheng et al., 2024) framework for SFT. Specifically, we adopt a learning rate of 1e-5, a batch size of 256, and conduct training for 2 epochs. For the GRPO phase, we implement the VeRL framework (Sheng et al., 2025) with a learning rate of 1e-6, a batch size of 128, and 2 training epochs. All experiments are performed on 8 RTX A6000 GPUs, each equipped with 48GB.

During evaluation, we set temperature to 0.0, maximum sequence lengths to 1024, and top-p to 0.95 to ensure the fairness of evaluation.

Case Question …I became general manager at the age of 30, was dedicated and conscientious in my work... Six months ago... I was not rehired. The newly appointed manager is only 32 years old... I feel extremely wronged and treated unjustly... I can't muster any energy all day long, have no appetite, and nothing tastes good anymore… When I see colleagues talking, I always suspect they're gossiping about me… I don't even want to go out anymore. I've declined all social gatherings… I really can't accept failing like this. This situation truly causes me great pain. Can you help me relieve this suffering?

Psyche-R1 …I can really sense the internal struggle and pain you are experiencing, which in psychology is referred to as "role transition" and "identity crisis"… This sudden shift in identity has left you feeling lost and defeated… Your emotional responses… are all common psychological reactions when facing major life changes. These symptoms indicate that you are experiencing adjustment disorder and... **First**, accepting reality is key... You need to learn to view your worth from a broader perspective… **Second**, adjust your mindset and maintain a positive attitude… **Third**, establish a strong social support system… **Finally**, if symptoms persist and affect your daily life, it is suggested that seeking help from professional psychologists…

Qwen2.5-72B-Instruct Your situation is truly understandable and deserves sympathy… **Accept Reality:** First, try to accept the current situation... **Seek Support:** Share your feelings with trusted friends or family members... **Adjust Your Mindset:** Maintaining a positive and optimistic attitude is crucial for overcoming difficulties. **Focus on Personal Growth:** Use this time to concentrate on personal interests and development… **Improve Lifestyle Habits:** Pay attention to healthy eating and... **Build Your Social Network:** … maintaining some social activities is very helpful for improving your mood... **Make a Plan:** Think about your next career move… Everyone's feelings when facing difficulties are real and valid…

Figure 5: A qualitative example from the CPsyExam test set comparing *Psyche-R1* and Qwen2.5-72B-Instruct. Highlights indicate empathetic expressions (red), psychological expertise (green), and reasoning (blue).

Case Question …三十岁就当上了总经理，对工作兢兢业业…半年前国企改制…我居然落聘了。新聘任的总经理才32岁…我受不了这种气…我烦，我气…一天到晚打不起精神，也没什么胃口，吃什么都不香…看到同事们在说话，总怀疑是不是在议论我…我都不想出门了，各种聚会我都给推掉了…我真不甘心就这样失败了，这事确实让我很痛苦。您能帮我解除痛苦吗？

Psyche-R1 …我能感受到你内心的挣扎和痛苦，这在心理学上被称为"角色转换"和"自我认同危机"…这种身份的突然转变让你感到失落和挫败…你的情绪反应…都是在面对重大生活变化时常见的心理反应。这些症状表明你正在经历心理适应障碍……**首先**，接受现实是关键…你需要学会从更广阔的视角看待自己的价值…**其次**，调整心态，保持积极的心态…**再次**，建立良好的社会支持系统…**最后**，如果症状持续存在并影响日常生活，建议寻求专业心理医生的帮助……

Qwen2.5-72B-Instruct 您的情况确实让人感到同情……**接受现实：**首先，尝试接受当前的情况…**寻求支持：**与信任的朋友或家人分享您的感受…**调整心态：**…保持积极乐观的态度对于克服困难至关重要。**专注于个人成长：**利用这段时间专注于个人兴趣和发展…**改善生活习惯：**注意饮食健康…**建立社交网络：**…维持一定的社交活动对恢复心情很有帮助…**制定计划：**思考接下来的职业规划…每个人面对困境时的感受都是真实合理的……

Figure 6: The Chinese version of the qualitative example presented in the Case Study. Highlights indicate empathetic expressions (red), psychological expertise (green), and reasoning (blue).

C.3 Details of Model Training

For SFT training, the hyperparameters utilized for training the model are configured as follows: the learning rate is set to 1e-05, the batch size is set to 256, the number of epochs is set to 2. We employ AdamW as the optimizer, configured with the epsilon set to 1e-08. The learning rate scheduler is set to cosine type with a warmup ratio of 0.1.

For GRPO training, the hyperparameters are configured as follows: the learning rate is set to 1e-06, the batch size is set to 128, and the number of epochs is set to 2. The PPO mini-batch size is configured to 32, with a micro-batch size per GPU of 20. We incorporate KL divergence regularization with the KL loss coefficient set to 1e-03, employing the low-variance KL loss type.

D Details of Prompts

We provide the prompts used throughout our data synthesis pipeline. Only the English version is presented due to compilation issues in LaTeX with non-English languages.

Prompts for Data Cleaning

You are a professional evaluator with extensive knowledge in psychology. Users on mental health platforms are facing difficulties in their lives, so they have provided questions and detailed descriptions and have received some responses from counselors. Please carefully analyze the given questions, descriptions, and responses, determine whether the responses are helpful to the users and have positive significance, and return "reasonable" or "unreasonable".

Prompts for Question Generation

You are an expert in designing psychology examination questions with extensive work experience. Your task is to generate {num_questions} clear and challenging psychology questions based on the text below. Do not add any information that is not mentioned in the provided text.
Note: When generating questions that reference the text, you must provide the detailed and complete textual evidence to offer suffi-

Dataset	Dialogue	Knowledge Question	CoT Rationales	Empathetic Dialogue	Expertise
CPSYCOUND (Zhang et al., 2024)	3,134	-	✗	✓	✗
PsyDTCorpus (Xie et al., 2025)	5,000	-	✗	✓	✗
SMILECHAT (Qiu et al., 2024)	55,165	-	✗	✓	✗
PsycoLLM (Hu et al., 2025a)	173k	9,106	✓	✗	✓
Ours	72,920	75,465	✓	✓	✓

Table 6: Comparison of psychological datasets.

Model	Param.	Version
MiniCPM4-8B	8B	openbmb/MiniCPM4-8B
Qwen2.5-7B	7B	Qwen/Qwen2.5-7B-Instruct
Qwen2.5-14B	14B	Qwen/Qwen2.5-14B-Instruct
Qwen2.5-72B	72B	Qwen/Qwen2.5-72B-Instruct
DeepSeek-R1	671B	deepseek-ai/DeepSeek-R1
DeepSeek-R1-70B	70B	deepseek-ai/DeepSeek-R1-Distill-Llama-70B
QwQ-32B	32B	Qwen/QwQ-32B
Qwen3-30B-A3B	30B	Qwen/Qwen3-30B-A3B
Qwen3-235B-A22B	235B	Qwen/Qwen3-235B-A22B
Magistral-Small	24B	mistralai/Magistral-Small-2506
GPT-4o	UNK	gpt-4o-2024-05-13
Gemini1.5-Pro	UNK	gemini-1.5-pro-latest
Claude3.7-Sonnet	UNK	claude-3-7-sonnet-20250219
CPsyCounX	7B	finetuned on Internlm-7B-Chat
EmoLLM	7B	finetuned on Qwen2-7B-Instruct
PsycoLLM	14B	finetuned on Qwen1.5-14B-Instruct
PsyDT	7B	finetuned on Qwen2-7B-Instruct

Table 7: Detailed information of baselines.

cient information.
 # Text Content: {text}
Please generate {num_questions} {type_instruction} questions based on the text above. These questions must be based on the text content, and you must ensure that the answers have clear evidence within the text. Please try to ensure diversity and variation among the generated questions.
You must strictly adhere to the following guidelines:

1. The questions should be challenging and require reasoning to test the candidate's reasoning skills and academic literacy, rather than being simple knowledge-recall questions.
2. The questions need to be clear, accurate, and well-structured, with reasonably set options and an appropriate distribution of difficulty.
3. Ensure that the questions and their corresponding answers have clear evidence in the text.

Follow the JSON format below to generate the questions:

```JSON
{
  "question": "...",
  "options": "...",
  "answer": "....",
  "type": "..."
}
```

You need to repeat the structure above to generate a total of {num_questions} questions.

Prompts for Question Control

You are an expert in psychology. Now, I have a batch of questions that were converted from book texts using large language models. However, some of these questions have missing information. Your task is to judge whether the following psychology questions are reasonable.
The criteria for judging question reasonableness are whether the question provides sufficient information for candidates to solve the problem. Since these questions are gen-

erated by large language models based on a batch of book texts, candidates can only see the questions and cannot access the original texts.
Therefore, a “reasonable” question should be: after reading the question, candidates can choose the correct answer from the options through deep thinking about the question content (i.e., the "question") combined with their existing knowledge, without needing to read the original text content. Conversely, an “unreasonable” question should be, there is missing information, and without reading the original text, it is impossible to choose the correct answer based solely on the question and one’s own knowledge.
Note: you need to carefully read the question, understand its content, and ensure that you give an accurate judgment!
 # Examples: {examples}
Now, please follow the above guidelines to judge whether the following question is reasonable. Note that you only need to return “reasonable” or “unreasonable” without any other text content:
 # Question Type: {type}
 # Question: {question}

Prompts for Rationale Generation (for rationale generation)

You are an expert in psychology with extensive professional experience.
Please carefully read the following psychology question, analyze and reason through it using psychological knowledge, and explain your reasoning step by step along with your final predicted answer. This requires comprehensive analysis, summarization, exploration, re-evaluation, reflection, backtracking, and iteration to develop a thoughtful reasoning process. In the reasoning section, each of your reasoning steps should be considered in detail from a professional psychological perspective, such as analyzing the problem, summarizing relevant findings, brainstorming, verifying the accuracy of the current step, improving any errors, and revisiting previous steps.
Now, you must follow the JSON format below to present your rationale and prediction:

```JSON
{
  "rationale": "...",
  "prediction": "..."
}
```

 # Question: {question}

Prompts for Rationale Generation (for candidate prompt generation)

You are an expert in prompt optimization with extensive professional experience.
Based on the following psychological question and initial prompt, please generate a better prompt to guide large language models to conduct more accurate and detailed analysis and reasoning for **this question**.
 # Current Prompt: {current_prompt}
 # Question: {question}

Prompts for Rationale Generation (for rationale comparison)

You are an expert in psychology exam grading with extensive work experience.
Below are different responses to the same psychology question. You need to objectively, thoroughly, and comprehensively evaluate these responses, ultimately choose the best one from among them, and provide your detailed explanation for the choice.
 # Rationales: {rationale_1} ... {rationale_n} ...
Please return your selection in the following JSON format:

```JSON
{
  "best_rational_index": "...",
  "reason": "..."
}
```

Prompts for Question Selection

You are participating in a psychology exam. Please choose an answer based on the provided question and options. Directly output the letter of the option. No explanation is needed.

 # Question: {question}
Please present the predicted answer directly with the letter of the option. No explanation is needed.

Prompts for Empathetic Dialogue Synthesis

 # Role: You are a psychological counselor with extensive theoretical knowledge and counseling experience. You possess strong empathy and compassion, keen observational skills, excellent listening abilities, and conversational techniques. Your aim is to help users improve their mood and overcome difficulties.
 # Your tasks are: Since users' questions commonly contain issues like inappropriate expressions and logical confusion, making the questions often unclear, you need to:

1. **Organize the sequence of events:** conduct detailed analysis of the context and content within the problem;
2. **Understand psychological confusion:** you need to combine your psychological knowledge and counseling experience to uncover the mental issues and states within the user's question;
3. **Adopt the user's perspective and refine the question:** You must refine the question from user's first-person perspective. Based on the given question, you need to polish and organize it into a complete, logically clear, and sufficiently detailed expression. This expression should highlight the user's psychological confusion or mental state to provide adequate substantive content.

Note: You only need to return the refined question without providing any other irrelevant text! Now, try to address the following problem using the above guidelines: