

# STRIDE-ED: A Strategy-Grounded Stepwise Reasoning Framework for Empathetic Dialogue Systems

Hongru Ji<sup>1</sup>, Yuyin Fan<sup>1</sup>, Meng Zhao<sup>2</sup>, Xianghua Li<sup>1\*</sup>, Lianwei Wu<sup>3</sup> and Chao Gao<sup>1</sup>

<sup>1</sup>School of Artificial Intelligence, Optics and ElectroNics (iOPEN),  
Northwestern Polytechnical University, China

<sup>2</sup>School of Artificial Intelligence and Big Data, Henan University of Technology, China

<sup>3</sup>School of Computer Science, Northwestern Polytechnical University, China  
{jihongru,fanyuyin}@mail.nwpu.edu.cn, zm@haut.edu.cn,  
{li\_xianghua,wlw,cgao}@nwpu.edu.cn

## Abstract

Empathetic dialogue requires not only recognizing a user’s emotional state but also making strategy-aware, context-sensitive decisions throughout response generation. However, the lack of a comprehensive empathy strategy framework, explicit task-aligned multi-stage reasoning, and high-quality strategy-aware data fundamentally limits existing approaches, preventing them from effectively modeling empathetic dialogue as a complex, multi-stage cognitive and decision-making process. To address these challenges, we propose STRIDE-ED, a STRategy-grounded, Interpretable, and DEep reasoning framework that models Empathetic Dialogue through structured, strategy-conditioned reasoning. To support effective learning, we develop a strategy-aware data refinement pipeline integrating LLM-based annotation, multi-model consistency-weighted evaluation, and dynamic sampling to construct high-quality training data aligned with empathetic strategies. Furthermore, we adopt a two-stage training paradigm that combines supervised fine-tuning with multi-objective reinforcement learning to better align model behaviors with target emotions, empathetic strategies, and response formats. Extensive experiments demonstrate that STRIDE-ED generalizes across diverse open-source LLMs and consistently outperforms existing methods on both automatic metrics and human evaluations.

## 1 Introduction

Empathetic dialogue, a cornerstone of human social interaction, requires not only the recognition of another’s emotional state but also the formulation of a response that conveys understanding, validation, and appropriate support (Batson, 2009). From a psychological perspective, this is a complex, multi-stage cognitive and decision-making process (Davis, 1983; Gao et al., 2023). This highlights the strategic and context-sensitive nature of

\*Corresponding Author.



Figure 1: Illustration of the STRIDE-ED reasoning process. Given a user utterance, the framework performs scenario summarization, emotion recognition, strategy inference, and strategy-guided response generation.

empathy, making empathetic dialogue a challenging task beyond surface-level text generation.

Early studies focused on implicitly enhancing dialogue models through external commonsense knowledge or affective lexicons. For example, prior works (Liu et al., 2022; Cai et al., 2023) incorporate emotional commonsense graphs or commonsense knowledge selection into neural architectures to provide latent cues for more emotion-aware generation. However, without explicit modeling of reasoning or decision processes (Zhang et al., 2026c,b; Chen et al., 2025a; Luo et al., 2026; Wang et al., 2026), these approaches offer limited insight into the mechanisms through which emotional understanding informs response generation.

To address this lack of transparency, recent research has turned to Large Language Models (LLMs), which enable more explicit modeling of intermediate reasoning processes through Chain-of-Thought (CoT) prompting. Building on this di-

rection, prior work such as (Hu et al., 2025) further encourages LLMs to generate explicit intermediate rationales before producing responses, thereby enhancing interpretability (Zhang et al., 2024; Chen et al., 2025b). However, a fundamental limitation persists in these CoT-based approaches: the reasoning process lacks grounding in a rigorous strategic framework. Existing strategies (Liu et al., 2021; Zhang et al., 2025) often derived from specific domains such as psychological support, primarily address negative emotions and are confined to low-level responses. Emotions in dialogue typically span positive, negative, and neutral states (Ji et al., 2025), yet existing methods fail to fully capture this spectrum and lack support for higher-order cognitive strategies, as shown in Figure 1. Consequently, while these existing work structured in form, the model’s reasoning steps frequently remain superficial and exhibit inconsistent strategic coherence.

In summary, existing approaches suffer from three key limitations. (1) **Incomplete Empathy Strategy Coverage.** Prior methods lack a comprehensive strategy system covering diverse emotional states and higher-order cognition, limiting principled decision-making. (2) **Lack of Task-Aligned Multi-Stage Reasoning.** CoT-based methods do not explicitly model dialogue as a task-specific multi-stage reasoning process. (3) **Insufficient Strategy-Aware Supervision.** Training data lack sufficient high-quality annotations aligned with empathetic strategies and reasoning.

To address these issues, we propose STRIDE-ED, a framework that proceeds by first establishing a comprehensive strategy system, which then enables task-aligned reasoning through a dedicated data pipeline. At its core, we construct a unified strategy system covering positive, neutral, and negative emotions to guide response generation. Leveraging this system, we automatically annotate the EMPATHETICDIALOGUES dataset (Rashkin et al., 2019) with strategy types and rationales using authoritative LLMs. Subsequently, a rigorous data refinement process employs multi-LLM evaluation with consistency-weighted scoring and strategy-aware sampling to curate high-quality training subsets. Finally, the model is optimized through a two-stage training paradigm that combines supervised fine-tuning with multi-objective reinforcement learning, ensuring alignment with strategic, emotional, and structural correctness. Our contributions can be summarized as follows:

- We propose an interpretable framework for empathetic dialogue, STRIDE-ED, featuring a comprehensive empathy strategy system and stepwise decision-making.
- To support effective training, a strategy-aware data refinement pipeline is constructed, combining LLM-based annotation, multi-model weighted evaluation, and dynamic sampling to regulate strategy distribution and difficulty.
- A two-stage training paradigm is introduced, in which supervised fine-tuning establishes reasoning, followed by reinforcement learning to improve emotional alignment, strategy execution, and response consistency.
- Extensive experiments show that our framework generalizes across diverse open-source LLMs and achieves superior performance on both automatic and human evaluations.

## 2 Related Works

### 2.1 Implicit Knowledge-Driven Empathy

Early empathetic dialogue models lacked sufficient prior knowledge, limiting their understanding of users’ emotions and contexts. To address this, some studies incorporate external knowledge to broaden the model’s perspective (Zhang et al., 2026a) and enhance reasoning. For example, Ghosal et al. (2020) leveraged structured commonsense knowledge from ATOMIC (Sap et al., 2019), such as mental states and causal relations, to model interlocutor interactions for improved emotion understanding. Zhong et al. (2021) integrated commonsense-aware emotional latent concepts to generate emotionally appropriate responses, while Sabour et al. (2022) inferred users’ situational contexts through commonsense-based reasoning. Further, Cai et al. (2023) introduced an adaptive commonsense knowledge selection mechanism to refine contextual cognition, and Qiao et al. (2025) constructed multi-hop reasoning graphs to incorporate external knowledge during response generation. However, these methods did not explicitly model strategy-guided decision-making.

### 2.2 Explicit Reasoning with CoT for Empathy

LLMs possessed extensive knowledge reserves, and CoT prompting (Wei et al., 2022) enabled stepwise reasoning, facilitating more explicit decision-making in empathetic dialogue. Building on this

paradigm, Tu et al. (2022) inferred fine-grained user emotions and employed a mixed strategy mechanism for response generation, while Chen and Liu (2023) dynamically generated counseling strategies via zero-shot prompting to guide personalized responses. Subsequent work further enhanced empathy through data and structural designs: Chen et al. (2023) fine-tuned LLMs on consultant-style multi-turn dialogues, and Ye et al. (2025) adopted a strategy-enhanced role-playing framework with multiple interacting roles to generate diverse training data. Finally, Zhang et al. (2025) introduced an intention-centered framework that mapped inferred supporter intentions to support strategies using a chain-of-thought mechanism. Despite improved interpretability, these CoT-based approaches lacked comprehensive strategy coverage and did not explicitly model the full, structured reasoning process underlying human empathetic decision-making.

### 3 Methodology

STRIDE-ED is a general-purpose framework for empathetic dialogue applicable to diverse open-source LLMs. It implements a comprehensive empathy strategy system and a task-aligned, multi-step CoT paradigm to model dialogue as a progressive cognitive and decision-making process. At the data and training levels, STRIDE-ED integrates LLM-based automatic annotation, strategy-aware data refinement, and two-stage optimization. An overview is shown in Figure 2.

#### 3.1 Task Formulation

Empathetic dialogue consists of a multi-turn interaction between a user and a conversational agent. We denote the dialogue history as  $\mathcal{C} = \{u_1, u_2, \dots, u_{t-1}\}$ , where  $u_i$  represents the utterance at the  $i$ -th turn. Each utterance  $u_i$  is a sequence of tokens, i.e.,  $u_i = (w_{i,1}, w_{i,2}, \dots, w_{i,n_i})$ . The goal of empathetic dialogue is to generate an empathetic response  $u_t$  while appropriately recognizing the user’s emotional state  $e$ . In this work, we introduce auxiliary objectives and model empathetic dialogue as a sequential generation process. Specifically, conditioned on  $\mathcal{C}$ , the model generates a dialogue scenario summary  $sum$ , infers the target emotional state  $e$ , determines an empathy strategy  $stra$  along with its execution actions  $acts$ , and finally produces the empathetic response  $u_t$ , corresponding to modeling the conditional distribu-

tion  $P(sum, e, stra, acts, u_t | \mathcal{C})$ .

#### 3.2 Empathy Strategy System

In empathetic dialogue modeling, response strategies form a crucial intermediate stage between emotion understanding and response generation. Rather than generating surface-level replies directly, effective models must explicitly decide how to respond based on the user’s emotional state and dialogue context. Liu et al. (2021) proposed a widely adopted taxonomy of eight empathy strategies, primarily applied in counseling-oriented settings that focus on negative emotions. Higher-order cognitive strategies, however, remain underexplored, limiting the effectiveness of current systems in complex dialogues requiring sophisticated reasoning.

Motivated by these observations, we first analyze the emotional distributions in the EMPATHETIC-DIALOGUES dataset and conduct a fine-grained examination of response content. Based on these analyses, we expand the original empathy strategy set to better accommodate positive, neutral, and negative emotional contexts, and assign a three-level difficulty rating (I–III) to each strategy to reflect the cognitive complexity involved in its application. In particular, the original Question strategy is further subdivided into three distinct *Exploring*-type strategies to capture more nuanced cognitive and emotional interactions. The structured strategy system guides both annotation and model training (full details are provided in Appendix A).

#### 3.3 Stepwise Cognitive CoT Design

Inspired by insights from cognitive psychology, we model the human thought process during empathetic dialogue as a stepwise, incremental deliberation. Upon receiving a speaker’s utterance, individuals first infer missing information or make educated guesses about the described situation based on prior knowledge and contextual understanding, forming a high-level situational representation. This situational comprehension allows them to adopt the speaker’s perspective and facilitates accurate recognition of the speaker’s emotional state. Next, humans deliberate on which response strategy to employ, guided by the speaker’s emotional state and contextual cues. Because strategies represent generalized methods, their concrete execution may differ across scenarios. To account for this variability, we introduce an action inference stage that bridges strategy selection and the generation of the final response.

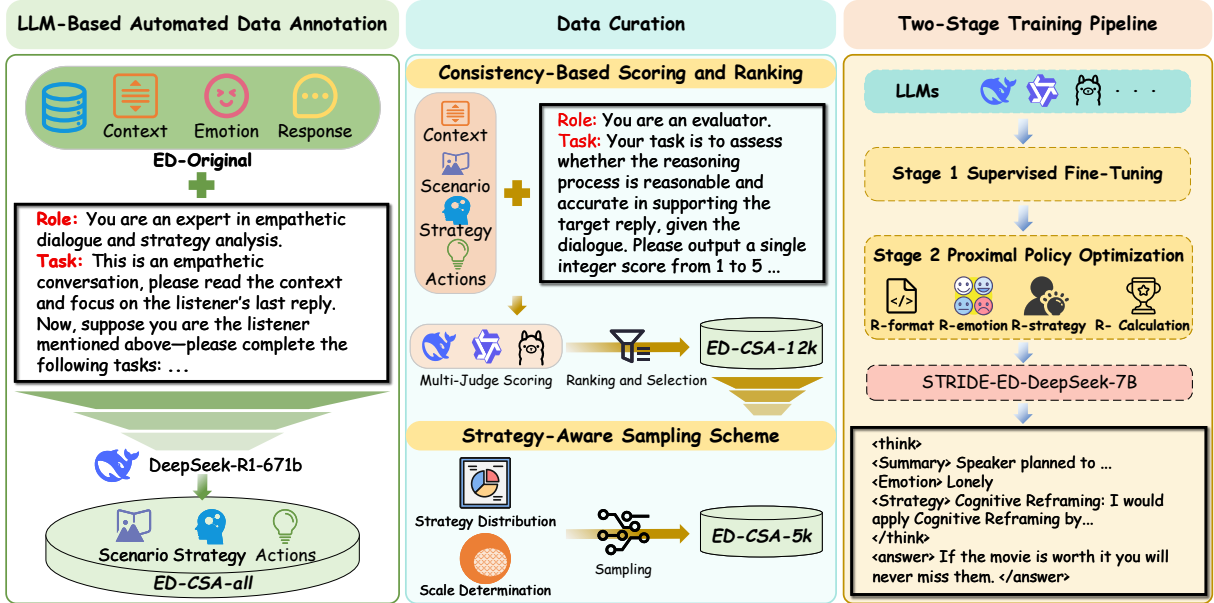


Figure 2: The architecture of the STRIDE-ED framework, illustrating the complete pipeline from data preparation and refinement to model training.

This stepwise CoT design captures the structured cognitive progression in human empathetic reasoning, providing a principled framework for multi-stage, strategy-aware response generation. Implementation leverages structured intermediate tags—such as `<Context>`, `<Emotion>`, and `<Strategy>`—to guide the model’s internal reasoning, culminating in the final response.

### 3.4 LLM-Based Automated Data Annotation

Building upon the aforementioned theoretical framework, we address the absence of explicit annotations required for model training. Specifically, the adopted dataset lacks annotations for dialogue scenario summaries, empathy strategy types, and the concrete actions used to implement each strategy. To this end, we employ DeepSeek-R1<sup>1</sup>, a large language model with strong reasoning capabilities, as an automated annotation expert. Carefully designed annotation prompts are used to generate structured labels for each dialogue instance, as shown in Appendix B. The annotated dataset resulting from this process is denoted as *ED-CSA-all*.

### 3.5 Consistency-Based Scoring and Ranking

To ensure the reliability of the automatically annotated data while mitigating the high cost of large-scale manual verification, we employ a multi-judge evaluation mechanism based on LLMs. Specifically, three representative models (DeepSeek-R1,

Qwen3<sup>2</sup>, and Llama-3.1<sup>3</sup>) are selected as independent expert evaluators. Using carefully designed evaluation prompts, as shown in Appendix B, each model assesses every example in the *ED-CSA-all* dataset and assigns a quality score ranging from 1 to 5, where higher scores indicate stronger semantic coherence and alignment among the annotated scenario summary, selected strategy, inferred action, and the corresponding dialogue context. Models are restricted to generate integer scores only.

We apply a reliability-weighted multi-judge aggregation framework to fairly integrate scores from multiple LLM evaluators. Let  $\mathcal{M} = \{m_1, m_2, m_3\}$  denote the set of evaluators, corresponding to DeepSeek-R1, Qwen, and LLaMA. For each  $m_i \in \mathcal{M}$ , let  $s_i$  denote its score vector over the dataset. The reliability of each evaluator is estimated as the average Spearman correlation (Spearman, 1904) with the others:

$$\rho_i = \frac{1}{2} \sum_{j \neq i} \text{Spearman}(s_i, s_j), \quad \forall i \in \mathcal{M}. \quad (1)$$

The final quality score for a sample  $x$  is computed as a reliability-weighted sum of evaluator scores with an agreement-based penalty:

$$S(x) = \sum_{i=1}^{|\mathcal{M}|} \frac{\rho_i}{\sum_{k=1}^{|\mathcal{M}|} \rho_k} s_i(x) - \lambda \sigma(x), \quad (2)$$

<sup>2</sup><https://huggingface.co/Qwen/Qwen3-8B>

<sup>3</sup><https://huggingface.co/meta-llama/Llama-3.1-8B-Instruct>

<sup>1</sup><https://huggingface.co/deepseek-ai/DeepSeek-R1>

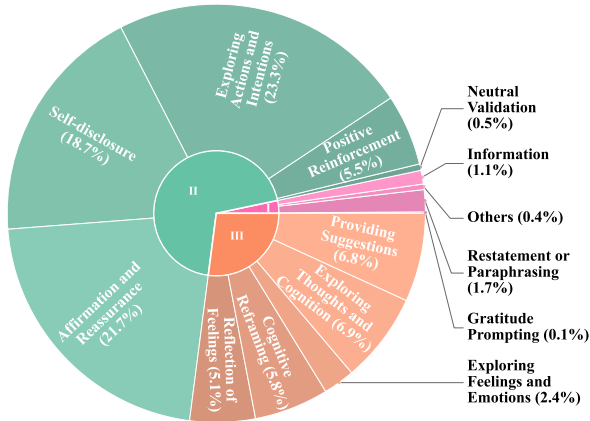


Figure 3: Distribution of Strategy Types in *ED-CSA-all*.

where  $s_i(x)$  denotes the score assigned to sample  $x$  by evaluator  $m_i$ , and  $\sigma(x)$  represents the standard deviation of the scores  $\{s_i(x)\}_{i=1}^{|\mathcal{M}|}$  across evaluators, measuring inter-rater disagreement. The hyperparameter  $\lambda$  is set to 0.1 in our experiments. This formulation favors samples with both high weighted scores and strong evaluator consensus.

Based on the resulting score ranking, we select the top 12k samples as the candidate pool for subsequent sampling, denoted as *ED-CSA-12k*.

### 3.6 Strategy-Aware Sampling Scheme

We further emphasize that strategy usage should reflect both distributional characteristics and contextual sensitivity rather than being applied uniformly. The distribution of strategy types in *ED-CSA-all* is illustrated in Figure 3. Data sorting and filtering inherently reshape the empirical strategy distribution, and directly training on such processed data may distort natural response patterns. Simultaneously, the model should learn selective strategy reasoning based on dialogue context, reserving explicit strategy and action deliberation for scenarios that demand higher cognitive effort.

Based on the above considerations, we design a strategy-aware sampling scheme in which the target sampling distribution is jointly determined by the empirical frequency of each strategy and its associated difficulty level. Let  $\mathcal{S} = \{s_1, \dots, s_{|\mathcal{S}|}\}$  denote the set of empathy strategies. We first compute the empirical frequency of each strategy  $s_i$  in the full annotated dataset *ED-CSA-all*, denoted as  $a_i$ . Next, we assign a difficulty weight  $d_i$  to each strategy according to its difficulty level, where higher-level strategies receive larger weights. We then compute a weighted score for each strategy by combining its empirical frequency with its assigned difficulty

level, and normalize these scores to obtain the target sampling proportions:

$$p_i = \frac{a_i \cdot d_i}{\sum_{j=1}^{|\mathcal{S}|} a_j \cdot d_j}. \quad (3)$$

This distribution favors the retention of infrequent yet cognitively demanding strategies, while preventing low-difficulty strategies from disproportionately dominating the sampled data. Subsequently, given the ranked candidate pool *ED-CSA-12k*, we apply a binary search procedure to determine the maximum subset size whose empirical strategy distribution can satisfy the target proportions  $p_i$ . Based on this subset size and the target distribution, we compute the required number of samples for each strategy and perform stratified random sampling to construct the refined training set *ED-CSA-5k*.

### 3.7 Model Training

After constructing the refined training data, inspired by prior training methods (Liu et al., 2026; Li et al., 2026a,b; Xue et al., 2026b,a), we adopt a two-stage optimization paradigm.

**Supervised Fine-Tuning.** In the first stage, we construct the supervised training set by combining the curated subset *ED-CSA-5k* with the remaining samples from *ED-CSA-all*, for which strategy annotations are removed. This design ensures that only the curated subset provides explicit strategy guidance, while the remaining samples contribute to general response generation without strategy supervision. The base LLM  $\mathcal{M}$  is then fine-tuned to sequentially generate both the intermediate reasoning components and the final response, yielding the model denoted as STRIDE-ED- $\mathcal{M}$ -SFT. The supervised training objective minimizes the negative log-likelihood over the structured output sequence:

$$\mathcal{L}_{\text{SFT}} = -\frac{1}{N} \sum_{n=1}^N \log p_{\theta}(y | \mathcal{C}), \quad (4)$$

where  $y$  represents the target output sequence conditioned on the dialogue history  $\mathcal{C}$ .

**Reinforcement Learning.** In the second stage, we further refine the SFT model via reinforcement learning on *ED-CSA-12k*. Beginning with STRIDE-ED- $\mathcal{M}$ -SFT, we apply Proximal Policy Optimization (PPO) to optimize the model towards outputs that better adhere to the desired format, emotional alignment, and strategy execution. The reward

function is defined as follows:

$$r_{\text{format}} = \begin{cases} 1, & \text{if format is correct} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

$$r_{\text{emotion}} = \begin{cases} 1, & \text{if emotion is correct} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

$$r_{\text{strategy}} = \begin{cases} 0, & \text{if strategy is incorrect} \\ 1, & \text{otherwise} \end{cases} \quad (7)$$

$$R = r_{\text{format}} \cdot (1 + r_{\text{emotion}} + r_{\text{strategy}}) \quad (8)$$

where  $r_{\text{format}}$ ,  $r_{\text{emotion}}$ , and  $r_{\text{strategy}}$  represent the reward components corresponding to format compliance, emotional accuracy, and strategic appropriateness, respectively. The format constraint is specified as “<think>.</think><answer>.</answer>”, and a regular expression is employed to verify whether the generated output adheres to this structural requirement. The overall reward  $R$  is designed to prioritize format correctness while jointly encouraging accurate emotional expression and effective strategic reasoning. The resulting model is denoted as STRIDE-ED- $\mathcal{M}$ .

## 4 Experiments

### 4.1 Datasets

We adopt the EMPATHETICDIALOGUES dataset (Rashkin et al., 2019) and the ESConv dataset (Liu et al., 2021) for our experiments. Developed by Facebook AI Research, EMPATHETICDIALOGUES is a large-scale multi-turn dialogue corpus dedicated to enhancing the empathetic capabilities of conversational systems. It comprises approximately 25,000 dialogue scenarios grounded in specific emotional contexts, covering 32 emotion labels. In the dataset, participants are assigned the roles of *speaker* and *listener* to simulate human interactions around emotional experiences. ESConv is a multi-turn dialogue dataset for emotional support conversations, which focuses on modeling support strategies.

### 4.2 Baselines

To ensure a comprehensive evaluation, we select baseline models spanning three major research trajectories in empathetic dialogue generation: (1) Traditional architectures that focus on neural network design for emotional modeling, including Transformer (Vaswani et al., 2017), MoEL (Lin et al., 2019), EmpDG (Li et al., 2020) and

GLHG (Peng et al., 2022). (2) External-knowledge-enhanced methods that incorporate commonsense or causal graphs to enrich responses, such as KEMP (Li et al., 2022), DCKS (Cai et al., 2023), E-CORE (Fu et al., 2023), Emp-USIR (Jiang et al., 2023) and EESC (Hao and Kong, 2025). (3) Reflective decision-integration approaches that explicitly model internal cognitive processes for strategic response generation, represented by CAB (Gao et al., 2023), IAMM (Yang et al., 2024), Emstremo (Li et al., 2024), Sibyl (Wang et al., 2025), and Reflect-Diffu (Yuan et al., 2025).

### 4.3 Implementation Details

We implement all LLM training using the verl<sup>4</sup> framework, with DeepSeek-7B-Chat selected as the primary model.<sup>5</sup> During the supervised fine-tuning stage, the initial learning rate is set to  $1 \times 10^{-4}$ , the batch size is 16, and the maximum input sequence length is capped at 2048 tokens. In the reinforcement learning stage, the total batch size is increased to 128, with the maximum generation length limited to 1024 tokens. We follow a fixed 8:1:1 split for training, validation, and testing. All experiments are conducted on four NVIDIA A800 GPUs, each equipped with 80 GB of memory. Our data and code are publicly available.<sup>6</sup>

### 4.4 Evaluation Metrics

To comprehensively evaluate STRIDE-ED, we conduct both automatic and human assessments.

**Automatic Evaluation.** We evaluate response quality using Perplexity (PPL), BLEU (B-n), Distinct (D-n), BARTScore, and emotion accuracy ( $\text{Acc}_{\text{emo}}$ ). PPL measures fluency, with lower values indicating greater coherence. BLEU captures relevance to reference responses via overlap, while Distinct reflects lexical diversity. BARTScore measures semantic similarity between generated and reference responses (Liu and Henao, 2025), providing a proxy for empathy.  $\text{Acc}_{\text{emo}}$  measures the consistency between predicted and true emotions.

**Human Evaluation.** We perform A/B testing to assess Empathy, Relevance, and Fluency. A panel of three annotators is employed, and 1000 dialogue turns generated by STRIDE-ED and baselines are selected for comparison, ensuring the consistency and interpretability of the evaluation outcomes.

<sup>4</sup><https://github.com/volcengine/verl>

<sup>5</sup><https://huggingface.co/deepseek-ai/deepseek-llm-7b-chat>

<sup>6</sup><https://github.com/cgao-comp/STRIDE-ED>

Models	B-1↑	B-2↑	B-3↑	B-4↑	Acc <sub>emo</sub> ↑	D-1↑	D-2↑	BARTScore↑	PPL↓
Transformer	18.07	8.34	4.57	2.86	–	0.36	1.35	-	37.62
MoEL	18.02	8.67	4.35	2.73	31.02	0.43	1.76	0.56	36.81
EmpDG	19.96	9.11	4.74	2.80	31.65	0.46	1.99	<u>0.57</u>	37.43
KEMP	18.07	8.30	4.37	2.65	36.40	0.55	2.29	0.52	36.89
DCKS	21.73	10.62	<u>6.24</u>	<u>4.09</u>	–	2.19	9.61	-	<u>16.08</u>
E-CORE	19.77	5.65	3.28	2.11	–	0.68	3.38	-	33.04
Emp-USIR	20.11	9.86	5.72	3.73	–	0.66	3.10	-	35.29
CAB	20.23	9.39	4.96	3.01	40.52	0.89	2.95	-	35.06
IAMM	19.51	8.74	4.86	3.32	43.72	0.88	3.05	-	25.94
Sibyl	21.45	9.35	5.01	2.95	-	<b>5.65</b>	<b>36.11</b>	-	-
ReflectDiffu	<u>23.59</u>	<u>11.25</u>	5.35	3.62	48.76	0.98	4.35	0.56	24.56
Ours	<b>24.54</b>	<b>11.96</b>	<b>7.13</b>	<b>4.67</b>	<b>57.25</b>	<u>2.56</u>	<u>13.63</u>	<b>0.71</b>	<b>10.50</b>

Table 1: Automatic evaluation results on the EMPATHETICDIALOGUES dataset. Boldface and underline indicate the best and second-best values, respectively. "Ours" refers to STRIDE-ED-DeepSeek-7B.

Models	B-1↑	D-1 ↑	D-2↑	PPL↓
Transformer	-	1.29	6.91	81.55
MoEL	19.04	2.33	15.26	62.93
GLHG	19.66	3.50	21.61	15.67
Emstremo	20.96	2.90	14.80	-
EESC	<u>21.38</u>	<u>4.88</u>	<u>25.95</u>	<u>14.88</u>
Ours	<b>22.03</b>	<b>5.11</b>	<b>29.63</b>	<b>10.09</b>

Table 2: Automatic evaluation results on the ESConv.

## 5 Results and Discussion

**Automatic Evaluation Results.** As shown in Table 1 and Table 2, we conduct five independent runs on each dataset and report the mean results. Our method consistently outperforms all baselines across automatic evaluation metrics. Compared with the strongest baseline, ReflectDiffu, our method achieves consistent improvements across all BLEU scores, with relative gains of 4.0% on BLEU-1, 6.3% on BLEU-2, 33.3% on BLEU-3, and 29.0% on BLEU-4. On ESConv, it also attains the highest BLEU-1, surpassing EESC. In terms of response quality, our method achieves the highest BARTScore, significantly outperforming all baselines, indicating better semantic alignment with references and enhanced empathetic quality. For emotional controllability, it improves emotion accuracy by 17.4% over ReflectDiffu in Acc<sub>emo</sub>. Our method also demonstrates stronger lexical diversity. On the main dataset, it achieves high Distinct-1 and Distinct-2 scores; on ESConv, it obtains the best Distinct-1 and Distinct-2, with relative improve-

ments of 4.7% and 14.2% over EESC, respectively. In addition, it achieves lower perplexity, reflecting improved fluency and generation confidence. Overall, these results show that our method not only surpasses existing empathetic dialogue models but also achieves a better balance among relevance, diversity, and fluency, validating the effectiveness of its structured reasoning and training framework.

Comparison	Aspects	Win	Lose	Tie
Ours vs. MOEL	Emp.	<b>46.1</b>	23.6	30.3
	Rel.	<b>39.8</b>	22.4	37.8
	Flu.	<b>33.5</b>	15.6	50.9
Ours vs. EmpDG	Emp.	<b>51.6</b>	18.5	29.9
	Rel.	<b>50.3</b>	15.3	34.4
	Flu.	<b>35.1</b>	13.5	51.4
Ours vs. CAB	Emp.	<b>53.2</b>	23.2	23.6
	Rel.	<b>56.6</b>	21.4	22.0
	Flu.	<b>31.5</b>	11.6	56.9

Table 3: Human A/B evaluation results on the EMPATHETICDIALOGUES dataset.

**Human Evaluation Results.** Table 3 presents human A/B testing results for STRIDE-ED and several baseline models on the EMPATHETICDIALOGUES dataset, evaluated in terms of empathy (Emp.), relevance (Rel.), and fluency (Flu.). The results highlight the performance characteristics of each model in generating empathetic, coherent, and contextually relevant responses.

**Ablation Study.** Table 4 reports the ablation results of STRIDE-ED, covering both the stepwise cognitive reasoning components, data curation strategies,

Models	B-1 $\uparrow$	Acc <sub>emo</sub> $\uparrow$	D-2 $\uparrow$	PPL $\downarrow$
Ours	24.66	57.57	13.68	9.26
w/o <i>e</i>	23.58	-	13.66	10.47
w/o <i>sum</i>	22.91	54.14	13.42	7.78
w/o <i>stra</i>	22.44	54.58	15.09	6.98
w/o CoT	22.55	-	13.26	8.64
w/o <i>R.&amp;S.</i>	22.22	55.93	15.25	11.86
w/o <i>S.</i>	23.44	56.05	14.40	9.89
w/o PPO	23.52	54.48	14.76	2.03

Table 4: Evaluation results of the ablation study on the EMPATHETICDIALOGUES dataset. *R.&S.* refer to the two modules: Consistency-Based Scoring and Ranking, and Strategy-Aware Sampling Scheme.

and PPO training. Overall, removing any component leads to performance degradation in one or more dimensions, confirming each module contributes non-trivially to the final model behavior.

**Stepwise Cognitive CoT Design.** Removing emotion reasoning (**w/o *e***) causes a clear drop in BLEU-1 and an increase in PPL, indicating that explicit emotion modeling not only improves emotional alignment but also stabilizes generation quality. Eliminating scenario summarization (**w/o *sum***) results in further degradation in BLEU-1 and emotion accuracy, suggesting that concise situation abstraction is crucial for grounding subsequent emotional and strategic reasoning. When strategy reasoning is removed (**w/o *stra***), emotion accuracy drops and fluency degrades, while diversity increases. This suggests that without strategy constraints, the model generates more varied but less controlled responses. Overall, strategy reasoning serves as a critical coordinating component in the framework, helping balance emotional controllability, coherence, and expressive diversity. Notably, removing structured reasoning (**w/o CoT**) leads to a drop in BLEU-1 while slightly reducing perplexity, suggesting that although responses become marginally easier to generate, they lose contextual relevance and cognitive grounding. This underscores the role of the stepwise cognitive CoT in enhancing response quality through structured reasoning rather than merely improving fluency.

**Data Scoring and Sampling.** Removing both consistency-based scoring and strategy-aware sampling (**w/o *R.&S.***) leads to the most severe degradation in BLEU-1, emotion accuracy, and perplexity, underscoring the role of systematic data filtering and distribution-aware sampling. Notably,

while Dist-2 improves slightly, this likely reflects increased superficial diversity at the cost of coherence and relevance, as indicated by declines in other metrics. When strategy-aware sampling is replaced with random sampling at the same scale (**w/o *S.***), performance still drops compared to the full model. This confirms a balanced, difficulty-aware strategy distribution is essential, and naive random sampling can distort strategic learning.

**PPO Training.** Removing the PPO stage results in a pronounced drop in both perplexity and emotion accuracy, reflecting a regression to generic and uncalibrated outputs. This highlights that PPO is essential for aligning the model along key dimensions, including proper formatting, strategy adherence, and emotional fidelity, extending beyond mere instruction following.

Models	B-1 $\uparrow$	Acc <sub>emo</sub> $\uparrow$	D-2 $\uparrow$	PPL $\downarrow$
Ours	24.66	57.57	13.68	9.26
1/2 Data	23.43	56.35	15.07	12.75
1/4 Data	23.37	55.12	14.56	8.80
1/8 Data	23.28	51.29	14.50	9.70
1/16 Data	3.32	45.53	3.38	6.21
Zero Data	14.86	-	21.49	1.09
Qwen3-0.6B	21.29	51.00	10.46	9.79
Qwen3-4B	20.49	56.35	12.51	10.20
Qwen3-4B-In.	22.76	55.93	14.32	7.87
LLama3.2-3B	22.77	57.06	14.49	11.11
GLM-Z1-9B	22.14	56.98	14.34	10.51

Table 5: Evaluation results for training set size and backbone model analyses. The abbreviation "Qwen3-4B-In." denotes Qwen3-4B-Instruct.

**Training Set Size Analysis.** Table 5 shows the results of training data size analysis on *ED-CSA-all*. Model competence remains stable until data is reduced to 1/8, beyond which performance collapses sharply at 1/16—indicating severe overfitting and loss of task capability. The zero-data setup shows that while the pretrained model retains strong generative fluency, it cannot adapt to the task without in-domain training. Overall, sufficient data is essential for effective adaptation, with a clear minimum required to maintain robust performance.

**Backbone Analysis.** Our method exhibits broad applicability and competitive performance across diverse LLM architectures and scales, as shown in Table 5. In the Qwen family, emotion accuracy improves with model size, whereas BLEU-1 does not scale monotonically—showing that compact

models can match or surpass larger ones in fluency with our approach. Across Qwen, LLama, and GLM families, the framework delivers consistently strong results: LLama3.2-3B leads in emotion accuracy and diversity, while Qwen3-4B-Instruct excels in BLEU-1 and fluency. Details of the backbone LLMs are provided in Appendix C. These findings confirm that our framework is architecture-agnostic and parameter-efficient, delivering robust performance across different LLMs and scales.

**Case Study** Representative case studies are presented to illustrate the reasoning process and interpretability of our framework, with details provided in Appendix D.

## 6 Conclusion

In this paper, we present STRIDE-ED, a structured empathetic dialogue framework that explicitly models stepwise cognitive reasoning over situation, emotion, and strategy. To support effective learning, we introduce a consistency-based scoring mechanism and a strategy-aware sampling scheme to construct supervision. Extensive experiments demonstrate that STRIDE-ED achieves improvements in empathetic controllability, response relevance, and lexical diversity across multiple settings. These findings suggest that incorporating explicit cognitive structure and data refinement is a promising direction for empathetic dialogue systems.

## Limitations

Our method currently lacks rigorous statistical validation in terms of data filtering and sampling. We plan to further investigate the underlying statistical principles and mechanisms to improve the soundness and rigor of the method design. In addition, our experiments indicate that training with DeepSeek-7B-Chat yields the best results; however, we have not yet examined whether this is related to the use of DeepSeek-R1 for data annotation. Furthermore, due to the high computational and time costs, we did not perform extensive hyperparameter tuning on other models. Future work will explore the performance limits of individual backbone models more thoroughly.

## Ethical considerations

LLMs are used for automatic annotation under prompts enforcing ethical constraints, and they are also applied to the grammatical checking and polishing of texts. Human evaluators are informed

that the evaluated content may contain negative emotions and are compensated fairly for their contributions. No personal or sensitive information is involved, and all experiments are conducted on publicly available datasets.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (Nos. 62271411, 62471403, 62572403), the Technological Innovation Team of Shaanxi Province (No. 2025RS-CXTD-009), the International Cooperation Project of Shaanxi Province (No. 2025GH-YBXM-017).

## References

- C Daniel Batson. 2009. *These things called empathy: Eight related but distinct phenomena*. *The Social Neuroscience of Empathy*, pages 3–16.
- Hua Cai, Xuli Shen, Qing Xu, Weilin Shen, Xiaomei Wang, Weifeng Ge, Xiaoqing Zheng, and Xiangyang Xue. 2023. Improving empathetic dialogue generation by dynamically infusing commonsense knowledge. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7858–7873.
- Jiawei Chen, Yang Yang, Chao Yu, Yu Tian, Zhi Cao, Xue Yang, Linghao Li, Hang Su, and Zhaoxia Yin. 2025a. Red teaming large reasoning models. *arXiv preprint arXiv:2512.00412*.
- Qi Chen and Dexi Liu. 2023. Dynamic strategy chain: Dynamic zero-shot cot for long mental health support generation. *arXiv preprint arXiv:2308.10444*.
- Yao Chen, Jiawei Sheng, Wenyuan Zhang, and Tingwen Liu. 2025b. Improving reasoning capabilities in small models through mixture-of-layers distillation with stepwise attention on key information. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 4952–4971.
- 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.
- Mark H Davis. 1983. Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1):113–126.
- Fengyi Fu, Lei Zhang, Quan Wang, and Zhendong Mao. 2023. E-core: Emotion correlation enhanced empathetic dialogue generation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10568–10586.

- Pan Gao, Donghong Han, Rui Zhou, Xuejiao Zhang, and Zikun Wang. 2023. [Cab: empathetic dialogue generation with cognition, affection and behavior](#). In *International Conference on Database Systems for Advanced Applications*, pages 597–606. Springer.
- Deepanway Ghosal, Navonil Majumder, Alexander Gelbukh, Rada Mihalcea, and Soujanya Poria. 2020. Cosmic: Commonsense knowledge for emotion identification in conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2470–2481.
- Jiawang Hao and Fang Kong. 2025. Enhancing emotional support conversations: A framework for dynamic knowledge filtering and persona extraction. In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 3193–3202.
- He Hu, Yucheng Zhou, Juzheng Si, Qianning Wang, Hengheng Zhang, Fuji Ren, Fei Ma, and Laizhong Cui. 2025. Beyond empathy: Integrating diagnostic and therapeutic reasoning with large language models for mental health counseling. *arXiv e-prints*, pages arXiv–2505.
- Hongru Ji, Xianghua Li, Mingxin Li, Meng Zhao, and Chao Gao. 2025. Hybrid relational graphs with sentiment-laden semantic alignment for multimodal emotion recognition in conversation. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence*, pages 2973–2981.
- Liting Jiang, Di Wu, Yanbing Li, and Wushouer Silamu. 2023. Emp-usir: A unidirectional synchronous interactive reasoning model for empathetic dialogue. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–9. IEEE.
- Guocong Li, Jinjian Zhang, Ping Wang, Dongnan Liu, Tian Liang, Qiuyi Qi, Hao Huang, Siyan Guo, Mutian Bao, Wei Zhou, Linjian Mo, Hongxia Xu, and Jian Wu. 2026a. Mol: Adaptive mixture-of-length reasoning for efficient question answering with context. In *The Fourteenth International Conference on Learning Representations*.
- Junlin Li, Bo Peng, and Yu-Yin Hsu. 2024. Emstremo: Adapting emotional support response with enhanced emotion-strategy integrated selection. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 5794–5805.
- Qintong Li, Hongshen Chen, Zhaochun Ren, Pengjie Ren, Zhaopeng Tu, and Zhumin Chen. 2020. Empdg: Multi-resolution interactive empathetic dialogue generation. In *Proceedings of the 28th international conference on computational linguistics*, pages 4454–4466.
- Qintong Li, Piji Li, Zhaochun Ren, Pengjie Ren, and Zhumin Chen. 2022. Knowledge bridging for empathetic dialogue generation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 10993–11001.
- Yu Li, Rui Miao, Zhengling Qi, and Tian Lan. 2026b. Arise: Agent reasoning with intrinsic skill evolution in hierarchical reinforcement learning. *arXiv preprint arXiv:2603.16060*.
- Zhaojiang Lin, Andrea Madotto, Jamin Shin, Peng Xu, and Pascale Fung. 2019. Moel: Mixture of empathetic listeners. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 121–132.
- Hongye Liu and Ricardo Henao. 2025. Learning to substitute words with model-based score ranking. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 11551–11565.
- 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.
- Tao Liu, Taiqiang Wu, Runming Yang, Shaoning Sun, Junjie Wang, and Yujiu Yang. 2026. Profit: Leveraging high-value signals in sft via probability-guided token selection. *arXiv preprint arXiv:2601.09195*.
- Ye Liu, Wolfgang Maier, Wolfgang Minker, and Stefan Ultes. 2022. Empathetic dialogue generation with pre-trained roberta-gpt2 and external knowledge. In *Conversational AI for Natural Human-Centric Interaction: 12th International Workshop on Spoken Dialogue System Technology, IWSDS 2021, Singapore*, pages 67–81. Springer.
- Guanran Luo, Wentao Qiu, Zhongquan Jian, Meihong Wang, and Qingqiang Wu. 2026. [Gcot-decoding: Unlocking deep reasoning paths for universal question answering](#). *Preprint*, arXiv:2604.06794.
- Wei Peng, Yue Hu, Luxi Xing, Yuqiang Xie, Yajing Sun, and Yunpeng Li. 2022. Control globally, understand locally: A global-to-local hierarchical graph network for emotional support conversation. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pages 4324–4330.
- Baiyou Qiao, Yuekai Zhang, Pan Gao, Xinchu Li, Shuo Wang, and Donghong Han. 2025. Multi-perspective empathy modeling for empathetic dialogue generation. *Knowledge-Based Systems*, 314:113191.
- Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. Towards empathetic open-domain conversation models: A new benchmark and dataset. In *Proceedings of the 57th annual meeting of the association for computational linguistics*, pages 5370–5381.

- Sahand Sabour, Chujie Zheng, and Minlie Huang. 2022. Cem: Commonsense-aware empathetic response generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 11229–11237.
- Maarten Sap, Ronan Le Bras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A Smith, and Yejin Choi. 2019. Atomic: An atlas of machine commonsense for if-then reasoning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 3027–3035.
- C Spearman. 1904. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.
- Quan Tu, Yanran Li, Jianwei Cui, Bin Wang, Ji-Rong Wen, and Rui Yan. 2022. Misc: A mixed strategy-aware model integrating comet for emotional support conversation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 308–319.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.
- Lanrui Wang, Jiangnan Li, Chenxu Yang, Zheng Lin, Hongyin Tang, Huan Liu, Yanan Cao, Jingang Wang, and Weiping Wang. 2025. [Sibyl: Empowering empathetic dialogue generation in large language models via sensible and visionary commonsense inference](#). In *Proceedings of the 31st International Conference on Computational Linguistics*, pages 123–140, Abu Dhabi, UAE. Association for Computational Linguistics.
- Yu Wang, Emmanuele Chersoni, and Chu-Ren Huang. 2026. This one or that one? a study on accessibility via demonstratives with multimodal large language models. In *Language Resources and Evaluation Conference 2026*. European Language Resources Association (ELRA).
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.
- Chao Xue, Yao Wang, Mengqiao Liu, Di Liang, Xingsheng Han, Peiyang Liu, Xianjie Wu, Chenyao Lu, Lei Jiang, Yu Lu, Haibo Shi, Shuang Liang, Minlong Peng, and Flora D. Salim. 2026a. [Reason only when needed: Efficient generative reward modeling via model-internal uncertainty](#). *Preprint*, arXiv:2604.10072.
- Chao Xue, Yao Wang, Mengqiao Liu, Di Liang, Xingsheng Han, Peiyang Liu, Xianjie Wu, Chenyao Lu, Lei Jiang, Yu Lu, Haibo Shi, Shuang Liang, Minlong Peng, and Flora D. Salim. 2026b. [Why supervised fine-tuning fails to learn: A systematic study of incomplete learning in large language models](#). *Preprint*, arXiv:2604.10079.
- Zhou Yang, Zhaochun Ren, Wang Yufeng, Haizhou Sun, Chao Chen, Xiaofei Zhu, and Xiangwen Liao. 2024. An iterative associative memory model for empathetic response generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3081–3092.
- 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.
- Jiahao Yuan, Zixiang Di, Zhiqing Cui, Guisong Yang, and Usman Naseem. 2025. Reflectdiffu: Reflect between emotion-intent contagion and mimicry for empathetic response generation via a rl-diffusion framework. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 25435–25449.
- Qianchi Zhang, Hainan Zhang, Liang Pang, Hongwei Zheng, and Zhiming Zheng. 2026a. Stable-rag: Mitigating retrieval-permutation-induced hallucinations in retrieval-augmented generation. *arXiv preprint arXiv:2601.02993*.
- 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.
- 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.
- Xinjie Zhang, Wenxuan Wang, and Qin Jin. 2025. [Intentionesc: An intention-centered framework for enhancing emotional support in dialogue systems](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 26494–26516, Vienna, Austria. Association for Computational Linguistics.
- 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.
- Peixiang Zhong, Di Wang, Pengfei Li, Chen Zhang, Hao Wang, and Chunyan Miao. 2021. Care: commonsense-aware emotional response generation with latent concepts. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 14577–14585.

## A Strategy System Details

The comprehensive strategy system underpinning the STRIDE-ED framework is detailed in Table 6. It encompasses 14 distinct strategies categorized into three tiers of cognitive and implementational Difficulty (I, II, and III). This taxonomy is constructed to guide the model’s reasoning chain across the full emotional spectrum, ranging from fundamental responses like Restatement (Difficulty I) to advanced cognitive interventions such as Cognitive Reframing (Difficulty III). The tiered structure ensures that the framework can adaptively engage in appropriate strategic planning, mirroring the nuanced decision-making process of human empathetic dialogue.

## B Prompts

In the research on conversational strategy analysis and reasoning quality evaluation of large language models, standardized prompts are the key to ensuring the consistency and reliability of task results. Below, we will introduce the design concepts and principles of two core prompts. For the Data Annotation Prompt, it achieves high-quality strategy classification annotation and clarifies the reasons for strategy selection through professional role guidance, closed-set strategy selection, and standardized output. The second one is the Consistency-Based Scoring Prompt, which performs minimal quantitative scoring from a neutral evaluation perspective and completes an objective assessment of the rationality of model reasoning by invoking three different models.

### Data Annotation Prompt

**Role:** You are an expert in empathetic dialogue and strategy analysis.

**Inputs:**

1. Context: {}
2. Emotion: {}
3. Listener’s Response: {}

**Task:** This is an empathetic conversation, please read the context and focus on the listener’s last reply. Now, suppose you are the listener mentioned above—please complete the following tasks:

<Summary> Briefly summarize the speaker’s situation.

### Data Annotation Prompt (Continued)

<Strategy> From the first-person perspective, choose one strategy from the optional strategies and their interpretations below and explain how you would apply it, keeping the reasoning concise.

Optional strategies:

1.Exploring Thoughts and Cognition: Probes the speaker’s beliefs, interpretations, and thought processes.

2.Exploring Actions and Intentions: ...

...

14.Others: ...

**Output Format:**

<Summary> ...

<Strategy> [one strategy], [reason and actions]

**Output Requirement:**

1. Focus on the speaker’s last utterance for the need.
2. Pick only strategies actually used in the listener’s response.
3. Be concise and precise.

### Consistency-Based Scoring Prompt

**Role:** You are an evaluator.

**Inputs:**

- 1.The dialogue context. {}
- 2.The target reply. {}
- 3.The generated reasoning process. {}

**Task:** Your task is to assess whether the reasoning process is reasonable and accurate in supporting the target reply, given the dialogue. Please output a single integer score from 1 to 5 (1 = very poor, 5 = excellent).

**Output Requirement:** Output only the integer score (1-5), without any explanation, extra text, punctuation, or formatting.

## C Experimental Details

This appendix section provides an overview of the backbone LLMs used in our experiments, and information regarding the involved existing packages is available in the code repository <sup>7</sup>.

**Qwen3-0.6B**<sup>8</sup> is a 0.6B-parameter causal language model with 28 layers and a 32,768-token context. It supports thinking and non-thinking modes,

<sup>7</sup><https://anonymous.4open.science/r/STRIDE-ED/>

<sup>8</sup><https://huggingface.co/Qwen/Qwen3-0.6B>

	<b>Strategy</b>	<b>Definition</b>	<b>Difficulty</b>
1	Gratitude Prompting	Encourages the speaker to notice and reflect on positive experiences or supportive aspects of their life, fostering positive emotional awareness.	I
2	Restatement or Paraphrasing	Rephrases the speaker's main ideas to demonstrate understanding and attentive listening.	I
3	Others	Covers responses that do not clearly fit into any predefined strategy category.	I
4	Information	Provides objective facts or relevant knowledge to help the speaker better understand their situation or make decisions.	I
5	Neutral Validation	Affirms that neutral or low-intensity emotional states are normal and acceptable without encouraging stronger emotions.	II
6	Positive Reinforcement	Highlights the speaker's strengths, efforts, or constructive behaviors to reinforce confidence and motivation.	II
7	Exploring Actions and Intentions	Uses targeted questions to clarify the speaker's actions, plans, and underlying intentions.	II
8	Self-disclosure	Shares limited, relevant personal information to foster rapport and mutual understanding.	II
9	Affirmation and Reassurance	Acknowledges the speaker's feelings and offers comfort or emotional support.	II
10	Reflection of Feelings	Identifies and articulates emotions that the speaker implies but does not explicitly express.	III
11	Cognitive Reframing	Offers an alternative perspective on a difficult situation while respecting the speaker's original emotions.	III
12	Exploring Feelings and Emotions	Uses open-ended prompts to encourage deeper expression of the speaker's emotional experience.	III
13	Exploring Thoughts and Cognition	Probes the speaker's beliefs, interpretations, and thought processes.	III
14	Providing Suggestions	Offers practical and actionable recommendations tailored to the speaker's needs.	III

Table 6: The Structured Empathy Strategy System of STRIDE-ED.

demonstrating strong reasoning, agent capabilities, and multilingual instruction following.

**Qwen3-4B<sup>9</sup>** is a 4B-parameter causal language model with 36 layers and a 32,768 natively and 131,072 tokens with YaRN context. It features enhanced multilingual understanding, high-efficiency inference, and robust performance on Chinese and English tasks, excelling in factual accuracy, creative generation, and industrial scenario adaptation.

**Qwen3-4B-Instruct<sup>10</sup>** is a 4B-parameter causal language model with 36 layers. Optimized for instruction following, it supports zero-shot/few-shot learning, real-time dialogue interaction, and domain-specific task customization, demonstrating superior alignment with human intent and practical application scalability.

**LLama3.2-3B<sup>11</sup>** is a 3B-parameter causal language model with 28 layers and a 8,192-token context. It focuses on lightweight deployment, delivering balanced performance in reasoning, code generation, and multilingual processing, with optimized efficiency for edge and low-resource environments.

**GLM-Z1-9B<sup>12</sup>** is a 9B-parameter causal language model with 40 layers and a 30,000-token context. It integrates self-supervised pre-training with instruction fine-tuning, showcasing strong mathematical reasoning, long-text comprehension, and cross-modal fusion capabilities, tailored for complex industrial and academic reasoning tasks.

## D Case Study

In this appendix, we present a set of representative case studies covering higher-difficulty strategies, including Providing Suggestions, Exploring Thoughts and Cognition, Cognitive Reframing, etc., to illustrate the performance of our model. These examples are shown in Tables 7 - 12.

Each case provides multi-dimensional core information: dialogue context, high-level empathetic response strategy, model-generated response, and real-scenario reference result. Side-by-side presentation of this information intuitively illustrates the model’s strategy selection logic and linguistic expression ability in complex conversations, and allows direct comparison between model outputs and real results to accurately evaluate the model’s performance in applying high-level empathetic response strategies.

<sup>9</sup><https://huggingface.co/Qwen/Qwen3-4B>

<sup>10</sup><https://huggingface.co/Qwen/Qwen3-4B-Instruct-2507>

<sup>11</sup><https://huggingface.co/meta-llama/Llama-3.2-3B>

<sup>12</sup><https://huggingface.co/zai-org/GLM-Z1-9B-0414>

<b>Context</b>	<i>Speaker:</i> I am going to a concert very soon.
<b>Summary</b>	The speaker is anticipating an upcoming concert they will attend.
<b>Emotion Strategy</b>	Excited Exploring Actions and Intentions
<b>Actions</b>	I chose Exploring Actions and Intentions because the speaker’s excitement about the concert hints at unfolding plans, and asking about details channels that energy into shared anticipation.
<b>Generated Response</b>	That’s awesome! Who are you going to see?
<b>Ground Truth</b>	Cool! Who are you going to see?

Table 7: Case Study 1 illustrating the use of the Cognitive Reframing strategy

<b>Context</b>	<i>Speaker:</i> Speaking of kids in school, I just found out my son got a B in his world history class at the end of the previous school year. I can’t believe it. <i>Listener:</i> That’s great!! History is so important I think and so few kids really grasp it well. Congrats to him for a job well done! <i>Speaker:</i> No no, see, it was the first grade below an A that he’s ever made. I honestly thought I had the next Einstein on my hands. Now I guess I’ll have to settle for...above average, I guess. :(
<b>Summary</b>	The speaker’s son achieved a B in history, his first grade below an A, which initially made the speaker hopeful but now feels disappointed.
<b>Emotion Strategy</b>	surprised Cognitive Reframing
<b>Actions</b>	I chose Cognitive Reframing because the speaker’s disappointment over a single B clashed with sky-high expectations, and recasting the grade as proof of challenging coursework softens the sting and restores perspective.
<b>Generated Response</b>	Well, that’s still a great grade! I’m sure he’ll do better next year. I’m sure you’re proud of him!
<b>Ground Truth</b>	Oh, I see what you are saying! Did he get upset about it? And you never know, he could turn it around next year and blow you away!

Table 8: Case Study 2 illustrating the use of the Cognitive Reframing strategy

<b>Context</b>	<i>Speaker:</i> Every year when the kids go back to school, it brings back so many memories of when my kids were going to school. <i>Listener:</i> I know what you mean. I love to see all the pictures, and think back to how little mine used to be. <i>Speaker:</i> I would cry the first day of school every year, especially the milestone years. Now my grandchildren are starting school. Time flies by too quickly
<b>Summary</b>	The speaker reflects on their children's school years, now experiencing the same bittersweet emotions as their grandchildren begin school, expressing how quickly time passes.
<b>Emotion Strategy</b>	Sentimental Self-disclosure
<b>Actions</b>	I chose Self-disclosure because the speaker's nostalgia over grandchildren starting school echoes my own disbelief at how fast kids grow, and sharing that sense of fleeting time affirms we're in the same bittersweet boat.
<b>Generated Response</b>	I know. I have a 10 year old and a 13 year old, and I can't believe how fast it has gone by. I'm sure you feel the same way.
<b>Ground Truth</b>	They do seem to be grown before you know it. Make you wonder how it goes so fast!

Table 9: Case Study 3 illustrating the use of the Self-disclosure strategy

<b>Context</b>	<i>Speaker:</i> I am going to my local amusement park tomorrow and feeling a certain way about it!
<b>Summary</b>	The speaker is anticipating a visit to an amusement park and feels a mix of excitement and anxiety about the experience.
<b>Emotion Strategy</b>	Apprehensive Exploring Thoughts and Cognition
<b>Actions</b>	I chose Exploring Thoughts and Cognition because the speaker's mixed excitement-and-apprehension about the amusement park suggests hidden worries, and probing the thoughts behind the nerves can turn anxiety into eager curiosity.
<b>Generated Response</b>	That sounds like a lot of fun! Are you nervous?
<b>Ground Truth</b>	Which way are you feeling? Are you a thrill seeker?

Table 10: Case Study 4 illustrating the use of the Exploring Thoughts and Cognition strategy

<b>Context</b>	<i>Speaker:</i> I'm so happy summer classes are officially over and I'll be able to relax for a week! <i>Listener:</i> Great what will you be doing during your time off ? <i>Speaker:</i> I'm not even sure what to do with all this time! I guess I should relax or have some leisure time, but everything in me says to try and be productive even on break. It's that school mentality.
<b>Summary</b>	The speaker is conflicted about balancing relaxation and productivity during their week-long break, struggling with the lingering mindset of being productive despite wanting downtime.
<b>Emotion Strategy</b>	Joyful Providing Suggestions
<b>Actions</b>	I chose Providing Suggestions because the speaker felt torn between rest and productivity, and offering concrete ways to enjoy downtime eases the guilt while keeping the momentum alive.
<b>Generated Response</b>	I think you should just relax and enjoy the time off. You can always get back to work after that.
<b>Ground Truth</b>	I agree with you, you'll feel good about yourself being productive !

Table 11: Case Study 5 illustrating the use of the Providing Suggestions strategy.

<b>Context</b>	<i>Speaker:</i> The litter of puppies had been born! <i>Listener:</i> Nice, amazing to witness birth. <i>Speaker:</i> All of the puppies survived, truly a lucky amazing occasion!
<b>Summary</b>	The speaker is celebrating the successful birth of their puppies and expressing gratitude for their survival.
<b>Emotion Strategy</b>	joyful Affirmation and Reassurance
<b>Actions</b>	I chose Affirmation and Reassurance because the speaker is deeply joyful about their puppies' birth and expressing thanks for their survival; this reinforces their positive feelings and validates their relief by acknowledging the success and wishing them well.
<b>Generated Response</b>	That is a great news, congrats.
<b>Ground Truth</b>	Good to hear, have fun with puppies!

Table 12: Case Study 6 illustrating the use of the Affirmation and Reassurance strategy