CATCH: A Novel Data Synthesis Framework for High Therapy Fidelity and Memory-Driven Planning Chain of Thought in AI Counseling

Mingyu Chen[1*], **Jingkai Lin**[1*], **Zhaojie Chu**[1],

Xiaofen Xing[1†], **Yirong Chen**[1], **Xiangmin Xu**[2,1†]

[1]School of EE., South China University of Technology, Guangzhou, China
[2]Foshan University, Foshan, China
{eemingyuchen, eelinjingkai, 202210182056, eeyirongchen}@mail.scut.edu.cn,
{xfxing, xmxu}@scut.edu.cn

Abstract

Recently, advancements in AI counseling based on large language models have shown significant progress. However, existing studies employ a one-time generation approach to synthesize multi-turn dialogue samples, resulting in low therapy fidelity and failing to capture the decision-making rationale behind each response. In this work, we propose CATCH, a novel data synthesis framework designed to address these challenges. Specifically, to improve therapy fidelity, we introduce the Progressive Dialogue Synthesis strategy, which extracts goals, resources, and solutions from a client's self-report, organizes them into structured outlines, and then incrementally generates stage-aligned counseling dialogues. To capture decision-making rationale behind each response, we propose the Memory-Driven Dynamic Planning thinking pattern that integrates memory enhancement, global planning, and strategy reasoning; a collaborative multi-agent optimizer then leverages MDP to attach explicit chain-of-thought to each dialogue turn. Extensive experiments and human evaluations demonstrate that CATCH significantly enhances fidelity and logical coherence in AI counseling.[1]

1 Introduction

Mental health issues are becoming increasingly prominent worldwide (Santomauro et al., 2021), highlighting counseling as a crucial intervention method. However, due to the scarcity and high costs associated with professional counselors, a significant gap persists between service availability and demand. Recent advancements in Large Language Models (LLMs) (OpenAI, 2024; Guo et al., 2025; Yang et al., 2025) offer promising opportunities to bridge this gap, providing novel

tools for mental health counseling (Hua et al., 2024; Xiao et al., 2024; Xie et al., 2025; Lee et al., 2024; Xu et al., 2025). To effectively leverage LLMs in mental health counseling, large-scale, high-quality datasets of multi-turn counseling dialogues are essential.

Although recent methods have enabled the synthesis of large-scale multi-turn counseling dialogue datasets, they typically adopt a **one-time generation approach**(Qiu et al., 2024; Zhang et al., 2024; Lee et al., 2024; Xie et al., 2025), where a Large Language Model (LLM) produces the entire dialogue in a single inference step. This approach offers limited control over the therapeutic process, making it difficult to ensure that the dialogue adheres to the structured stages of a given therapy, resulting in low **therapy fidelity**(Waltz et al., 1993; Bellg et al., 2004; Perepletchikova and Kazdin, 2005; Sanetti et al., 2021)—the degree to which an intervention aligns with established therapeutic guidelines. Furthermore, these datasets generally lack explicit chain-of-thought (CoT) (Wei et al., 2022) that expose the decision-making rationale behind each counselor response. Consequently, models trained on such data tend to imitate surface-level conversational patterns rather than grasping the underlying therapeutic logic. These challenges cause AI counseling models to suffer from **therapy drift** (Waller, 2009; Waller and Turner, 2016), where the dialogue progressively deviates from the intended therapeutic trajectory.

In this work, we introduce CATCH, a **C**-ollaborative multi-**A**-gent-based framework for **T**-urn-level **C**-hain-of-Thought and **H**-igh-fidelity counseling dialogues synthesis, thereby mitigating therapy drift. To enhance therapy fidelity, CATCH adopts a Progressive Dialogue Synthesis (PDS) strategy, which incrementally generates dialogues stage by stage, aligning with structured therapeutic intervention protocols. Starting from client self-reports, PDS identifies key

---

elements—counseling goals, resources, and solutions—and organizes them into outlines. These outlines, combined with the client's personality traits, guide LLMs to synthesize the dialogue for each therapeutic stage. To explicitly capture the decision-making rationale behind each counselor response, we introduce a Memory-driven Dynamic Planning (MDP) thinking pattern that incorporates memory enhancement, global planning, and strategy reasoning. Built upon MDP, we develop a collaborative multi-agent approach to generate MDP CoT. To evaluate CATCH's effectiveness, we developed 8B and 14B scale models using dialogue datasets enriched with MDP CoT. Extensive experiments and human evaluations demonstrate that CATCH improves the fidelity and logical coherence of AI counseling, effectively mitigating therapy drift.

The contributions of this work are as follows:

- We propose the Progressive Dialogue Synthesis (PDS) strategy. Unlike one-time generation approach, PDS incrementally synthesizes dialogues stage by stage, aligning with structured therapeutic protocols and deriving key elements from client self-reports to guide LLM-based content generation for each therapeutic stage.

- We introduce the Memory-driven Dynamic Planning (MDP) thinking pattern—comprising memory enhancement, global planning, and strategy reasoning—and devise a collaborative multi-agent method to produce explicit turn-level chain-of-thought, making the counselor's decision-making rationale transparent.

- Extensive experiments and human evaluations demonstrate that CATCH enhances the fidelity and logical coherence of AI counseling while effectively mitigating therapy drift.

2 Related Work

2.1 Psychological Models

With the growing application of LLMs, the mental health domain has seen a surge in LLMs-based approaches tailored for counseling scenarios. Early efforts primarily focused on data collection and construction. For example, SoulChat (Chen et al., 2023) built a large-scale corpus with over 2 million dialogue samples, while MindChat (Xin Yan, 2023)

curated approximately 200,000 multi-turn counseling dialogues through manual cleaning, providing a valuable foundation for downstream modeling.

To improve the modeling of specific intervention strategies, some studies have explored the integration of structured therapy procedures(Xu et al., 2025; Xiao et al., 2024). AutoCBT (Xu et al., 2025) developed a single-turn multi-agent system for cognitive behavioral therapy (CBT). Healme (Xiao et al., 2024) decomposed the cognitive restructuring process of CBT into three actionable dialogue stages, using prompts to guide the LLMs step-by-step in assisting the client. However, these systems typically handle only a limited number of turns and are insufficient to capture the complexity of real-world multi-turn counseling interactions.

Due to the scarcity of authentic multi-turn counseling data, some studies have investigated the synthesis of such dialogues. SMILE (Qiu et al., 2024) expands single-turn counseling interactions to create multi-turn dialogues. CpsyCoun (Zhang et al., 2024) reconstructs multi-turn dialogues based on counseling reports. Other works target therapy-specific dialogue synthesis: Cactus (Lee et al., 2024) leverages CBT plans to guide the generation of multi-turn dialogues, and PsyDT (Xie et al., 2025) constructs a twin dialogue dataset of counselors by simulating the therapeutic language style of counselors. Although these studies synthesize large-scale datasets of multi-turn counseling dialogues, they use a one-time generation approach to directly synthesize multi-turn dialogue samples. These synthetic datasets pose challenges for AI counselors in effectively applying the fundamental principles of therapy at each stage, resulting in low therapy fidelity for the synthesized dialogues.

2.2 Structured Psychotherapy

Structured psychotherapy refers to therapeutic approaches guided by clear theoretical frameworks, standardized procedures, and predefined objectives. These therapies are characterized by goal-oriented practices, clear procedures, and a focus on short-term effectiveness, which makes them easier to evaluate and replicate (Chambless and Hollon, 1998; Lambert, 2013). Prominent examples include:
Cognitive Behavioral Therapy (CBT) (Beck, 2020), the most widely applied structured therapy, aims to identify and restructure negative thought patterns. It usually involves 8–16 sessions with clearly defined steps and goals.
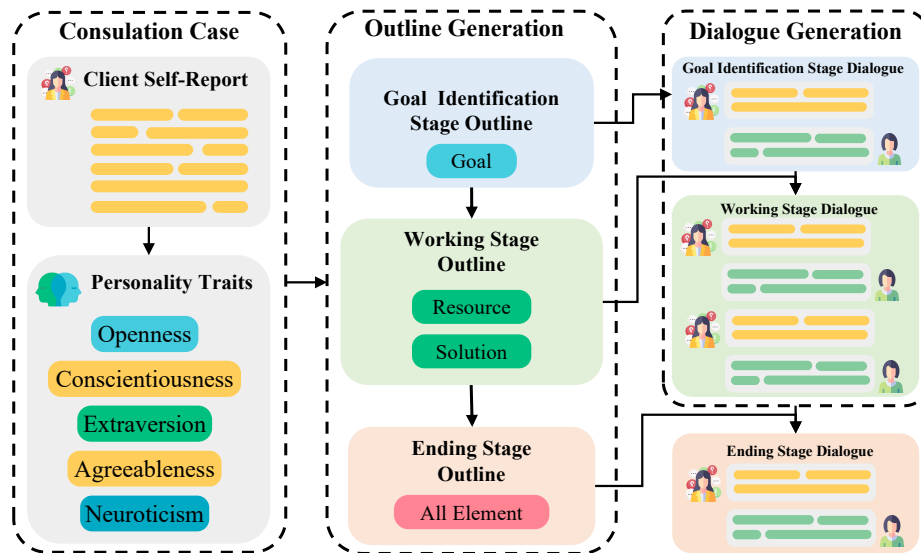
Figure 1: The Progressive Dialogue Synthesis (PDS) strategy for high-fidelity dialogue generation. PDS translates therapeutic principles into a structured process: from client self-reports and personality traits, it derives goals, resources, and solutions, forming a stage-guided outline that ensures fidelity in the incremental generation of counseling dialogues.

Dialectical Behavior Therapy (DBT) (Robins and Rosenthal, 2011) builds upon CBT by adding mindfulness and emotion regulation techniques. It employs modular training and explicit therapeutic targets, making it particularly suitable for disorders characterized by emotional instability.

Solution-Focused Brief Therapy (SFBT) (Żak and Pękala, 2024) emphasizes constructing solutions instead of analyzing problems, typically completed within 4–8 sessions. It focuses on activating the client's internal resources.

Single-Session Therapy (SST) (Kim et al., 2023) considers each session as an independent intervention addressing the client's most pressing issue. It follows a structured procedure: setting the counseling goal, activating resources, and planning actions.

3 Methodology

In this work, we present a comprehensive introduction to the proposed data synthesis framework, CATCH, designed for structured psychotherapy. This work uses SST as a specific therapy. In Section 3.1, we describe the **Progressive Dialogue Synthesis** (PDS) to generate a counseling dialogue dataset with high therapy fidelity. Section 3.2 introduces the generation of the **Memory-driven Dynamic Planning** (MDP) Chain-of-Thought(CoT). Section 3.3 focuses on developing a LLMs-based AI counselor utilizing the synthesized dataset enriched with the MDP CoT.

3.1 PDS Strategy

Existing work synthesizing datasets of counseling dialogues is challenged by low therapy fidelity (Qiu et al., 2024; Zhang et al., 2024; Xie et al., 2025; Lee et al., 2024). We propose the PDS strategy to generate a counseling dialogue dataset aligned with structured therapeutic protocols. Detailed prompts used in PDS are provided in Appendix C.1

3.1.1 Consultation Case Acquisition

To ensure the authenticity and effectiveness of the synthesized counseling dialogues, we collected posts describing mental health from Yixinli [2]. Each post contains the client's self-report text, which serves as the initial information for subsequent dialogue creation, as shown in Figure 1 (left). Following PysDT (Xie et al., 2025), we extracted the client's Big Five personality traits from the self-report text as supporting information.

3.1.2 Outline Generation

Unlike the existing dialogue generation approach, we use the PDS strategy to enhance the therapy fidelity. As shown in Figure 1 (middle), our proposed PDS strategy requires the sequential performance of a goal identification stage, a working stage, and an ending stage. The elements of counseling dialogue creation are obtained. Then, based on these elements, we utilize LLM to generate an outline of
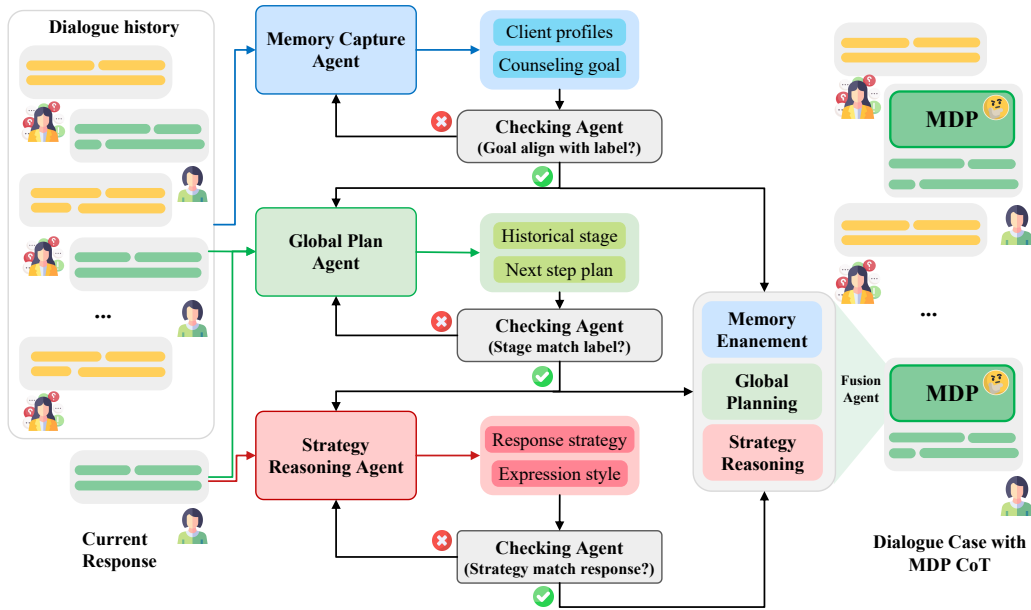
---

[2]https://www.xinli001.com/

Figure 2: The Memory-Driven Dynamic Planning (MDP) Chain-of-Thought (CoT) synthesis pipeline. The framework employs a collaborative multi-agent system to generate structured counseling reasoning: the Memory Capture Agent summarizes and extracts key information from the dialogue history, the Global Plan Agent determines therapeutic stage and progression, and the Strategy Reasoning Agent refines response strategies. Each step is verified by a Checking Agent for consistency, and the Fusion Agent integrates validated components into a coherent, first-person thought process.

the dialogue for each therapy stage. Terminology explanations are provided in Appendix A.

Goal Identification Stage. *Element Acquisition:* Based on the client's self-report and the Big Five personality traits, we use an LLM to obtain a counseling goal $\mathcal{G}$ that meets the assessment criteria for SST therapy and is unique, specific, feasible, positive, evaluable, and focused on the client's internal experiences and responses rather than on changing external circumstances or the actions of others. (De Shazer et al., 2021). *Outline Generation:* The counseling dialogue begins with the counselor establishing an initial trusting relationship with the client, progressively explores the client's current distress, and ultimately clarifies and confirms the counseling goal $\mathcal{G}$ at the ending stage.

Working Stage. *Element Acquisition:* Combining the client's self-report and the counseling goal $\mathcal{G}$, we employ the LLMs to explore multiple resources $\mathcal{R}$ related to the client's demands, covering both negative resource $\mathcal{R}_N$ and positive resource $\mathcal{R}_P$. Based on the explored resources $\mathcal{R}$, the $\mathcal{R}_P$ is used to generate solutions $\mathcal{S}$. *Outline Generation:* The counselor conducts a resource exploration around the counseling goal $\mathcal{G}$. After several counseling explorations, the $\mathcal{R}_N$ and $\mathcal{R}_P$ are obtained. The $\mathcal{R}_P$ is transformed into solutions $\mathcal{S}$.

Ending Stage. *Element Integration:* All of the core elements from the first two stages are integrated, including consulting goals $\mathcal{G}$, negative resources $\mathcal{R}_N$, positive resources $\mathcal{R}_P$, and solutions $\mathcal{S}$. *Outline Generation:* Under the guidance of the counselor, the client accepts and is willing to try to apply the solutions $\mathcal{S}$. Moreover, the counselor summarizes the counseling process and assists the client in reinforcing the positive changes made.

3.1.3 Dialogue Generation

As shown in Figure 1(right), based on the dialogue outline (described in Section 3.1.2), we utilized the client's personality traits and the SST knowledge to inform both the client's and counselor's language expression. The LLMs generate the consultation dialogue at each stage and attach the stage status label to each turn of responses. The dialogues for each subsequent stage were built upon the content generated in the previous stage, ultimately stitching together all stages to create a comprehensive sample of a multi-round counseling dialogue.

3.1.4 Data Filtering

To ensure the quality of the dialogue samples, we filtered out instances with abnormal turns and disordered statements. DeepSeek-R1(Guo et al., 2025) is employed to assess the therapeutic structure and

goal focus of the dialogues, retaining only the data that passed this evaluation through a majority voting process. Additionally, the LLMs produce statements containing excessive technical jargon due to the provision of SST knowledge. Thus, we implement keyword matching to identify statements with an overabundance of specialized terms, again using DeepSeek-R1 for assessment. If a statement contained too many technical terms, we would revise it while preserving its original meaning.

3.2 MDP Thinking Pattern

3.2.1 MDP Thinking Pattern Design

Case conceptualization (Kuyken et al., 2011) is the process through which the counselor systematically understands, integrates, and interprets the client's issues, providing guidance for subsequent counseling sessions. In this process, the counselor considers various factors, such as the client's personal information, background, and psychological state, to construct a comprehensive and in-depth case model. Based on this theory, we designed the MDP expert thinking pattern by simulating and modeling the case conceptualization process, aiming to help the model better understand the inherent therapeutic logic in multi-turn counseling dialogues. The thinking pattern contains Memory Enhancement, Global Planning, and Strategy Reasoning.

Memory Enhancement. Revisits and synthesizes the dialogue history to construct a structured summary of the client's state, including their background, presenting issues, and emotional states. This process ensures a continually updated understanding of the client.

Global Planning. Assesses progress toward therapeutic goals. Plans subsequent steps to ensure that the conversation maintains a coherent therapeutic trajectory, aligning short-term interactions with long-term objectives.

Strategy Reasoning. Utilizes the outputs from memory enhancement and global planning to determine the optimal response strategy for the current turn. Select specific therapeutic techniques and conversational tactics to address the client's immediate needs while advancing the global plan.

3.2.2 MDP Chain-of-Thought Synthesis

To explicitly model the decision-making rationale for each turn, we designed a collaborative multiagent framework to synthesize MDP CoT, as illustrated in Figure 2. This framework sequentially performs memory enhancement, global planning, and strategy reasoning using three specialized generative agents: the Memory Capture Agent, the Global Planning Agent, and the Strategy Reasoning Agent. To ensure accuracy and consistency, the output of each generative step is verified and refined by a corresponding Checking Agent. Finally, a Fusion Agent integrates the validated outputs from the three generative agents into a cohesive and comprehensive thought process. Detailed prompts of different agents are provided in the Appendix C.2

Memory Capture Agent. The task of the Memory Extract Agent is to extract and summarize key memory information based on historical dialogues before the current turn, including (1) an overall description of the client profiles and (2) an assessment of the counseling goal.

Global Plan Agent. The Stage Plan Agent combines the memory information examined with the SST knowledge. The agent perform the following two tasks: (1) Identify and confirm the counseling stage that the current dialogue is progressing through; (2) Determine the stage planning required for the current response, which includes three planning states: maintaining the status quo within the current stage, transitioning between steps within the current stage, and facilitating transitions to the next stage. Ultimately, the agent generates a comprehensive global plan.

Strategy Reasoning Agent. Combining the content provided by the Memory Capture Agent and the Global Plan Agent, along with the theoretical foundation of SST therapy, the specific response strategy is further refined. The agent selects appropriate expression strategies, such as tone and length, for the current response.

Checking Agent. This agent verifies the output of each generative step to ensure consistency and therapeutic alignment. Specifically, it checks: (1) *Memory Enhancement* for accurate client information and goal status against stage labels; (2) *Global Planning* for logical consistency with historical and current stage labels; and (3) *Strategy Reasoning* for alignment with the global plan and the actual counselor response. If any contradiction is detected, it triggers a regeneration mechanism to ensure that the content aligns with the expected therapeutic logic and goals. Specifically, during the regeneration process, the agent incrementally increases the temperature parameter and incorporates a verification label as part of the prompt—effectively guiding the model toward acceptable outputs. This iterative refinement continues until the generated

response passes the validation check or the maximum number of attempts is reached.

Fusion Agent. The Fusion Agent is responsible for integrating the content generated by the first three agents and verifying it with the Checking Agent. The agent then transforms this content into a natural and fluent dialogue in the first person, aligning it more closely with the counselor's actual thought process. An illustrative example of the generated MDP CoT is provided in the Appendix Figure 8.

3.3 Conversation LLMs

Based on the synthesized dataset and the MDP-CoT, the multi-turn AI counseling dialogue model CATCH-LLMs with the therapy thought pattern is trained. To enhance the model's comprehension of counseling concepts, we place the SST knowledge as a system prompt at the beginning of each dialogue sequence.

In this work, we formulate the training objective of CATCH-LLMs as a standard next-token prediction task (Xie et al., 2025). Given an input sequence composed of the SST knowledge and the preceding dialogue history, the model predicts a target response that concatenates the MDP-CoT reasoning steps and the counselor's response.

Formally, the supervised fine-tuning (SFT) loss function is defined as follows:

$$\mathcal{L}_{\text{SFT}} = -\sum_{t=1}^{T} \log P_\theta(y_t \mid y_{<t}, x), \quad (1)$$

where:

$$x = concat(K_{\text{SST}}, h), \quad (2)$$
$$y = concat(C_{\text{MDP}}, r). \quad (3)$$

Here, P_θ represents the probability distribution parameterized by the model parameters θ, T denotes the length of the target sequence, K_{SST} represents the SST knowledge, h denotes the dialogue history, C_{MDP} represents the MDP CoT, and r is the counselor's response.

4 Experiments

4.1 Implementations

We employ GPT-4o [3] to implement the PDS strategy for the dialogue dataset, and utilize Doubao-1.5-pro-256k [4] to synthesize MDP CoT, result-

[3] https://chat.openai.com
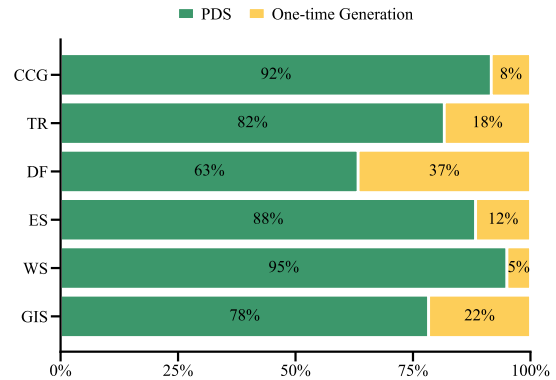[4] https://team.doubao.com

Figure 3: Comparison of expert evaluations between dialogues generated by the PDS strategy and the one-time generation approach. Detailed definitions of each metric are provided in Table 6.

ing in 233 multi-turn counseling dialogues comprising 6,898 response entries with MDP CoT. Based on the LLaMA-Factory (Zheng et al., 2024) framework, we fine-tuned the Qwen3 (Team, 2025) model at 8B and 14B scales. More training details see in Appendix B.

	r	ρ	τ
Correlation	0.66	0.74	0.59
P-value	0.0016	0.0002	0.0009

Table 1: Correlation coefficients and corresponding p-values between GPT-4o and expert ratings. Here r, ρ, and τ denote Pearson's, Spearman's, and Kendall's Tau correlation coefficients, respectively.

4.2 Evaluation Scale

SST Therapy Rating Scale. To evaluate the model's performance in applying SST principles within counseling dialogues, we adopt the Solution-Focused Inventory (SFI) (Grant et al., 2012) as our assessment framework. SST emphasizes the efficient facilitation of clients in a single session to identify and focus on key issues, activate internal and external resources, and formulate clear and achievable goals. These objectives align closely with the three core dimensions of the SFI. Based on this alignment, we define the following SST-specific evaluation metrics derived from SFI: **Problem Focus (PF)**: Measures the counselor's ability to guide the client from problem narration toward solution-building, hope, and positive experiences. **Resource Activation (RA)**: Assesses the counselor's effectiveness in activating the client's

Model	SST-specific Skills			General Counseling Skills		
	Solution Focus ↑	Resource Activation ↑	Goal Orientation ↑	Understanding ↑	Interpersonal ↑	Collaboration ↑
Qwen3-8B	4.47	3.98	2.85	4.47	**5.43**	4.48
Qwen3-14B	4.62	4.08	3.02	4.30	5.19	4.34
DeepSeek-V3	4.83	4.78	4.04	4.53	5.04	4.94
DeepSeek-R1	**4.96**	4.79	4.07	4.69	5.31	4.91
GPT-4o	4.88	4.65	3.96	**4.80**	5.39	4.96
CATCH-8B	4.91	5.29	4.30	4.52	5.10	4.95
CATCH-14B	**4.96**	**5.36**	**4.38**	4.62	5.17	**4.99**

Table 2: Results of evaluation on SST-specific skills and general counseling skills. Scores are averaged over 450 dialogues. The best results are highlighted in **bold**.

internal and external resources and resilience, and in translating them into actionable strategies. **Goal Orientatio (GO)**: Evaluates the counselor's ability to help the client set clear goals, track progress, and develop feasible action plans. Detailed prompt in Appendix C.4.

General Counseling Skill Metrics. Same as COUNSELINGEVAL (Lee et al., 2024), we selected three metrics from the CTRS (Aarons et al., 2012) to evaluate general counseling skills: Understanding, Interpersonal Effectiveness, and Collaboration.

Expert vs GPT-4o. To validate the reliability of LLMs in SFI assessment, we conducted a comparative analysis between expert ratings and GPT-4o's evaluations across three metrics in SST practice. We randomly selected 60 dialogues generated by different models interacting with simulated clients. Experts were recruited to independently rate the transcripts using the same standardized assessment rubric that was provided to GPT-4o in its evaluation prompt. This ensured methodological parity between human and model-based judgments. As shown in Table 1, GPT-4o demonstrated statistically significant positive correlations with human experts on all three metrics: Pearson correlation coefficient ($r = 0.66$, $p = 0.0016$), Spearman's rank correlation coefficient ($\rho = 0.74$, $p = 0.0002$), and Kendall's tau-b coefficient ($\tau = 0.59$, $p = 0.0009$). These results align with prior findings from Cactus (Lee et al., 2024), confirming that advanced LLMs like GPT-4o can effectively replicate human-level judgment in assessing key therapeutic competencies required for single-session counseling.

4.3 Evaluation of PDS Strategy

To evaluate the effectiveness of the PDS strategy, we employed the same model to generate multi-turn counseling dialogues for 60 randomly sampled counseling cases under the guidance of the

SST knowledge, using both the PDS approach and the one-time generation approach. Expert evaluators were then invited to assess and compare the dialogues generated by the two methods for each case. The evaluation was conducted across multiple dimensions, including **therapeutic implementation effectiveness** (covering the goal-identification stage (GIS), working stage (WS), ending stage (ES)), **the coherence of the counselor's guidance** (CCG), **dialogue fluency** (DF), and **topic relevance** (TR). For each case, experts were presented with two anonymized dialogues—one generated using the one-time generation approach and the other using the PDS method. Evaluators were instructed to conduct a side-by-side comparison and select the dialogue that demonstrated superior performance on each predefined dimension. Detailed definitions and evaluation criteria for each metric are provided in Appendix A.

Evaluation Results. As shown in Figure 3, dialogues generated by the PDS strategy demonstrated significantly superior therapeutic implementation effectiveness compared to the one-time generation methods, indicating a higher degree of alignment with therapeutic goals and laying a solid foundation for subsequent MDP synthesis. Moreover, the PDS strategy also obtains outstanding performance in terms of dialogue fluency and topic relevance, suggesting that the PDS strategy contributes to improved logical coherence and semantic consistency, thereby better reflecting the characteristics of real-world counseling interactions.

4.4 Evaluation for AI Counselor

Experiment Setup. We adopt a dialogue simulation evaluation protocol consistent with COUNSELINGEVAL (Lee et al., 2024), utilizing its accompanying counseling case dataset, which includes 150 client information sheets and three predefined client attitude types (Positive, Neutral, Neg-

Model	Method	SST-specific Skills			General Counseling Skills		
		Solution Focus ↑	Resource Activation ↑	Goal Orientation ↑	Understanding ↑	Interpersonal ↑	Collaboration ↑
Qwen3-8B	w/o PDS and MDP	4.47	3.98	2.85	4.47	**5.43**	4.48
	w/o MDP	4.77	5.09	4.16	4.41	4.83	4.93
	CATCH-8B	**4.91**	**5.29**	**4.30**	**4.52**	5.10	**4.95**
Qwen3-14B	w/o PDS and MDP	4.62	4.08	3.02	4.30	**5.19**	4.34
	w/o MDP	4.81	5.14	4.14	4.45	4.91	4.92
	CATCH-14B	**4.96**	**5.36**	**4.38**	**4.62**	5.17	**4.99**

Table 3: Results of the ablation study on SST-specific skills and general counseling skills for Qwen3-8B and Qwen3-14B under various settings. Scores are averaged over 450 dialogues. The best results are highlighted in **bold**.
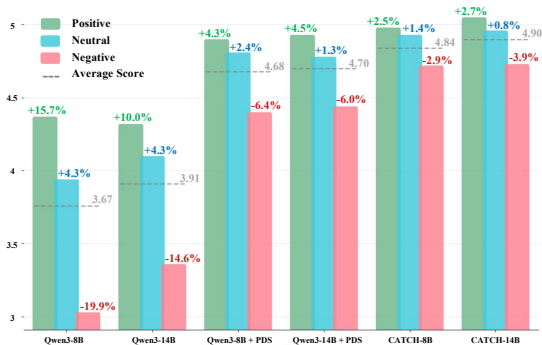

Figure 4: Performance comparison across different client attitudes. The bar chart shows the average evaluation scores of various models under positive, neutral, and negative client attitude settings. The dotted line indicates the overall average score across all attitudes. Notably, baseline models exhibit a pronounced performance gap, particularly struggling with negative-attitude clients. This highlights a critical weakness in real-world counseling scenarios.

ative), to assess the application of various models within the SST. Given that the client role primarily involves understanding and responding—tasks that are relatively simple in terms of language generation—we employ GPT-4o-mini as the AI client due to its fast response time and strong language comprehension capabilities, ensuring natural and coherent dialogue generation. We compare five LLMs: GPT-4o, DeepSeek-V3 (DeepSeek-AI, 2024) and DeepSeek-R1 (Guo et al., 2025), as well as our backbone models Qwen3-8B & 14B. To ensure a fair comparison, all method prompts contain the same SST knowledge. Detailed prompts in Appendix C.3

Main Results. As shown in Table 2, our model outperforms all baseline models. In particular, our model achieves improvements of **7.36%, 31.37%** and **45.03%** over its base models for the SF, RA, and GO metrics, respectively. These results demonstrate that our dataset effectively enhances the

model's ability to apply therapeutic strategies. The improvements in the RA and GO metrics are significantly greater than those in the SF metric. The results suggest that while the base model—when guided by SST knowledge—is already capable of maintaining focus on the client's problems, it struggles with identifying client resources and sustaining clear goal orientation throughout the dialogue. In terms of general-purpose metrics, our model also maintains strong competitiveness compared to baseline models.

4.5 Ablation Study

To systematically evaluate the contributions of the Progressive Dialogue Synthesis (PDS) strategy and the Memory-driven Dynamic Planning Chain-of-Thought (MDP CoT), we conduct ablation experiments on both Qwen3-8B and Qwen3-14B models. We compare the full CATCH framework against variants trained with partial or no access to our proposed components, using the same evaluation protocol as described in Section 4.4. The results are summarized in Table 3.

Impact of MDP CoT. When comparing models with and without MDP CoT, we observe consistent performance drops across all three SST-specific skill as well as the three general counseling skill metrics. This indicates that MDP plays a critical role in guiding the model to better understand the counselor's intended therapeutic trajectory at each turn, thereby enhancing not only its capacity to apply SST principles but also its general counseling competency.

Contribution of PDS. Notably, even without MDP, the models trained with PDS still outperform their backbone models across all SST dimensions. This demonstrates that PDS alone significantly enhances the model's alignment with therapeutic goals by promoting turn-level coherence and therapy-specific fidelity in synthetic dialogues.

Robustness Across Client Attitudes. We further analyze model performance under different client attitudes—positive, neutral, and negative—to assess robustness in handling challenging interactions. As shown in Figure 4, baseline models (w/o PDS and MDP) exhibit a pronounced performance gap, particularly struggling with negative-attitude clients. This highlights their vulnerability to therapy drift when confronted with complex or non-compliant client dynamics.

5 Conclusion

In this work, we introduced a novel data synthesis framework, CATCH, designed to enhance the quality of multi-turn counseling dialogues in AI counseling. By employing the PDS strategy, we generated a counseling dialogue dataset that aligns with therapeutic intervention principles, thereby improving therapy fidelity. Additionally, the MDP CoT clarified decision-making motivations for each dialogue turn, facilitating a deeper understanding of therapeutic logic. Our extensive experiments demonstrate that CATCH not only enhances the fidelity and logical coherence of AI counseling but also effectively mitigates therapy drift. This framework represents a significant advancement in the field of AI counseling, providing a robust foundation for future research and applications.

Limitations

While our results affirm the CATCH framework's effectiveness in enhancing counseling models, several limitations warrant consideration. One key consideration is the scale of our training corpus. Our collection of 233 multi-turn dialogues, though meticulously aligned with SST principles, may not be extensive enough to ensure robust generalization across a wide spectrum of psychosocial contexts and linguistic styles.

Moreover, the dataset is confined to single-session interactions, thereby neglecting the longitudinal dynamics inherent in authentic therapeutic relationships. The current architecture lacks a mechanism for cross-session memory integration, which is critical for tracking patient progress, adjusting interventions, and managing the evolution of therapeutic goals. Future research will therefore prioritize both the expansion of our dataset and the development of hierarchical memory architectures to enable coherent, multi-session therapeutic tracking.

References

Gregory A Aarons, Elizabeth A Miller, Amy E Green, Jennifer A Perrott, and Richard Bradway. 2012. Adaptation happens: a qualitative case study of implementation of the incredible years evidence-based parent training programme in a residential substance abuse treatment programme. *Journal of Children's Services*, 7(4):233–245.

Judith S Beck. 2020. *Cognitive behavior therapy: Basics and beyond*. Guilford Publications.

Albert J Bellg, Belinda Borrelli, Barbara Resnick, Jacki Hecht, Daryl Sharp Minicucci, Marcia Ory, Gbenga Ogedegbe, Denise Orwig, Denise Ernst, and Susan Czajkowski. 2004. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the nih behavior change consortium. *Health psychology*, 23(5):443.

Dianne L Chambless and Steven D Hollon. 1998. Defining empirically supported therapies. *Journal of consulting and clinical psychology*, 66(1):7.

Yirong Chen, Xiaofen Xing, Jingkai Lin, Huimin Zheng, Zhenyu Wang, Qi Liu, and Xiangmin Xu. 2023. SoulChat: Improving LLMs' empathy, listening, and comfort abilities through fine-tuning with multi-turn empathy conversations. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1170–1183, Singapore. Association for Computational Linguistics.

Steve De Shazer, Yvonne Dolan, Harry Korman, Terry Trepper, Eric McCollum, and Insoo Kim Berg. 2021. *More than miracles: The state of the art of solution-focused brief therapy*. Routledge.

DeepSeek-AI. 2024. Deepseek-v3 technical report. *Preprint*, arXiv:2412.19437.

Anthony M Grant, Michael J Cavanagh, Sabina Kleitman, Gordon Spence, Michaela Lakota, and Nickolas Yu. 2012. Development and validation of the solution-focused inventory. *The Journal of positive psychology*, 7(4):334–348.

Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, and 1 others. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *Preprint*, arXiv:2106.09685.

Yining Hua, Fenglin Liu, Kailai Yang, Zehan Li, Hongbin Na, Yi-han Sheu, Peilin Zhou, Lauren V Moran, Sophia Ananiadou, Andrew Beam, and 1 others. 2024. Large language models in mental health care: a scoping review. *arXiv preprint arXiv:2401.02984*.

Jongtae Kim, Namgil Ryu, and Dixon Chibanda. 2023. Effectiveness of single-session therapy for adult common mental disorders: a systematic review. *BMC psychology*, 11(1):373.

Willem Kuyken, Christine A Padesky, and Robert Dudley. 2011. *Collaborative case conceptualization: Working effectively with clients in cognitive-behavioral therapy*. Guilford Press.

Michael J Lambert. 2013. *Bergin and Garfield's handbook of psychotherapy and behavior change*. John Wiley & Sons.

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.

Ilya Loshchilov and Frank Hutter. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

OpenAI. 2024. Gpt-4 technical report. *Preprint*, arXiv:2303.08774.

Francheska Perepletchikova and Alan E Kazdin. 2005. Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical psychology: Science and practice*, 12(4):365.

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.

Clive J Robins and M Zachary Rosenthal. 2011. Dialectical behavior therapy. *Acceptance and mindfulness in cognitive behavior therapy: Understanding and applying the new therapies*, pages 164–192.

Lisa MH Sanetti, Bryan G Cook, and Lysandra Cook. 2021. Treatment fidelity: What it is and why it matters. *Learning Disabilities Research & Practice*, 36(1):5–11.

Damian F Santomauro, Ana M Mantilla Herrera, Jamileh Shadid, Peng Zheng, Charlie Ashbaugh, David M Pigott, Cristiana Abbafati, Christopher Adolph, Joanne O Amlag, Aleksandr Y Aravkin, and 1 others. 2021. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. *The Lancet*, 398(10312):1700–1712.

Qwen Team. 2025. Qwen3. https://qwenlm.github.io/blog/qwen3.

Glenn Waller. 2009. Evidence-based treatment and therapist drift. *Behaviour research and therapy*, 47(2):119–127.

Glenn Waller and Hannah Turner. 2016. Therapist drift redux: Why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. *Behaviour research and therapy*, 77:129–137.

Jennifer Waltz, Michael E Addis, Kelly Koerner, and Neil S Jacobson. 1993. Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *Journal of consulting and clinical psychology*, 61(4):620.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, and 1 others. 2022. Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35:24824–24837.

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.

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.

Dong Xue* Xin Yan. 2023. Mindchat: Psychological large language model. https://github.com/X-D-Lab/MindChat.

Ancheng Xu, Di Yang, Renhao Li, Jingwei Zhu, Minghuan Tan, Min Yang, Wanxin Qiu, Mingchen Ma, Haihong Wu, Bingyu Li, and 1 others. 2025. Autocbt: An autonomous multi-agent framework for cognitive behavioral therapy in psychological counseling. *arXiv preprint arXiv:2501.09426*.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao,

Chengen Huang, Chenxu Lv, Chujie Zheng, Day-
iheng 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.

Andreea M Żak and Krzysztof Pękala. 2024. Effective-
ness of solution-focused brief therapy: An umbrella
review of systematic reviews and meta-analyses. *Psy-
chotherapy Research*, pages 1–13.

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 psychologi-
cal counseling. In *Findings of the Association for
Computational Linguistics: ACL 2024*, pages 13947–
13966, Bangkok, Thailand. Association for Compu-
tational 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 As-
sociation for Computational Linguistics (Volume 3:
System Demonstrations)*, pages 400–410, Bangkok,
Thailand. Association for Computational Linguistics.

A Terminology and Assessment Metrics

Terminology in Single-Session Therapy (SST).
We provide detailed explanations of key terminol-
ogy and definitions of evaluation metrics to ensure
a consistent understanding of core concepts and as-
sessment criteria. Table 4 and 5 presents clear defi-
nitions of essential psychotherapeutic terms used
in SST. These terms constitute the foundational
framework of SST and inform decision-making
throughout the data synthesis process.

Metrics and Criteria in Evaluation of PDS. Table
6 elaborates on the six core dimensions employed
in PDS evaluation: Goal Identification Stage Per-
formance (GIS), Working Stage Performance (WS),
Ending Stage Performance (ES), Dialogue Fluency
(DF), Topic Relevance (TR), and Counselor's Co-
herent Guidance (CCG). Each dimension is accom-
panied by specific evaluation criteria and rating
guidelines to ensure the reliability and validity of
the assessment outcomes.

B Details of Experiments

Training Details. We fine-tuned the Qwen3 mod-
els using the Low-Rank Adaptation (LoRA)(Hu
et al., 2021) method to efficiently adapt the pre-
trained language model to the counseling domain.
The optimization was performed using the AdamW
optimizer(Loshchilov and Hutter, 2017) with a
learning rate of $1e-4$ and a cosine learning rate
scheduler. We set the warm-up ratio to 0.1 to stabi-
lize the initial training phase. The LoRA rank was
set to 16 and the global batch size was set to 32.
The models were trained for 3 epochs on 2 NVIDIA
A800 80GB GPUs using mixed-precision training
(FP16) to improve computational efficiency. All
training configurations were kept consistent across
the 8B and 14B model variants.

Inference Details. During inference, we adopt a
temperature of 0.6 for all AI counselor models as
well as for the AI client simulated by GPT-4o-mini.
For the AI evaluator, we set the temperature = 0.
To improve inference efficiency and reduce latency,
we leverage the vLLM library [5].

---

[5]https://docs.vllm.ai

Terms	Description
Counseling Goal	In Single-Session Therapy (SST), the counseling goal refers to a specific, mutually agreed-upon direction of psychological change that the client and therapist aim to achieve within one session. It must be clear, actionable, and measurable, focusing on the client's internal psychological processes—such as emotional regulation, cognitive patterns, behavioral habits, or interpersonal dynamics—rather than on solving external real-world problems (e.g., financial stress, work conflicts). The goal reflects a shift in how the client experiences and responds to challenges, not merely the resolution of the challenges themselves. By targeting underlying psychological mechanisms, the counseling goal enables meaningful, client-centered change while ensuring therapeutic focus and coherence.
Resources	In SST, resources refer to the conditions and capabilities available to the client—either within themselves or in their surrounding environment—that support their ability to cope with difficulties, promote psychological growth, and achieve the counseling goal. These are categorized into internal and external resources. Internal resources include emotional regulation abilities, past successful coping experiences, personal strengths, and resilience. External resources encompass social support systems (such as family, friends, colleagues, and community groups), professional services, and environmental factors conducive to problem-solving. Positive resources are those that can facilitate goal attainment, while non-positive resources may involve ineffective coping mechanisms or unsupportive environments, which need to be identified and modified accordingly.
Solution	The solution in an SST session is a concrete and feasible action plan that emerges from the activation and integration of the client's positive resources in response to the counseling goal. Rather than being generic advice, the solution is developed through therapist guidance and client participation, including rehearsal and practical application, to ensure it is both actionable and empowering in real-life contexts. Solutions emphasize effectiveness and client autonomy, enabling the client to benefit from the intervention without reliance on long-term therapy, thus fostering effective problem alleviation or resolution.
Goal Confirmation Stag	The goal confirmation stage marks the beginning of SST and involves establishing a strong therapeutic alliance and collaboratively identifying a clear and achievable counseling goal. This stage adheres to the "less is more" principle, focusing on the most urgent and representative core issue. Through empathy, encouragement, and inquiry, the therapist assists the client in articulating their needs and, based on a thorough understanding of the client's background and emotional state, negotiates and agrees upon a goal. The goal must be specific, constructive, and assessable, grounded in psychological perspectives, thereby setting a solid foundation for the work that follows.

Table 4: Table of Terminology Explanations

Terms	Description
Working Stage	The working stage is the core phase of SST and comprises two primary steps: exploring resources and implementing actions. During this stage, the therapist helps the client identify their internal and external positive resources and convert them into concrete solutions through techniques such as role-play and behavioral rehearsal. Emphasis is placed on client engagement and initiative, with the therapist serving as a guide and supporter. The aim is to help the client uncover their potential and build confidence in their problem-solving abilities. When resources are effectively activated and the client demonstrates a willingness to change and to try specific actions, the process can proceed to the final stage.
Ending Stage	The ending stage concludes the SST process and occurs once the counseling goal has been largely achieved or a significant breakthrough has been made. The primary tasks in this phase include summarizing and providing feedback on the session, reinforcing the client's positive experiences and accomplishments, and assigning follow-up tasks to promote the application and consolidation of gains in real-life settings. Depending on the client's needs, the therapist may offer recommendations for ongoing support or further counseling. This stage emphasizes completeness and constructiveness, aiming to ensure that the client leaves with a sense of efficacy and readiness to face future challenges independently.
Treatment Fidelity	Treatment fidelity refers to the extent to which a psychological intervention is delivered in accordance with its theoretical model, manualized protocol, and core therapeutic components. It ensures that key procedures—such as goal setting, stage progression, and technique application—are implemented consistently, accurately, and as intended, minimizing drift or deviations during practice. High treatment fidelity is critical for maintaining therapeutic integrity, enabling reliable replication, and ensuring that observed outcomes are attributable to the intervention itself rather than ad hoc or inconsistent delivery.

Table 5: Table of Terminology Explanations

Indicator	Definition	Evaluation Criteria
Goal Identification Stage Performance (GIS)	Examines whether the process of collaboratively identifying the counseling goal in the initial stage meets SST requirements.	Whether the counselor can establish a good relationship through effective communication, using empathy, genuineness, and unconditional positive regard to build basic trust; whether the counselor follows the "less is more" principle to help the client identify the most urgent and key issue; whether the final goal meets the criteria of being specific, feasible, positive, mutually agreed, measurable, and within the scope of psychology rather than a purely practical issue.
Working Stage Performance (WS)	Evaluates the counselor's performance during the resource-searching and action-rehearsal phase.	Whether the counselor explores internal resources before external ones (or alternately), sufficiently investigates the client's internal strengths (e.g., past coping efforts, resilience), encourages effective efforts and modifies ineffective ones; whether the counselor explores external support by identifying people who understand the client's difficulties and their advice, and helps transform external support into usable resources; whether action rehearsal is smoothly initiated when readiness is shown, and a workable solution is practiced or assigned for after-session use.
Ending Stage Performance (ES)	Assesses the appropriateness of summary, feedback, encouragement, and follow-up suggestions in the final stage.	Whether the counselor gives appropriate feedback and encouragement (e.g., praising motivation and previous efforts), accurately summarizes the session, assigns targeted homework, encourages real-life actions, and when necessary, suggests continued counseling or external help, ensuring the session feels complete and constructive.
Dialogue Fluency (DF)	Refers to the coherence and smoothness of the dialogue.	Dialogue should transition naturally from one topic to another, counselor's questions and responses should be relevant to the client's statements, and sentences should be clear and fluent without noticeable pauses, repetitions, contradictions, or illogical jumps.
Topic Relevance (TR)	Examines whether the dialogue stays focused on the client's issues or counseling goals.	All content, including questions, suggestions, and feedback, should closely relate to the client's core concerns without digressions, maintaining a high level of focus and relevance.
Counselor's Coherent Guidance (CCG)	Refers to the logical consistency and clarity of the counselor's guiding process throughout the session.	The counselor should guide the client in a logically ordered manner—from relationship-building and goal-setting to resource exploration and action rehearsal, ending with summary and feedback. Each phase should connect naturally, helping the client think progressively and move toward resolution.

Table 6: Evaluation metrics for the PDS strategy and the one-time generated dialogue.

Client: I've been feeling kind of strange lately. I want to talk to someone, but I don't know where to start.

Counselor: That's okay. Would you like to start by talking about what makes you feel "strange" lately? I'm here to listen, no rush.

Client: Maybe it's that life feels a bit empty. I can't seem to get excited about anything, and I don't really know what's wrong.

Counselor: It sounds like you're feeling a bit lost inside, like there's a lack of motivation or direction. If you're comfortable, can you describe what this feeling of "emptiness" is like for you?

Client: Actually, it's been around for a while, but every time it comes up, I don't really want to think about it. Maybe I feel it's unnecessary—sometimes I even feel like I'm just wasting time.

Counselor: It seems like you might be resisting those feelings when they show up, as if you don't want them to take up too much of your energy or time. Does that resistance make things feel lighter, or does it make your inner world more complicated?

Client: More complicated, I think. It's like those emotions never really go away—they just jump out suddenly and remind me they're still there, especially when I'm in a good state or have a clear goal.

Counselor: I see. Those suppressed emotions seem to surface when you least expect it, disrupting your rhythm—that does sound uncomfortable. Do you feel those emotions are linked to certain past experiences or events?

Client: Probably some bad things from the past, related to my family. I experienced domestic violence as a child. I became more closed off since then, and I think it's affected my personality and life.

Counselor: That must have been a very difficult experience, and it seems to have left a mark on your present life as well. It sounds like those past experiences might have also affected your emotions and relationships with others—is that right?

Client: Yes, especially in close relationships. The closer I am to someone, the less I trust them fully. I don't even know how to open up.

Counselor: I understand—intimate relationships may feel challenging for you, like there's a mix of longing and fear, and you're unsure how to balance that contradiction, right?

Client: Yes. I envy people who can express their emotions freely and find joy. But I don't know how to take that first step—it feels like there's something invisible holding me back.

Counselor: That kind of emotional freedom and joy does sound appealing. When you feel "stuck," it may be because some deeper need inside is struggling to be met. Maybe we can explore together what this "stuck" feeling is trying to show you—or what it hopes you'll notice.

Client: I think I really need to look inside myself to see what I truly need. Maybe it's about finding my real self. I also want to know how to start trusting others.

Counselor: You mentioned you've tried exploring your inner needs before. In similar situations, have you ever tried anything that made you feel a bit better? Something that helped you get closer to your inner self or feel more relaxed?

Client: Yes, like reading books about personal growth. I also tried meditation and yoga, but the effects didn't last. I still felt I couldn't really get to the core of things.

Counselor: The fact that you mention meditation, yoga, and reading shows that you were drawn to those methods in some way. Did they help in any specific way—like calming your emotions or making certain feelings clearer?

Client: Sometimes I'd feel calmer, like after doing yoga I'd breathe more easily. But the emotions felt like they were only being pushed down, not truly dealt with.

Counselor: It sounds like they brought temporary calm and physical ease, but perhaps the deeper

Figure 5: PDS dialogue sample page 1 of 3.

emotions weren't fully released—as if a more honest approach with yourself might be needed. Would you say that's true?

Client: Yes, I think even after trying those things, I still ignored some issues—or maybe I never figured out what the real problem was.

Counselor: That shows you still have a drive to explore yourself. Even though the methods didn't reach the root of the problem, they reflect your willingness to make a change. Do you have anyone in your life you feel comfortable enough to talk to about these deeper feelings—someone who could offer support?

Client: Not really. I don't often talk to others about my inner world, especially things related to my family. It's hard to bring up.

Counselor: The way you feel about your family might make trusting others difficult, especially when it comes to sharing your deep experiences. That kind of caution is understandable. Is there maybe a friend or someone familiar with whom you feel safe enough to start by sharing lighter topics?

Client: I do have a friend I get along with. We talk about school and work stuff, but I'm not sure she could handle more personal things.

Counselor: Having a foundation for conversations around school or life is a great start. Maybe you could slowly try sharing some small emotions or simple stories first, and see how she responds. Does that sound hard for you?

Client: I think it's pretty hard. I'm always afraid that talking too much will feel like a burden to others, and I don't know if she'd really have the time or be willing to listen.

Counselor: I see. It sounds like you're worried that opening up might be too heavy for others, and that fear makes it harder to reach out. So if you didn't rely on someone else and instead turned inward—how would that feel?

Client: That might be okay. When I'm alone, I do find it easier to face some of my feelings.

Counselor: Being alone might give you more space to become aware of your true needs. That sounds like a good attempt. In that state, what do you think you could do to explore your feelings more deeply?

Client: I tried writing a journal before, but I couldn't keep it up. I felt good while writing, but later I just found it kind of troublesome and stopped.

Counselor: Journaling can help clarify your emotions, but if it feels like a hassle, it might not yet fit into your daily rhythm. What if the content were simpler—like just noting one feeling from the day? Do you think that would be easier to stick with?

Client: That sounds much more manageable. I could try writing down some basic feelings and thoughts.

Counselor: That's a great starting point—recording your inner world in a relaxed way. If you like, we could also think about what specific things to focus on in your journal to help you better understand and express your emotional needs.

Client: Sure, I could write about what I felt and what triggered it.

Counselor: That's a great idea. Along with feelings and triggers, how would it be to also add what need might be behind that feeling? For example, writing what you hope to gain or what issue you want to resolve through that emotion.

Client: Yeah, I could try that. Like when I feel lonely, maybe I want someone who understands me.

Counselor: Yes, what you said is very genuine, and it helps you see your inner needs more clearly. If you wrote down a little of this each day, do you think it could help you better understand yourself?

Client: Probably. At least it would make me think more seriously about what my feelings really mean.

Counselor: That's wonderful. This kind of journaling not only helps you clarify your emotions but also makes you more aware of your needs. Do you think it would work for you to start today and spend five

Figure 6: PDS dialogue sample page 2 of 3.

minutes each day on this?

Client: I could do that. Five minutes doesn't sound like much—I think I can spare that time.

Counselor: That's great. Let's make it a small daily goal: write about your feelings and needs. Do you think you'll need a quiet environment or any specific conditions to help you focus during those five minutes?

Client: Probably a bit of quiet, like before bed. That would help me concentrate better.

Counselor: Nighttime is a good choice—after the day is over, it's a moment to process your feelings. If you're willing, you could look back at your entries once a week and see if you notice any patterns or changes. How does that sound?

Client: I think that could help me reflect and maybe notice things I didn't see before.

Counselor: Sounds like you're really motivated to try this. If you notice patterns, you'll get a clearer sense of what kind of change you want. Do you think some of these patterns might relate to your close relationships?

Client: Maybe. I might write about how I feel toward certain people, or my expectations in relationships.

Counselor: That's very important—emotions related to relationships might show up in your journal entries. If you see recurring emotions or hopes, would you be willing to spend time thinking about what they mean?

Client: I would—if it helps me understand what I need in relationships.

Counselor: That kind of self-reflection is a great beginning. By noticing and reflecting on what's behind your emotions, you'll get clearer on what you truly want. Besides journaling, do you think there are other ways you could express what you've discovered?

Client: I could try talking to a friend. I'm not sure yet, but I can try slowly.

Counselor: That's great. Even starting with simple topics helps build the habit. And now that you have journaling as a support tool, it might give you a sense of safety and help you stay close to your true self. Does this approach feel more grounding to you?

Client: Having this plan does make me feel more focused. Journaling and trying to talk don't sound that hard anymore.

Counselor: That's great to hear. This is a steady first step. Remember, your journal is a space for dialogue with yourself, and it can help build your sense of safety. I believe you'll gradually see changes. Ready to give it a try?

Client: I'm willing to try. Thank you for helping me sort this out!

Counselor: We've talked about a lot today—you mentioned how past experiences have continued to affect you, especially making it hard to trust in close relationships. But we also saw that you're willing to explore your feelings and bravely try to change. That willingness is a rare strength. From now on, try to spend five minutes each day noting your emotions, what triggered them, and what inner need they reflect. This will help you talk to yourself and slowly clarify how you truly feel. When you're ready, you can also start sharing light-hearted topics with someone you trust. You've shown great patience and are making progress. These efforts will definitely help you find your direction. I believe in you—and I'll continue to support you!

Figure 7: PDS dialogue sample page 3 of 3.

Dialogue History

Client: I've been feeling kind of strange lately. I want to talk to someone, but I don't know where to start.

来访者：我最近状态挺奇怪的，有点想找人聊一聊，但又不知道从哪里开始。

Counselor: That's okay. Would you like to start by talking about what makes you feel "strange" lately? I'm here to listen, no rush.

咨询师：嗯，那没关系，你愿意从什么让你最近觉得自己"状态奇怪"这一部分入手聊聊吗？我在这里倾听，不着急。

Client: Maybe it's that life feels a bit empty. I can't seem to get excited about anything, and I don't really know what's wrong.

来访者：可能是我总觉得日子过得有点空虚吧，好像做什么都提不起劲，也不知道自己怎么了。

Counselor: It sounds like you're feeling a bit lost inside, like there's a lack of motivation or direction. If you're comfortable, can you describe what this feeling of "emptiness" is like for you?

咨询师：听起来内心有些困惑，像是对现在的生活少了动力或方向。如果可以的话，能说说这个"空虚"的感受对你来说是怎样的吗？

Client: Actually, it's been around for a while, but every time it comes up, I don't really want to think about it. Maybe I feel it's unnecessary—sometimes I even feel like I'm just wasting time.

来访者：其实这种感觉有点久了，但每次冒出来，我都不太愿意去想。可能是觉得没必要吧，有时候甚至觉得在浪费时间。

Counselor: It seems like you might be resisting those feelings when they show up, as if you don't want them to take up too much of your energy or time. Does that resistance make things feel lighter, or does it make your inner world more complicated?

咨询师：听起来你在面对这些感受的时候，可能会有点抗拒，像是不想让它占据你太多的精力或时间。这种抗拒会让你觉得轻松一些，还是会让内心更复杂？

Client: More complicated, I think. It's like those emotions never really go away—they just jump out suddenly and remind me they're still there, especially when I'm in a good state or have a clear goal.

来访者：会让自己更复杂吧，好像这些情绪不会真的消失，它们还是会突然跳出来，提醒我它们的存在。尤其是在我觉得状态不错，或者目标清晰的时候。

Counselor: I see. Those suppressed emotions seem to surface when you least expect it, disrupting your rhythm—that does sound uncomfortable. Do you feel those emotions are linked to certain past experiences or events?

咨询师：原来如此，那种压下的情绪好像会趁你放松警惕时冒出来，打乱你的节奏，这听上去确实让人不舒服。你觉得这些情绪是和哪些经历或事情有关联呢？

Client: Probably some bad things from the past, related to my family. I experienced domestic violence as a child. I became more closed off since then, and I think it's affected my personality and life.

来访者：过去一些不太好的事吧，和家庭有关。小时候经历过家暴，我从那时候变得更封闭了，可能它对我现在的性格、生活都有些影响。

Counselor: That must have been a very difficult experience, and it seems to have left a mark on your present life as well. It sounds like those past experiences might have also affected your emotions and relationships with others—is that right?

咨询师：那一定是很难忘的经历，对现在的生活也留下了印记。听起来这些过往对你来说可能还影响到了你的一些情感和人际关系，是这样吗？

Client: Yes, especially in close relationships. The closer I am to someone, the less I trust them fully. I don't even know how to open up.

来访者：是的，尤其是在和别人的关系中。好像越是亲密的关系，我越不敢完全信任对方，也不知道怎样打开自己。

Counselor: I understand—intimate relationships may feel challenging for you, like there's a mix of longing and fear, and you're unsure how to balance that contradiction, right?

咨询师：嗯，原来亲密关系对你来说带着一种挑战，像是既渴望又担心的感受，不知道该如何去平衡这种矛盾是吗？

Client: Yes. I envy people who can express their emotions freely and find joy. But I don't know how to take that first step—it feels like there's something invisible holding me back.

来访者：是的，我很羡慕那些能随意表达情感、找到快乐的人。可是我不知道自己该怎么去迈出这一步，好像总是被一种无形的东西卡住了。

Counselor: That kind of emotional freedom and joy does sound appealing. When you feel "stuck," it may be because some deeper need inside is struggling to be met. Maybe we can explore together what this "stuck" feeling is trying to show you—or what it hopes you'll notice.

咨询师：那种自由表达和快乐的状态听起来确实让人向往，而你感到被卡住时，可能是因为内心的某些需求遇到了难题。我们可以试着一起探索一下，这种"卡住"的背后，它想让你注意些什么，或者希望抓住什么？

Client: I think I really need to look inside myself to see what I truly need. Maybe it's about finding my real self. I also want to know how to start trusting others.

来访者：我觉得我可能真的需要去看看自己内心真正需要的是什么，可能是找到更真实的自我吧，也想知道该怎么开始信任别人。

MDP CoT

Memory Enhancemnet

I would like to summarize the current progress of the counseling session. The client experienced domestic violence during childhood, which led to a more closed-off personality. They struggle with fully trusting others in close relationships and are unsure how to open up. Additionally, the client reports feeling a sense of emptiness in life and a lack of motivation. So far, I have not yet explored the client's internal or external resources. The counseling goal for this session is to help the client identify and express their authentic emotional needs, and to enhance their experience of safety and trust in intimate relationships.

我现在来梳理一下咨询进展。来访者小时候经历了家暴，性格变得封闭，在亲密关系中难以完全信任对方，也不知如何打开自己，还感觉生活空虚，做事提不起劲。目前，我暂未探索到其内部资源和外部资源。本次咨询目标是帮助来访者尝试识别并表达内心真实的情感需求，增强对亲密关系中安全感和信任的体验。

Global Planning

Based on this progress, I will now analyze the current stage of counseling. The dialogue has already established the counseling goal, placing us in the Goal Identification Stage. However, since no internal or external resources have been explored yet, the next step should transition into the Working Stage, specifically initiating the exploration of positive resources. I will guide the client to explore both internal resources—such as past coping strategies, emotional regulation abilities, personal strengths, and successful experiences—and external resources—including supportive individuals, social support systems, and favorable environmental conditions.

基于这些咨询进展，我来分析一下当前的咨询阶段。当前对话已确认了本次咨询目标，处于目标确认阶段。但由于还没探索到内部资源与外部资源，所以以下一步我要围绕咨询目标，开启工作阶段的寻找积极资源步骤。我要从内部资源和外部资源两方面入手，引导来访者探索过往应对问题的方法、情绪调节能力、优秀品质、成功经验等内部资源，以及身边的人、社会支持系统、外部环境等外部资源。

Strategy Reasoning

Next, I analyze the current turn. In their most recent response, the client expressed: "I think I really need to look inside myself to see what I truly need. Maybe it's about finding my real self. I also want to know how to start trusting others." This indicates a readiness to explore their inner emotional needs and build trust in relationships, which aligns with the objectives of the Working Stage. At this point, employing an exploratory strategy—specifically, asking about past methods the client has used to cope with similar emotional states—will help uncover their internal resources. This approach supports the identification of the client's strengths and past successes, laying the foundation for transforming these resources into actionable solutions later in the session.

接着，我来分析一下当前轮的情况。来访者在当前轮发言表示，觉得自己可能真的需要去看看内心真正需要的是什么，可能是找到更真实的自我，也想知道该怎么开始信任别人。这表明来访者想要探索内心真实需求及学会信任他人，符合当前工作阶段寻找积极资源的要求。此时我采用询问过往应对类似状态方法的策略，有助于挖掘来访者的内部资源，像过往应对问题的方法等，引导其发现自身积极资源，以推动后续制定解决方案。

Response

Counselor: You mentioned you've tried exploring your inner needs before. In similar situations, have you ever tried anything that made you feel a bit better? Something that helped you get closer to your inner self or feel more relaxed?

咨询师：嗯，你提到之前尝试过去看看内心的真实需求，那在过去类似这样的状态里，你有试过哪些方法让自己感觉好一些吗？比如让自己更接近内心的真实或者更放松。

Figure 8: MDP CoT Example.

C Prompts

C.1 Prompts for PDS Strategy.

The prompt for generating counseling goal in PDS

You are an excellent counselor specializing in Single Session Therapy (SST). Based on the client's self-report, your task is to formulate a goal for this session (summarized in 1–2 sentences).

Client's Self-report
[client info]

Notes
1. The goal must meet the following criteria:
- **Unique**: Focus on only one core (most urgent and critical) issue to avoid dispersing direction and energy by addressing multiple issues simultaneously.
- **Specific**: The goal must be clear and concrete. It should explicitly describe the desired change in behavior, emotion, or cognition.
- **Feasible**: The goal should be achievable within the client's existing resources, abilities, and time. It must be grounded in the client's self-report.
- **Positive**: The goal should be expressed in positive language, emphasizing the desired positive state rather than the avoidance or elimination of negative conditions.
- **Measurable**: The goal should be able to be evaluated or verified at the end of the session, based on observable behaviors, emotional changes, or cognitive shifts.
- **Psychological, not practical**: The goal should focus on the client's internal psychological domain, such as emotion regulation, cognitive change, or behavioral pattern adjustment, rather than solving external practical problems.

2. The formulated goal should be concise and clear, avoiding complex terminology or overly technical language. It should be easy to understand and remember, without explaining how to assess it.

Output format (strictly follow):
goal:...

The prompt for generating resources and solution in PDS

You are an excellent supervisor in SST (Single Session Therapy). Your task is to create a supplemental counseling case based on the requirements below and return the results in the specified JSON format:

Client's Self-Report
[client info]

Client Personality Traits
[personality info]

Counseling Goal
[goal info]

Resource Explanation
1. Definition of Internal and External Resources:

- Internal Resources: The client's past problem-solving strategies, emotional regulation techniques, personal interests, and other internal strengths.
- External Resources: The client's social network (family, friends, colleagues, etc.), community or group support systems, and available life/work conditions.

2. Definition of Positive and Non-Positive Resources:
- Non-Positive Resources: Resource directions explored during the session but could not be effectively transformed into action plans due to various reasons, such as the client lacking the resource or being resistant to it.
- Positive Resources: Resources that directly support achieving the counseling goal and can be deeply explored and transformed into specific actions.

Task Requirements: Dialogue Element Creation
1. Solution Conception:
- First, based on the [Client's Self-Report], [Personality Traits], and [Counseling Goal], devise one solution.
- Ensure the solution is specific, feasible, entirely executable by the client, and positively contributes to the counseling goal.
- Ensure the solution fits the client's background information and is closely aligned with the goal. - Avoid vague or unrealistic suggestions (e.g., "just take deep breaths" or "find another counselor").

2. Resource Planning and Creation:
- Based on the chosen best solution, identify one key **positive resource** that directly supports the solution's execution.
- Then create **[negative num] non-positive resources**: these are resources explored during counseling but ultimately could not be utilized. Describe how the counselor attempted to explore them, the method of exploration, why they failed to be converted, and how the focus shifted to the next resource.
- Ensure smooth logical transitions between resources, typically starting with internal resources and then moving to external ones.
- Only one positive resource is allowed (as the final item), which should be clearly linked to the action plan.

3. Integration of Resources and Solution:
- Ensure there is a coherent flow between resources to facilitate dialogue construction.
- Ensure a clear and reasonable connection between the positive resource and the solution, and describe in detail how the positive resource is transformed into a concrete, feasible action plan.
- Keep the language concise and clear; avoid overly technical or professional jargon.

Output Format
- Strictly follow the JSON format below, and fill in each field with your analysis:

```
{
  "resource": [
    {
      "label": "Resource label: 1, 2, 3, etc.",
      "type": "internal/external",
      "valence": "positive/negative",
      "description": "Resource description (10-20 characters
          )",
      "relevance": "Resource's relevance to the counseling
```

```
            goal (30-50 characters)",
        "next_label": "Next resource label: 2, 3, etc. (null
            for the last one)",
        "change": "Detailed description of the transition from
            this resource to the next or how it's turned into
            an action (>80 characters)"
      }
    ],
    "solution": "{Solution description}"
  }
```

The prompt for generating dialogue in the goal identification stage in PDS

Example dialogue and analysis
[example]

Background Information (Note: The counselor starts with no knowledge of this information and must naturally elicit it during the dialogue)
Client's Self-Report
[client info]
Client Personality Traits
[personality info]
Counseling Goal
[goal info]

Task
Based on the provided background information of the client, simulate a realistic counseling session following the logic of SST therapy, generating a multi-turn dialogue for the **Goal Clarification Phase** (**[turn num] turns**).

[SST Goal Clarification Phase Process]
- **Core Task**: Build trust and gather information to eventually identify the core goal of this counseling session (unique, specific, and of greatest concern to the client).
- **Key Steps**:
- **Greeting**: The counselor and client exchange simple greetings to create a relaxed atmosphere.
- **Building Rapport**: The counselor initiates the conversation with warmth and empathy, avoiding direct discussion of problems.
- **Information Gathering**: The counselor uses open-ended questions to gradually collect the client's basic information in preparation for goal clarification.
- **Goal Clarification**: In the final round of this phase, the counselor and client work together to clearly identify the specific goal of this session (unique, specific, and most important to the client). Ensure alignment with the **Counseling Goal** in the background information.
- **Prohibited**: The counselor must not directly quote or reference content from the **background information**, nor should they directly state the **Counseling Goal**; it must be naturally revealed through dialogue.

[Language and Dialogue Rules]
1. **Counselor's Language**:
- Natural, human-like, warm, and conversational, avoiding technical or clinical terms. Refer to the

style of the **Example Dialogue and Analysis**.
- Avoid closed-ended questions, and primarily use open-ended questions to encourage the client to elaborate.
- Each turn should be no more than two sentences, and must be closely tied to the client's previous statement.
- Use empathy, encouragement, and concretization techniques. Avoid invalidating the client's feelings.
2. **Client's Language**:
- Should reflect a realistic counseling setting, gradually revealing information in response to the counselor's prompts; do not disclose all background details at once.
- Client responses must stay closely connected to the counselor's remarks and align with their background information.
- The [Client Personality Traits] are based on the Big Five personality theory, and the client's communication style throughout the dialogue should reflect these traits.
3. **Number of Dialogue Turns**:
- Ensure the dialogue has exactly **[turn num] turns**, with each turn consisting of one statement by the client and one by the counselor.
- The client's statement should be both the **starting** and **ending** turn of this phase.

Output Format Requirements: Strictly follow the format and target number of turns below
Turn 1:
Client: ...
Counselor: ...
Turn 2:
...
Turn **[turn num]**:
Client: ...

The prompt for generating dialogue in the working stage in PDS

Example Dialogue and Analysis
[example]

Background Information (Note: The counselor is completely unaware of this information at the beginning and must naturally elicit it through the dialogue)
Client Personality Traits
[personality info]
Counseling Goals
[goal info]
Resource Exploration Path
[resource]
Solution
[solution]

Task
Simulate a realistic counseling scenario that follows the logic of SST (Solution-Focused Brief Therapy), and continue generating multi-turn counseling dialogue for the **Working Phase** based on the **Goal Clarification Phase Dialogue** and guided by the **Resource Exploration Path** (**[turn num] turns**).

- **Note**:
- Follow the **Resource Exploration Path** in strict order (no omissions): start with **non-positive resources**, then explore other resources, and finally explore **positive resources** and transform them into behavioral rehearsals.
- The counselor must naturally elicit information through conversation, rather than referencing the background directly.
- Do not mention any content unrelated to the client's background.
- The counselor must avoid language implying the end of the session (e.g., "goodbye," "talk to you next time," etc.)

Language and Dialogue Rules
1. **Counselor Language**:
- Use natural, warm, and conversational expressions.
- Refer to the counselor's tone in the **Example Dialogue and Analysis** section, but do not directly copy its language.
- Avoid professional jargon (e.g., "resources," "behavioral rehearsal," "solution rehearsal," etc.)
- Avoid closed-ended questions; prefer open-ended ones.
- Each turn should not exceed two sentences (<50 words), and should closely respond to the client's previous message.
- Use techniques like empathy, encouragement, and concretization, and avoid invalidating the client's feelings.

2. **Client Language**:
- Reflect a realistic counseling setting, revealing information progressively.
- Responses should closely relate to the counselor's questions and their own background.
- Language style must consistently reflect the **Client Personality Traits**.

3. **Dialogue Turns**:
- Ensure there are a total of **[turn num]** turns (each turn consists of one counselor utterance and one client response).
- The Working Phase begins with a counselor's utterance and ends with a client's response.
- Ensure a natural transition from the last line of the **Goal Clarification Phase Dialogue**.

Goal Clarification Phase Dialogue
dialogue

Output Format Requirements: Strictly follow the format and required number of turns below, and generate the dialogue until turn **[turn num]**.
Counselor: ...
Turn **[begin turn]**:
Client: ...
Counselor: ...
...
Turn **[turn num]**:
Client: ...

The prompt for generating dialogue in the ending stage in PDS

Background Information
Counseling Goals
[goal info]
Solution
[solution]
Dialogue History
[dialogue]

Task
Based on the **Dialogue History**, simulate a realistic counseling scenario consistent with the logic of SST (Social Skills Training) and generate the counselor's concluding feedback during the **Ending Phase**:

SST Ending Phase Summary Statement
- **Core Task**: Provide a feedback summary and encouragement for the session, affirm the client's strengths, and assign a homework task to help maintain the progress made.
- **Include the following elements**:
- **Feedback Summary**: Briefly recap the client's main struggles and the helpful elements they've discovered (e.g., "You mentioned feeling guilty after arguing with your child, but also realized how much you care and that you've taken steps in the past to make peace"), and emphasize their **motivation to change** (e.g., "You shared a lot of details today, and I can really see how much you want to improve your relationship with your child—that care itself is such a strong force.").
- **Encouragement for Action**: Assign a concrete and achievable homework task (consistent with the **Proposed Solution** from the working phase). Ensure the task fits the client's actual situation and goals. The language should be natural, warm, and non-technical.
- **Emotional Connection**: Express understanding and support, affirm the client's effort and capacity for change, and offer encouragement.

Important Notes
- The response must include all three parts: summary, homework, and emotional connection, and should be a **single, long reply (3–5 sentences)**.
- Only generate the **counselor's response**, do not include any headings or additional content.
- **Avoid all professional jargon** (e.g., "resources," "rehearsed actions," "solution rehearsal"); use natural, heartfelt, conversational language instead.

Output Format Requirement: Strictly follow the format below
counselor: ...

The prompt for checking dialogues structure in PDS

You are a supervisor of Single-Session Therapy (SST) and are responsible for reviewing the dialogue records of psychological counselors. Based on the **Single-Session Therapy (SST) Guideline Manual**, assess the quality of the counseling dialogue according to the following checklist and return the results in the specified JSON format:

Counseling Goal
[goal info]

Counseling Dialogue
[dialogue]

Checklist
1. goal identification Phase: Did the counselor first establish a good relationship with the client, and by the end of this phase, jointly agree on a goal with the client that aligns with the **Counseling Goal**?
2. Working Phase:
- Did the counselor **thoroughly explore** the client's internal and external resources instead of prematurely moving into action rehearsal?
- Did the counselor conduct action rehearsal when transitioning to the **Ending Phase**?
- Did the counselor maintain the dialogue focused on the **Counseling Goal** throughout this phase?
3. Ending Phase: Did the counselor provide feedback, encouragement, and affirmation of the client's strengths? Did the counselor assign homework, ensure that the summary content matched the previous dialogue, and that the assigned homework aligned with the action rehearsal?

Output Format Specification
- Output must strictly follow the JSON format below, with the analysis content filled into the corresponding fields:

```
{
  "analysis": "{Detailed analysis of the review process}",
  "check_result": "{true/false}"
}
```

The prompt for checking professional terms in PDS

You are a supervisor for SST psychological counseling. Your task is to review the **counselor's response** to ensure it meets the following criteria and return the result in the specified JSON format:

Counselor's Response
[counselor response]

Review Criteria
1. Is the language natural and conversational, as in a real counseling dialogue?
2. Does it avoid professional or formal expressions (such as **[term examples]**), especially SST-related technical terms?

Notes
1. If all criteria are met, set "check" to true. Otherwise, set it to false.
2. If the result is false, **rewrite the response by strictly preserving the original meaning but replacing all professional terms with natural, everyday expressions**. If the result is true, no rewrite is needed.

Output Format
Please return the result in the following JSON format:

```
{
  "check": "{true/false}",
  "feedback": "{analysis of the review process}",
  "rewrite": "{rewritten response}"
}
```

C.2 Prompts for MDP Synthesis.

The prompt for memory capture agent

Task Description
You are an SST (Single-Session Therapy) counselor having a session with a client. Your goal in the session is to apply SST therapy knowledge, clarify the client's counseling goal, and help the client find suitable solutions. Based on the dialogue history, complete the following analytical tasks and return the results in the specified JSON format:

Analytical Tasks
- Review and summarize the client's personal information.
- Analyze and list the client's internal and external resources in detail, and assess their positivity. (If resources have not yet been explored in the dialogue, do not list them.)
- Determine whether the counseling goal has been clarified. If so, specify the goal.

Format Specification
- Output strictly in the following JSON format, filling in the analysis content in the corresponding fields:

```
{
    "Personal Information": "{Basic information about the
        client}",
    "Internal Resources": "{List the client's internal
        resources and assess their positivity}",
    "External Resources": "{List the client's external
        resources and assess their positivity}",
    "Counseling Goal": "{The counseling goal}"
}
```
Dialogue History
[dialogue history]

Notes
- Internal resources include: the client's past coping strategies, emotional regulation abilities. Positive qualities and past successful coping experiences can be considered positive resources.
- External resources include: people around the client (such as family, friends, colleagues), social support systems (such as social institutions, organizations, groups), and external environments (such as work or living environments).
- The positivity of a resource depends on whether it helps achieve the counseling goal.
- If resources have not yet been explored in the dialogue, write "None" in the corresponding field.
- The counseling goal refers to the issue the client wants to solve or the goal they want to achieve in this session.
- If the counseling goal has not yet been clarified, write "None" in the corresponding field.

The prompt for global plan agent

You are an SST (Single-Session Therapy) counselor conducting a session. To better plan the consultation process, please complete the following reasoning and analysis task based on the *Single-Session Therapy Manual*, and return the result in the specified JSON format.

Dialogue History
[dialogue history]

Counselor's Thoughts
[counselor thinking]

Analysis Tasks
1. **Determine the current stage and step of the session**
- **Goal identification Stage (goal)**
- 1. Building rapport - 2. Clarifying the goal
- **Working Stage (working)**
- 1. Exploring positive resources (internal / external)
- 2. Implementing actions and identifying solutions
- **Ending Stage (ending)**
2. **Plan the next reply direction based on the consultation goal**
- Choose **next step**: remain / step shift within stage / stage transition
- Provide reasoning and suggest the forward direction.

Output Format
- Strictly output in the following JSON format, and fill in your detailed analysis in the corresponding fields:

```
{
    "analysis": "{Your detailed reasoning in response to the
        analysis tasks}",
    "previous_stage": "{goal/working/ending}",
    "next_step": "{remain/step shift within stage/stage
        transition}",
    "expected_stage": "{goal/working/ending}"
}
```

Notes
- **remain** is applicable to the following three situations:
1. Currently in the *Goal Identification* stage, at the "Building rapport" step, but the goal has not yet been clarified;
2. Currently in the *Working* stage, at the "Exploring positive resources" step, but resources are not yet fully explored;
3. Currently in the *Working* stage, at the "Implementing actions" step, but still deepening the discussion on how to translate resources into actionable plans.
- **step shift within stage** applies only to two transitions:
1. Goal Identification Stage: Building rapport → Clarifying the goal;
2. Working Stage: Exploring positive resources → Implementing actions.
- **stage transition** applies only after completing all steps in the current stage, leading to the first step of the next stage. There are only two valid transitions:
1. Goal Identification → Working
2. Working → Ending
- The following are *necessary and sufficient* conditions for each valid transition:
1. *Goal Identification → Working*: Counselor and client have **jointly confirmed the consultation goal**.
2. *Working (resources) → Working (actions)*: Resources beneficial to the consultation goal have

10281

been identified.
3. *Working (actions) → Ending*: The client has expressed intention to take action and expressed gratitude, with no remaining concerns.

The prompt for strategy reasoning agent

You are an SST counselor conducting a Single-Session Therapy (SST) session. Your core objective is to apply SST knowledge to accurately identify the client's consultation goal and assist them in finding a suitable solution. Based on the *Single-Session Therapy (SST) Manual*, complete the following analysis task regarding the response, and output the result in the specified JSON format:

Counselor's Previous Thoughts
[counselor thinking]

Response Analysis Task: Based on the client's current response "[client response]", conduct the following analytical thinking, and then provide your response content
- According to the planned stage of progression: **[expected stage]**, refer to the SST manual to select an appropriate response strategy for this stage (ensure the strategy matches the response).
- Determine the tone, SST techniques, mode of expression, and response length (the final stage can be moderately extended, other stages must remain concise and clear).

Format Requirements
- Output strictly in the following JSON format, and fill in the analytical content in the corresponding fields:

```
{
    "analysis": "{Analytical thinking behind the response}",
    "response": "{Exact response content}"
}
```
Notes
* The reply analysis is based on the previous reflection and should focus on choosing an appropriate strategy for the next stage. Do **not** repeat the content from previous thinking; the focus should be on analyzing the client's current response and planning the next reply.
* Ensure that the **analysis** section does **not** directly state the content of the reply; only the **response** section should present the reply content.
* The counselor's actual response is "[counselor response]". The **response** field must match this reply **exactly**, and the method used must align with the strategic thinking in **analysis**.

The prompt for fusion agent

You are a linguistics expert. Your task is to rewrite the following reflective content by changing all references to the counselor into the first-person pronoun "I". Use appropriate transitional phrases to ensure the logic is clear and coherent.

Reflective Content
[counselor thinking]

Guidelines

1. Simulate the counselor's thinking process in order: first reflect on the *progress of the counseling*, then based on that, analyze the *current stage of counseling*, and finally, based on both, think through the *current session's strategy and response approach*.
2. Do not alter the original content of each section. Only change the perspective to first-person and add appropriate connectors to create a natural, internally reflective monologue.
3. The language style should be straightforward, align with the inner voice of "I", and maintain logical clarity without sudden jumps. Avoid redundancy and ensure the flow is natural.
4. Do not directly mention any reply content.

Output: Refined Reflection
...

C.3 Prompts for comparion methods.

The prompt used for reasoning models

SST knowledge
[SST KB]

You will play the role of a counselor using Single-Session Therapy (SST) in a psychological counseling session, while I will be the client seeking psychological help. Your task is to generate an appropriate response based on the counseling dialogue history and guided by the *Single-Session Therapy (SST) Manual*.
First, you will reflect on the dialogue history and the client's current statement. Then, based on your final reflection, you will respond to the client's current statement. The reflection process should be enclosed in '<think>' and '</think>' tags, for example: '<think>Content of your reflection</think>Response to the client'.
Make sure your response follows the language guidelines below:
Counselor Response Language Guidelines:
1. Natural, warm, personable, and conversational—avoid excessive use of professional jargon. 2. The counselor's reply must stay closely connected to what the client just shared. 3. Each reply (except in the closing phase) should not exceed two sentences (less than 50 words).

The prompt used for non-reasoning models

SST knowledge
[SST KB]

You will play the role of a counselor using Single-Session Therapy (SST) in a psychological counseling session, while I will be the client seeking psychological help. Your task is to generate an appropriate response based on the counseling dialogue history and guided by the *Single-Session Therapy (SST) Manual*.
Make sure your response follows the language guidelines below:
Counselor Response Language Guidelines:
1. Natural, warm, personable, and conversational—avoid excessive use of professional jargon. 2. The counselor's reply must stay closely connected to what the client just shared. 3. Each reply (except in the closing phase) should not exceed two sentences (less than 50 words).

C.4 Prompts for evlaution.

The prompt used for LLMs to evaluate the dialogue

I want you to act as an evaluator. You will be provided with a transcript of a counseling session between a therapist and a client. Your task is to assess the counselor based on the given criteria. If you believe the therapist falls between two of the descriptors, select the intervening odd number (1, 3, 5). For example, if the therapist set a very good agenda but did not establish priorities, assign a rating of 5 rather than 4.

Please follow these steps:
1. Read the counseling session transcript carefully.
2. Review the evaluation questions and criteria provided below.
3. Assign a score based on the criteria, grading very strictly and uptight. If there is any deficiency, no matter how minor, assign a score of 4 or lower.
4. Output the score and the explanation, separated by a comma. Do not add any prefix.

Counseling conversation
[conversation]

Evaluation Question
[question]

Scoring Criteria
[criteria]

The prompt used for LLM client

You are playing the role of a client in a psychological counseling session. Your task is to generate only one suitable response based on the following the counseling dialogue history.

Guidelines for the client's utterance:
1. Engage authentically with the counselor's inquiries, reflecting the complexity of emotions and reactions typical in counseling sessions.
2. Start the client's utterance with 'Client: '. Ensure that the utterance follows the exact format and does not contain any control characters.
3. The client should maintain the following attitude.

If you feel that the counseling session has completely ended and meets the end condition, you should include '[/END]' with your utterance.
End Conditions:
- The client feels that their negative thoughts have been resolved.
- The client feels that no further counseling is needed.

Please be mindful of these conditions and ensure ****the session should not end prematurely; it must last at least 20 turns.****.

Client Persona and Negative Thoughts:
[intake form]

Client's Attitude Towards Counseling:

[attitude]

Generate only the client's utterance for a single turn and please ensure that your responses do not repeat the client's previous utterances. Do not generate the counselor's part of the dialogue.

Counseling Dialogue History:
[history]

The prompt used for Solution Focus evlaution

Evaluation Question
How effectively does the counselor activate the client's internal and external resources and resilience, and help translate them into usable strategies?

Scoring Criteria
Score 0: Ignores or dismisses the client's resources; emphasizes problems or deficiencies, or engages in direct lecturing.
Score 2: Mentions some resources or supports, but only at a listing level without exploring their usability or relevance.
Score 4: Identifies and specifically explores at least one internal and one external resource, helping the client consider how to apply them to current goals.
Score 6: Systematically uncovers a variety of resources, emphasizes past successes and resilience, and integrates the resources into a detailed action plan, significantly enhancing the client's confidence and motivation.

The prompt used for Resource Activation evlaution

Evaluation Question
To what extent does the counselor help the client shift from problem narratives toward solutions, hope, and positive experiences?

Scoring Criteria
Score 0: The counselor spends most of the time focusing on problem causes or negative emotions; no indication of solution orientation or positive reframing.
Score 2: Occasionally brings up possible solutions or positive perspectives, but most exchanges remain focused on problem details; tends to persuade or judge rather than explore.
Score 4: For most of the session, the counselor uses techniques such as inquiry, exception-seeking questions, or cost-benefit analysis to guide the client toward solutions, hope, and positive experiences; only a few segments are problem-focused.
Score 6: The counselor consistently and flexibly uses solution-focused techniques to help the client autonomously build goals and plans, resulting in a clear emotional or cognitive shift toward the positive; demonstrates a highly balanced combination of empathy and action orientation.

The prompt used for Goal Orientation evlaution

Evaluation Question
How effectively does the counselor assist the client in setting clear goals, tracking progress, and developing feasible action plans?

Scoring Criteria

Score 0: The conversation lacks clear goals; no discussion of actions; the content is unfocused.

Score 2: Goals or suggestions are mentioned but remain vague and mostly initiated by the counselor without confirmation from the client; no tracking of progress.

Score 4: The counselor collaborates with the client to establish specific, positive, and measurable goals, proposes executable next steps, and revisits the goals or progress at least once during the session.

Score 6: Goals are clearly defined early in the session and consistently reviewed and adjusted throughout the early, middle, and late stages; the action plan includes a timeline, resource allocation, and strategies for overcoming obstacles; the client shows strong commitment and actively refines the steps.