

# A Principle-Driven Adaptive Policy for Group Cognitive Stimulation Dialogue for Elderly with Cognitive Impairment

Jiyue Jiang<sup>1,\*</sup>, Yanyu Chen<sup>1,\*</sup>, Pengan Chen<sup>1</sup>, Kai Liu<sup>2</sup>, Jingqi Zhou<sup>2</sup>, Zheyong Zhu<sup>1</sup>, He Hu<sup>3</sup>, Fei Ma<sup>3,†</sup>, Qi Tian<sup>3,4</sup>, Chuan Wu<sup>2,†</sup>

<sup>1</sup>The Chinese University of Hong Kong

<sup>2</sup>The University of Hong Kong

<sup>3</sup>Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ)

<sup>4</sup>Huawei

{jiangjy, chenyanu.cse}@link.cuhk.edu.hk, mafei@gml.ac.cn, cwu@cs.hku.hk

## Abstract

Cognitive impairment is becoming a major public health challenge. Cognitive Stimulation Therapy (CST) is an effective intervention for cognitive impairment, but traditional methods are difficult to scale, and existing digital systems struggle with group dialogues and cognitive stimulation principles. While Large Language Models (LLMs) are powerful, their application in this context faces key challenges: cognitive stimulation dialogue paradigms, a lack of therapeutic reasoning, and static-only user modeling. To address these issues, we propose a principle-driven adaptive policy actualized through a Group Cognitive Stimulation Dialogue (GCSD) system. We first construct a dataset with over 500 hours of real-world CST conversations and 10,000+ simulated dialogues generated via our Principle-Guided Scenario Simulation strategy. Our GCSD system then integrates four core modules to overcome LLM limitations: (i) a multi-speaker context controller to resolve role confusion; (ii) dynamic participant cognitive state modeling for personalized interaction; (iii) a cognitive stimulation-focused attention loss to instill cognitive stimulation reasoning; and (iv) a multi-dimensional reward strategy to enhance response value. Experimental results demonstrate that GCSD significantly outperforms baseline models across various evaluation metrics. Future work will focus on long-term clinical validation to bridge the gap between computational performance and clinical efficacy.

## Introduction

As the global population ages, cognitive impairment has become a significant public health challenge, affecting over 55 million people worldwide (Chib, Sharma, and Singh 2025). Cognitive stimulation therapy (Morley and Cruz-Oliver 2014), particularly in group settings (Sun, Zhang, and Wang 2022; Zhang et al. 2025), is an effective non-pharmacological intervention that maintains or improves cognitive function by fostering social interaction and collective reminiscence (Spector et al. 2003). Despite the proven efficacy of CST, the previous models are constrained by a re-

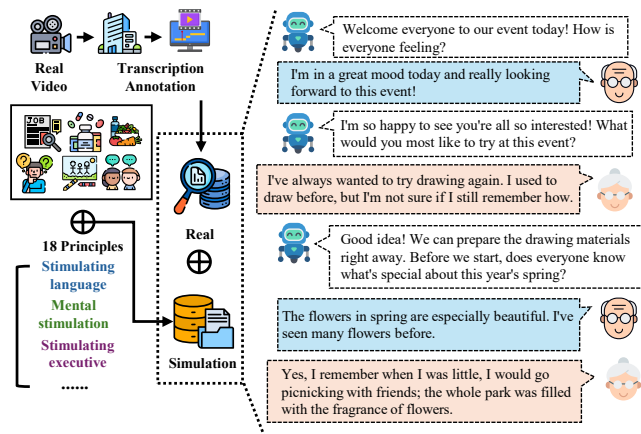


Figure 1: An example of a group cognitive stimulation dialogue. The left panel shows the construction process of real and simulated data, while the right panel presents a sample dialogue provided by the therapist/robot (left) to multiple elders (right).

liance on professional therapists, fixed schedules, and physical locations, which makes them costly and difficult to scale.

Previous cognitive training systems have largely been confined to rigid, question-and-answer formats, making it difficult for them to generate natural, open-ended dialogue (Magyar et al. 2019; Hirose et al. 2020; Tokunaga, Tamura, and Otake-Matsuura 2021; Kim et al. 2021). Due to their limited model scales, they also lack the capacity to learn and internalize the complex, multi-dimensional therapeutic principles behind the 18 cognitive stimulation therapy (Woods et al. 2006; Jiang et al. 2023a). Critically, these systems are entirely incapable of handling complex interactive scenarios such as multi-party dialogue, fundamentally neglecting the core therapeutic elements of group therapy, such as social interaction and peer support.

While large language models have brought revolutionary breakthroughs in recent years (Jiang et al. 2025c,b; Wang et al. 2025; Jiang et al. 2025a), they still face three core challenges when applied to principle-driven multi-party/group

\*These authors contributed equally to this work.

†Corresponding authors: mafei@gml.ac.cn, cwu@cs.hku.hk.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

cognitive stimulation dialogue scenarios: (1) Mismatched Dialogue Paradigm: The core architectures and training paradigms of mainstream LLMs are predominantly designed for dyadic (two-person) interactions, leading to issues such as speaker confusion and context loss when handling multi-party interactions. (2) Lack of Cognitive Stimulation Reasoning: As general-purpose models, they can generate fluent and empathetic text, but they lack the deep, cognitive stimulation-oriented reasoning and strategic guidance required by CST. (3) Static User State Modeling: Existing LLMs struggle to dynamically and continuously model the cognitive state of each elder, a capability that is crucial for providing personalized and adaptive cognitive stimulation.

This paper introduces a principle-driven adaptive policy designed to guide a dialogue system that simulates the cognitive stimulation process of group cognitive stimulation, and enables “anytime, anywhere” access to cognitive stimulation through its digital implementation. To build such a system, we first construct a Cantonese multi-party dialogue dataset by meticulously transcribing and reviewing about 500 hours of real group cognitive stimulation video recordings in collaboration with a third-party company (responsible for the data transcription and annotation). Subsequently, we design the Principle-Guided Scenario Simulation (PGSS) strategy to generate extensive data. This dual-data method serves two purposes: (1) It enriches the diversity of training scenarios by covering specific themes and interaction patterns that are sparse or absent in the real data. (2) It allows the model to pre-learn foundational patterns, enabling it to better grasp the basic framework and principles of multi-party cognitive stimulation dialogue.

Our key contributions are as follows:

- We construct a large-scale dataset with over 500 hours of real-world Cantonese conversations and 10,000+ principle-guided simulated dialogues, providing a key resource for multi-party therapeutic dialogue research.
- We introduce GCSD, a principle-driven adaptive policy for multi-party cognitive stimulation dialogues. It integrates four core modules: a multi-speaker context controller, dynamic participant state modeling, a cognitive stimulation-focused attention loss, and a multi-dimensional reward algorithm to align responses with CST principles.
- Extensive evaluations demonstrate that GCSD significantly outperforms strong commercial and open-source baselines across automatic and human metrics, confirming the effectiveness of our proposed framework.

## Related Work

### Cognitive Training Dialogue Systems

Recent advancements in dialogue systems have shifted from task-oriented designs to open-domain chatbots powered by LLMs (Lee et al. 2016; Mateos-Sanchez et al. 2022; Ni et al. 2023; Sánchez Cuadrado et al. 2024). While significant progress has been made in English-language systems (Yi et al. 2024; Pan et al. 2025), Chinese dialogue systems have also seen notable developments in general-purpose applica-

Categories	Num. (Real)	Num. (Stimulation)
TRAINING DATASET		
<b>Data Count</b>	10,502	9,429
<b>Total Tokens</b>	407,699	295,902
Assistant Count	160,619	119,370
Assistant Tokens	1,644	119,511
Human/Elder Number	9	7
Total Avg Length	38.82	31.38
Assistant Avg Length	1.024	1.001
Human Avg Length	1.006	1.001
TESTING DATASET		
<b>Data Count</b>	1,166	1,047
<b>Total Tokens</b>	45,848	32,862
Assistant Count	18,083	13,238
Assistant Tokens	18,515	13,262
Human/Elder Number	9	6
Total Avg Length	39.32	31.39
Assistant Avg Length	1.024	1.002
Human Avg Length	1.006	1.000

Table 1: Statistics of the multi-party cognitive stimulation dialogue dataset.

tions (Zhou et al. 2021; Gu et al. 2023). However, existing cognitive training dialogue systems often overlook the unique needs of elderly users with cognitive impairment. Current solutions, including robotic photo-based interactions and metamemory therapy (Magyar et al. 2019; Hirose et al. 2020; Tokunaga, Tamura, and Otake-Matsuura 2021; Kim et al. 2021), provide limited cognitive stimulation. Although recent work has introduced specialized datasets for Chinese cognitive training (Jiang et al. 2023a), their effectiveness remains constrained, particularly for low-resource languages like Cantonese.

## Data Construction

### Real Data Construction

Cognitive stimulation therapy is an evidence-based, non-pharmacological intervention for elders with mild to moderate cognitive impairment (Morley and Cruz-Oliver 2014; Zhang et al. 2025). It utilizes themed group sessions to enhance cognitive function, memory, and quality of life.

We process a dataset of about 500 hours of Cantonese group cognitive stimulation video data. The data undergoes manual transcription and annotation by a third-party company. To clean the spoken text, we apply rule-based processing: (1) truncating text exceeding 1000 characters; (2) removing special characters like [S] and [UNK]; (3) retaining only the final symbol in a series of punctuation marks.

To construct multi-party dialogue data, we insert an [Assistant] token before each therapist’s utterance and a [Human\_i] token (where “i” represents a unique individual) before each elder’s utterance. This process yields structured, authentic multi-party cognitive stimulation dialogue data.

## Simulated Data Construction - Principle-Guided Scenario Simulation Strategy

To address the sparsity of specific themes and interaction patterns in real-world data, and to enable the model to pre-learn the foundational frameworks and cognitive stimulation tenets of multi-party dialogue, we construct a large-scale simulated dataset. For this purpose, we design and implement the principle-guided scenario simulation strategy to generate high-quality dialogues using the GPT-4o API<sup>1</sup>.

The core of the PGSS strategy is the construction of a detailed prompt, a process that follows these steps (the complete prompt is provided in *Appendix Prompt*):

- **Define Task and Output Format:** We first define the model’s generation task: to produce multi-party dialogues exceeding 30 turns. We specify that the output must be a structured JSON object containing three fields: turn, speaker, and dialogue, ensuring data consistency and usability.
- **Establish Roles and Scenarios:** We establish the participant roles, comprising one therapist responsible for facilitation and five to six patients with cognitive impairments, each assigned a unique personal history and interests. The scenario is set in a well-equipped activity room to support diverse interactions.
- **Provide Activities and Prompts:** We provide a list of common activity categories from CST (art creation, thematic discussion, etc.) and instruct the model to randomly select from this list. In addition, we supply dialogue prompts for the opening, middle, and closing phases of a session to guide the model in generating conversations that align with the cognitive stimulation flow.
- **Explicitly List Constraining Principles:** The most critical step is that we explicitly list the 18 CST principles within the prompt (“Encourage new ideas”, “Value opinions”, “Use reminiscence”). We then require that the model’s generation, particularly the therapist’s facilitative utterances, must strictly adhere to these listed principles throughout the entire process.

The statistics of the real data and the simulated data (training and testing dataset) are presented in Table 1.

## Methodology: GCSD

### Overview

Figure 2 gives an overview of our principle-driven adaptive policy, which is instantiated by the GCSD system. This policy is operationalized through four key stages: (1) multi-party context control; (2) dynamic participant cognitive state modeling; (3) cognitive stimulation-focused attention loss; and (4) multi-reward policy optimization.

### Multi-Party Context Control

Effectively managing speaker roles is a foundational requirement for coherent multi-party dialogue. To enable

our model to learn the distinct characteristics of multiple elders, we adopt an established data formatting approach by employing special tokens. Specifically, we use [Assistant] to delineate utterances from the therapist and [Human.i] ([Human\_1], [Human\_2]) for each elder. For instance, a typical dialogue segment is structured as follows ( [|im\_start|> is the start symbol, and <|im\_end|> is the end symbol):

```
<|im_start|> [Human_1] Mid-Autumn Festival is coming, which reminds me of the mooncakes we used to eat.
[Human_2] Yes, and we used to play with lanterns when we were young. <|im_end|>
[Assistant] Both are wonderful memories. If you had to choose one, which do you feel better represents the festival? <|im_end|>
[Human_2] I would choose the lanterns, because my whole family would make them together.
<|im_end|> .....
```

This method ensures the model can accurately attribute utterances to speaker, maintaining role consistency and contextual coherence. In addition, we adopt a two-stage training strategy: first, continue training on the simulated dataset to learn structured, principle-aligned dialogue flows via PGSS; then, fine-tuning on real-world data to capture authentic linguistic nuances and conversational dynamics.

### Dynamic Participant Cognitive State Modeling

To overcome the static user state modeling limitations of standard LLMs, the dynamic participant state modeling module generates a dynamic soft prompt for each participant. This prompt, a continuous vector, is produced by a compact neural network designed to distill complex user state information (i.e., cognitive state, historical interactions) into a concise representation. This network operates in parallel to the dialogue model, dynamically modulating the generation to achieve personalized and adaptive responses.

**Soft Prompt Generation Network.** The generator is an efficient multi-layer perceptron (MLP) that takes a concatenated input vector  $X_{in}$  comprising various user state features. The network architecture is as follows:

- **Input Layer:** A linear layer transforms the input features  $X_{in}$  into a latent representation  $L_1$ , followed by a GELU activation (Hendrycks and Gimpel 2016).

$$L_1 = \text{GELU}(X_{in}W_1 + b_1) \quad (1)$$

where  $W_1 \in \mathbb{R}^{D_{in} \times 512}$ .  $D_{in}$  is the dimension of the concatenated input features.

- **Hidden Layers:** The attention block and the final projection can be written compactly in two balanced lines:

$$L_2 = \text{GELU}\left(\text{softmax}\left(\frac{(L_1W_Q)(L_1W_K)^T}{\sqrt{d_k}}\right)\right) \quad (2)$$
$$(L_1W_V)W_2 + b_2)$$

<sup>1</sup><https://openai.com/index/hello-gpt-4o/>

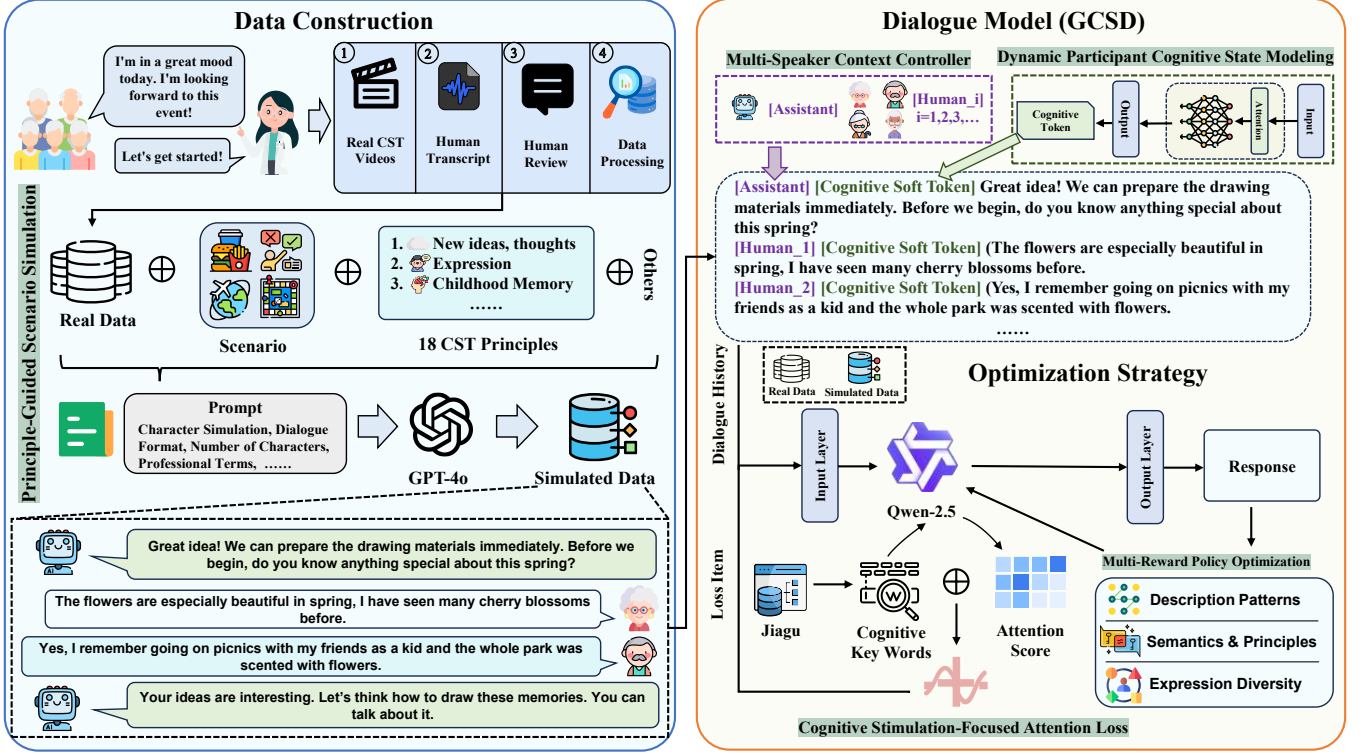


Figure 2: This is an overview of the GCSF. On the left side is the data construction phase, which includes the workflows for both real data and simulated data construction (principle-guided scenario simulation strategy), as well as an example shown at the lower left. On the right side is the GCSF model phase, where the model first learns patterns and scenarios based on simulated data, and then trains on real data. The model consists of four parts: multi-speaker context controller, dynamic participant cognitive state modeling, cognitive stimulation-focused attention Loss, and multi-reward policy optimization.

where  $Q = L_1 W_Q$ ,  $K = L_1 W_K$ ,  $V = L_1 W_V$  with  $W_Q, W_K, W_V \in \mathbb{R}^{512 \times d_k}$  ( $d_k = 512$ ), and the output layer parameters are  $W_2 \in \mathbb{R}^{512 \times 256}$  and  $b_2 \in \mathbb{R}^{256}$ .

- **Output Layer:** A final linear layer produces the soft prompt vector  $P_{\text{soft}} \in \mathbb{R}^{D_{\text{prompt}}}$ , where the dimension  $D_{\text{prompt}}$  matches the main model’s embedding dimension (i.e., 512). A tanh activation function bounds the output values, enhancing training stability.

$$P_{\text{soft}} = \tanh(L_2 W_{\text{out}} + b_{\text{out}}) \quad (3)$$

where  $W_{\text{out}} \in \mathbb{R}^{256 \times D_{\text{prompt}}}$ .

**Integration and Smoothing.** The generated soft prompt  $P_{\text{soft}}$  is prepended to the input token embeddings of the main Transformer-based dialogue model at each decoding step:

$$\text{Input}_{\text{Transformer}} = [P_{\text{soft}}; E_{\text{token}_1}; E_{\text{token}_2}; \dots; E_{\text{token}_k}] \quad (4)$$

where  $E_{\text{token}_i}$  are the learned embeddings of the input tokens. By treating  $P_{\text{soft}}$  as a “virtual token” embedding, it influences the self-attention mechanisms, conditioning the model’s output on the elder’s dynamically updated state.

To prevent erratic fluctuations in the prompt, which is critical when interacting with a vulnerable population, we apply a temporal smoothness regularization term:

$$\mathcal{L}_{\text{Smoothness}} = \|P_{\text{soft},t} - P_{\text{soft},t-1}\|_2^2 \quad (5)$$

This ensures that the personalized guidance evolves smoothly over the course of the interaction.

### Optimization Strategy

Our training process is divided into two core phases: (1) Supervised Fine-Tuning (SFT), which teaches the model to generate fluent, coherent, and therapeutically-focused responses; and (2) Policy Optimization, which further aligns the model’s behavior with multi-dimensional cognitive stimulation objectives using a reward-based algorithm.

**Phase 1: Supervised Fine-Tuning with Cognitive Stimulation-Focused Attention.** In the supervised fine-tuning phase, we optimize the model using a joint loss function composed of three components:

**1. Generation Loss ( $\mathcal{L}_{\text{Gen}}$ ):** This is the standard autoregressive cross-entropy loss, ensuring linguistic fluency and contextual relevance (Radford et al. 2019; Yang et al. 2025).

$$\mathcal{L}_{\text{gen}} = - \sum_{i=1}^N \log P(y_i | y_{<i}, \mathcal{U}) \quad (6)$$

where  $y_i$  is the target token at step  $i$ , given the context  $\mathcal{U}$  and the preceding tokens  $y_{<i}$ .  $N$  is the total sequence length.

**2. Cognitive Stimulation-Focused Attention Loss ( $\mathcal{L}_{\text{CSFAL}}$ ):** To instill cognitive stimulation reasoning, we

---

**Algorithm 1: GCSD Training Pipeline**


---

```

1: Input: Datasets  $D_{\text{real}}, D_{\text{simulated}}$ ; Base model  $M(\theta)$ .
2: Hyperparams: SFT weights  $\gamma_1, \gamma_2, \gamma_3$ ; MRPO group size  $G$ , KL penalty  $\beta$ .
3: Output: Optimized  $\theta_{\text{final}}$ .
4: procedure TRAINGCSD( $D_{\text{real}}, D_{\text{simulated}}, \theta$ )
5:    $D_{\text{SFT}} \leftarrow \text{ProcessData}(D_{\text{real}} \cup D_{\text{simulated}})$ 
6:    $D_{\text{prompt}} \leftarrow \text{ExtractPrompts}(D_{\text{SFT}})$ 
7:    $\theta_{\text{SFT}} \leftarrow \text{SFT\_Phase}(D_{\text{SFT}}, \theta)$ 
8:    $\theta_{\text{final}} \leftarrow \text{MRPO\_Phase}(D_{\text{prompt}}, \theta_{\text{SFT}})$ 
9:   return  $\theta_{\text{final}}$ 
10: end procedure
11: procedure SFT_PHASE( $D_{\text{SFT}}, \theta$ )
12:    $\theta_{\text{SFT}} \leftarrow \theta, P_{\text{soft, prev}} \leftarrow \mathbf{0}$ .
13:   for batch  $(U, Y) \in D_{\text{SFT}}$  do
14:      $P_{\text{soft, curr}} \leftarrow \text{SoftPromptNet}(\text{UserFeatures}(U))$ 
15:      $L_{\text{smooth}} \leftarrow \|P_{\text{soft, curr}} - P_{\text{soft, prev}}\|_2^2$ 
16:     Logits,  $A \leftarrow M_{\theta_{\text{SFT}}}(\text{Concat}(P_{\text{soft, curr}}, \text{Embed}(U)))$ 
17:      $L_{\text{total}} \leftarrow \gamma_1 \text{CrossEntropyLoss}(\text{Logits}, Y) + \gamma_2 \text{CalculateCSFAL}(A, \text{Keywords}(Y)) + \gamma_3 L_{\text{smooth}}$ 
18:     Update  $\theta_{\text{SFT}}$  via  $\nabla L_{\text{total}}$ .
19:      $P_{\text{soft, prev}} \leftarrow P_{\text{soft, curr}}$ 
20:   end for
21:   return  $\theta_{\text{SFT}}$ 
22: end procedure
23: procedure MRPO_PHASE( $D_{\text{prompt}}, \theta_{\text{SFT}}$ )
24:    $\pi_{\theta} \leftarrow \theta_{\text{SFT}}; \pi_{\text{ref}} \leftarrow \text{frozen } \theta_{\text{SFT}}$ .
25:   for training step do
26:     Sample  $q \in D_{\text{prompt}}$ .
27:     Generate  $\{o_i\}_{i=1}^G \sim \pi_{\theta}(\cdot|q)$ .
28:     Compute  $R_i \leftarrow \text{CompositeReward}(o_i, \text{reference})$ .
29:     Compute advantage estimates  $\{\hat{A}_{i,t}\}$ .
30:     Compute MRPO objective  $J_{\text{MRPO}}(\theta)$ .
31:     Update  $\theta$  by ascending  $\nabla J_{\text{MRPO}}(\theta)$ .
32:   end for
33:   return  $\theta$ 
34: end procedure

```

---

design a loss to guide the model’s attention mechanism. We first use a keyword extraction tool<sup>2</sup> to identify keywords from the reference response. We then compute a weighted mean squared error between the model’s attention weights and a target distribution derived from these keywords.

$$\mathcal{L}_{\text{CSFAL}} = \frac{1}{M} \sum_{j=1}^M \lambda_j (a_j - \eta_j)^2 \quad (7)$$

where  $M$  is the length of the token sequence;  $a_j$  is the model’s attention weight on token  $j$  (averaged across heads);  $\eta_j$  is the target attention score for token  $j$ . If the token is an extracted keyword,  $\eta_j > 0$ ; otherwise, it is 0;  $\lambda_j$  is a saliency weight, such as  $\exp(\kappa \cdot \eta_j)$ , that amplifies the penalty for critical items, where  $\kappa$  is a tunable hyperparameter.

<sup>2</sup><https://github.com/ownthink/Jiagu>

**3. Smoothness Loss ( $\mathcal{L}_{\text{Smoothness}}$ ):** This term, from the Dynamic Participant Cognitive State Modeling module, ensures the generated soft prompts evolve smoothly over time.

The joint loss ( $\mathcal{L}_{\text{SFT}}$ ) for the supervised fine-tuning phase is a weighted sum of these components:

$$\mathcal{L}_{\text{SFT}} = \gamma_1 * \mathcal{L}_{\text{Gen}} + \gamma_2 * \mathcal{L}_{\text{CSFAL}} + \gamma_3 * \mathcal{L}_{\text{Smoothness}} \quad (8)$$

where  $\gamma_1, \gamma_2$ , and  $\gamma_3$  are hyperparameters that balance the influence of each loss component.

**Phase 2: Multi-Reward Policy Optimization (MRPO).** Following SFT, the model’s policy is further refined through our multi-reward optimization framework, which is adapted from the GRPO algorithm (Guo et al. 2025).

For each prompt  $q$ , we sample a group of  $G$  candidate outputs  $\{o_1, \dots, o_G\}$  from the current policy  $\pi_{\theta}$ . The objective is to maximize the following function:

$$J_{\text{MRPO}}(\theta) = \mathbb{E}_{q, \{o_i\}} \left[ \frac{1}{G} \sum_{i=1}^G \frac{1}{|o_i|} \sum_{t=1}^{|o_i|} \left( \frac{\pi_{\theta}(o_{i,t}|q, o_{i,<t})}{\pi_{\theta_{\text{old}}}(o_{i,t}|q, o_{i,<t})} \hat{A}_{i,t} - \beta \cdot \text{KL}_{\text{term}} \right) \right] \quad (9)$$

where  $\hat{A}_{i,t}$  is the advantage estimate,  $\pi_{\theta_{\text{old}}}$  is the old policy (set to  $\pi_{\theta}$  in practice), and  $\beta$  is a coefficient controlling the KL penalty. The KL term regularizes the divergence between the policy  $\pi_{\theta}$  and a reference model  $\pi_{\text{ref}}$  (i.e., the SFT model) to prevent catastrophic forgetting.

**Multi-Dimensional Reward Mechanism:** The advantage estimate  $\hat{A}_{i,t}$  is derived from a rule-based reward model that combines several signals into a single scalar value. To capture established conversational patterns, the model is rewarded for n-gram overlap with reference responses, measured via BLEU-4. To encourage the learning of latent semantics and principles, we evaluate semantic similarity using BERTScore. Response diversity, which is crucial for engagement, is promoted by rewarding unique tokens, as quantified by Distinct-2 metrics. Finally, a binary reward enforces structural correctness, such as the proper use of [Assistant] tokens. This composite reward is the core mechanism for enforcing the ‘principle-driven’ aspect of our policy, driving the MRPO process to optimize the model’s outputs for cognitive stimulation relevance and quality.

## Experiments

### Implementation Details

The base model of the GCSD is Qwen-2.5-3b<sup>3</sup>. All models are implemented with PyTorch (Paszke et al. 2019) on one NVIDIA A100-80G GPU. Training is orchestrated using the AdamW optimizer (Loshchilov 2017) with decoupled weight decay ( $\lambda = 0.01$ ). We employ a *Linear Warmup with Cosine Annealing Learning Rate Schedule* (Waswani et al. 2017), where the learning rate linearly increases from  $10^{-7}$  to a peak of  $5 \times 10^{-5}$  over the initial 5% of training steps, subsequently annealing down to  $10^{-7}$  following

<sup>3</sup><https://huggingface.co/Qwen/Qwen2.5-3B-Instruct>

Models	ROUGE-L	BLEU-2	BLEU-4	BERTScore	Distinct-2	Relevance	Empathy	Fluency
SIMULATED DATASET								
ERNIE	17.24	8.32	7.38	65.35	53.92	3.05	3.20	2.90
Doubao-Pro	23.58	22.39	21.51	73.19	71.03	4.00	3.55	3.33
GPT-4o	21.78	20.38	20.05	72.46	67.83	3.67	3.65	3.45
DeepSeek-671B	21.67	22.13	22.09	77.29	<b>73.79</b>	3.59	3.45	3.36
LLaMA-3.1-405B	22.56	22.13	21.08	76.52	71.26	3.75	3.70	3.42
GLM-4-Plus (5-Shot)	17.69	18.24	18.11	70.18	72.97	3.45	3.15	3.15
ChatGPT (5-Shot)	21.57	18.06	15.44	72.18	52.86	3.95	3.50	3.33
DeepSeek-671B (5-Shot)	22.89	24.22	23.59	76.26	72.31	4.02	3.57	3.40
Qwen-Max (5-Shot)	21.54	21.34	20.15	74.27	64.41	3.65	3.60	3.38
GCS D w/o DPSM	21.53	24.45	22.18	79.33	56.41	3.60	3.54	3.32
GCS D w/o CSFAL	23.58	25.16	22.64	74.19	68.42	3.58	3.58	3.33
<b>GCS D-3b</b>	<b>26.92</b>	<b>28.32</b>	<b>26.83</b>	<b>79.45</b>	72.64	<b>4.20</b>	<b>3.67</b>	<b>3.50</b>
REAL DATASET								
ERNIE	18.32	11.32	9.16	66.93	56.28	3.33	2.95	3.00
Doubao-Pro	24.87	23.59	22.36	75.16	71.46	3.95	3.33	3.20
GPT-4o	25.76	21.53	20.14	73.79	69.15	4.00	3.45	3.35
DeepSeek-671B	22.54	23.58	22.42	79.98	<b>76.86</b>	4.08	3.45	3.46
LLaMA-3.1-405B	23.87	23.68	23.52	77.13	71.52	4.00	3.45	3.40
GLM-4-Plus (5-Shot)	18.54	18.53	17.39	70.32	72.18	3.50	3.27	3.33
ChatGPT (5-Shot)	22.32	18.23	16.19	73.54	55.38	3.67	3.33	3.25
DeepSeek-671B (5-Shot)	23.58	24.83	24.36	78.27	73.27	4.10	3.48	3.42
Qwen-Max (5-Shot)	22.87	21.56	20.53	73.63	64.76	3.67	3.40	3.38
GCS D w/o CT	24.33	27.14	26.51	75.42	74.19	3.99	3.45	3.40
GCS D w/o DPSM	23.57	25.63	23.15	79.48	60.13	3.74	3.40	3.32
GCS D w/o CSFAL	25.18	26.16	24.98	76.33	70.46	3.69	3.39	3.36
<b>GCS D-3b</b>	<b>27.63</b>	<b>28.53</b>	<b>27.93</b>	<b>80.12</b>	74.82	<b>4.15</b>	<b>3.50</b>	<b>3.53</b>

Table 2: Evaluation results between baselines and our GCS D on the **cognitive stimulation simulated and real testing dataset**. The first four metrics are automatic metrics, while the last three metrics are human metrics. **Bold** denotes the best results.

a cosine schedule. A micro-batch size of 1 is utilized, accumulated to an effective batch size of 16 through gradient accumulation steps. Furthermore, *mixed-precision training (FP16)* is employed to accelerate computation and reduce memory footprint. During inference, a temperature of 0.6 and a top- $p$  value of 0.95 are set for stochastic decoding, balancing generation diversity with coherence, consistent with best practices in large language model inference (Liu et al. 2024; Guo et al. 2025; Shao et al. 2024). For evaluation, the ROUGE-chinese package (version 1.0.3) is used to compute the ROUGE metric, the NLTK package (version 3.9.1) is used to compute the Bleu metric, and the bert-score package (version 0.3.10) is used to compute BERTScore.

## Evaluation

**Automatic Evaluation.** In the generation of cognitive stimulation dialogues, we utilize the ROUGE-L metric (Lin 2004; Ganesan 2018) and BLEU-4 metric (Papineni et al. 2002) to evaluate the quality of the text. In order to better evaluate semantic similarity, we also implement BERTScore (Zhang et al. 2019a). BERTScore assists in measuring the similarity between the embeddings of the generated sentence and the reference sentence. In addition, we employ Distinct (Li et al. 2015), which quantify the pro-

portion of uni-gram, respectively, across all generated outputs, thereby providing an indicator of diversity.

**Human Evaluation.** To qualitatively examine model performance, human evaluations are conducted. Dialogues from the GCS D and the baselines are sampled. Six elders and their relatives participate in evaluating the responses generated by different models. Following (Li et al. 2020; Liu et al. 2021; Li et al. 2022; Jiang et al. 2023b), all models undergo evaluation in terms of **Relevance**, **Empathy**, **Fluency**. **Relevance** assesses whether the generated responses align with the dialogue context. **Empathy** determines if the LISTENER understands the SPEAKER’s feelings. **Fluency** assesses the grammatical correctness and readability of the SPEAKER’s responses. Each metric is rated on a five-point scale, where 1, 3, and 5 represent unacceptable, moderate, and excellent performance, respectively.

## Baselines

For a comprehensive comparison, we evaluate our GCS D system against a suite of powerful, mainstream large language models. These include our base model, **Qwen-2.5-3b** (Yang et al. 2025), as well as other prominent models such as Baidu’s **ERNIE** (Zhang et al. 2019b), ByteDance’s

**Doubao-Pro<sup>4</sup>**, **DeepSeek-671B** (Liu et al. 2024), Meta’s **LLaMA-3.1-405B** (Dubey et al. 2024), Zhipu AI’s **GLM-4-Plus** (GLM et al. 2024), Alibaba’s **Qwen-Max** (Bai et al. 2023), and OpenAI’s **ChatGPT<sup>5</sup>** and its flagship model, **GPT-4o** (Hurst et al. 2024).

To validate the contribution of each component and our data strategy, we conduct an ablation study. **w/o CT** bypasses the initial training on simulated data, fine-tuning the base model exclusively on the real dataset. **w/o DPSM** ablates the dynamic participant state modeling module. **w/o CSFAL** excludes the cognitive stimulation-focused attention loss.

## Results

### Experimental Results

Our GCSD model achieves state-of-the-art performance on the real-world dataset, consistently outperforming strong baseline models (Table 2). The core insight from the results is that GCSD’s primary advantage stems not merely from generating semantically relevant content, but from its mastery of the specific, principle-guided interaction structures inherent to cognitive stimulation therapy. This is quantitatively demonstrated by its **BLEU-4** score of **27.93**, a **14.7%** improvement over the strongest 5-shot baseline, providing clear evidence that it effectively learns cognitive stimulation phrasing and multi-party turn-taking patterns. This structural advantage is particularly significant when contrasted with semantic similarity, where the gap between GCSD (**BERTScore: 80.12**) and the top baseline (79.98) is narrow. This structural command is further validated by human evaluations, where GCSD attains the highest score in Relevance (4.15), a critical metric for multi-party dialogue that indicates the model successfully overcomes challenges like speaker confusion. Ultimately, human A/B testing confirms its practical superiority, showing a decisive win rate against established models like ERNIE (75%) and a net positive preference over the powerful GPT-4o (50% win vs. 39% loss), validating its suitability for this specialized domain (Table 3).

The ablation study confirms that GCSD’s effectiveness stems from the synergistic integration of its components, rather than any single module. Bypassing the simulated data pre-training (w/o CT) leads to a noticeable performance drop (**BLEU-4** from **27.93** to **26.51**), highlighting the value of the dual-data strategy in teaching the model the foundational framework of CST. The performance degradation observed when removing the dynamic participant state modeling (w/o DPSM) and the cognitive stimulation-focused attention loss (w/o CSFAL) also confirms their respective contributions to personalized interaction and therapeutic reasoning. Therefore, GCSD’s success is not attributable to a single “silver bullet” but to the methodical integration of all four complementary modules.

<sup>4</sup><https://www.doubao.com/chat/>

<sup>5</sup><https://chatgpt.com/>

Models	Win	Loss	Tie
GCSD-3b vs ERNIE	75%	10%	15%
GCSD-3b vs GLM-4-Plus	73%	12%	15%
GCSD-3b vs Doubao-Pro	60%	23%	17%
GCSD-3b vs ChatGPT	58%	27%	15%
GCSD-3b vs GPT-4o	50%	39%	11%
GCSD-3b vs LLaMA-3.1-405B	58%	34%	8%
GCSD-3b vs DeepSeek-671B	43%	40%	17%
GCSD-3b vs Qwen-Max	55%	29%	16%

Table 3: Result of human A/B test.

### Ablation Study

The results of the ablation study (Table 2) clearly demonstrate that GCSD’s superior performance stems from the synergistic effect of its integrated components, rather than the contribution of any single module. In addition, the value of our dual-data strategy is confirmed, as removing the simulated data pre-training (w/o CT) resulted in a competitive but lower BLEU-4 score of 26.51, highlighting this phase’s role as a crucial “primer” for the model.

The study also validated the contributions of the more granular architectural components. Removing the Dynamic Participant Cognitive State Modeling (w/o DPSM) led to a noticeable performance drop (BLEU-4 to 23.15), confirming its importance in achieving personalized and adaptive interactions. Similarly, excluding the cognitive stimulation-focused Attention Loss (w/o CSFAL) also resulted in a lower BLEU-4 score of 24.98, showing its effectiveness in instilling the necessary therapeutic reasoning. In summary, GCSD’s leading performance is not due to a single “silver bullet” but is the result of the methodical integration and synergy of all four complementary components.

## Conclusion and Outlook

This paper proposes a principle-driven adaptive policy for multi-party cognitive stimulation dialogue, namely GCSD. By leveraging a large-scale real and simulated dataset and four core model modules, GCSD addresses the challenges of multi-party dialogue management, personalized cognitive state modeling, and alignment with cognitive stimulation principles. Experimental results show that GCSD achieves strong performance on both automatic and human evaluation in cognitive stimulation dialogue.

Looking ahead, our future work will proceed in three directions. First, we aim to conduct long-term clinical trials to validate the therapeutic efficacy of GCSD in real-world care settings, bridging the gap between computational metrics and clinical outcomes. Second, we plan to extend the current text-based framework to a multimodal system that incorporates visual and acoustic cues, enabling more empathetic and responsive interactions for elders with varying degrees of impairment. Finally, we will investigate safety mechanisms to prevent hallucinations in medical contexts, ensuring that the system remains a reliable and ethical companion for cognitive health support.

## Acknowledgments

We want to thank our anonymous AC, SPC and reviewers for their feedback. We acknowledge support from The University of Hong Kong and Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ). This work was supported by Hong Kong Innovation and Technology Commission's Innovation and Technology Fund (Award No. ITS/269/22FP).

## References

- Bai, J.; Bai, S.; Chu, Y.; Cui, Z.; Dang, K.; Deng, X.; Fan, Y.; Ge, W.; Han, Y.; Huang, F.; et al. 2023. Qwen technical report. *arXiv preprint arXiv:2309.16609*.
- Chib, S.; Sharma, K.; and Singh, R. 2025. Interventions of cognitive impairment in older adults: A comprehensive review. *Aging and Health Research*, 100223.
- Dubey, A.; Jauhri, A.; Pandey, A.; Kadian, A.; Al-Dahle, A.; Letman, A.; Mathur, A.; Schelten, A.; Yang, A.; Fan, A.; et al. 2024. The llama 3 herd of models. *arXiv e-prints*, arXiv-2407.
- Ganesan, K. 2018. Rouge 2.0: Updated and improved measures for evaluation of summarization tasks. *arXiv preprint arXiv:1803.01937*.
- GLM, T.; Zeng, A.; Xu, B.; Wang, B.; Zhang, C.; Yin, D.; Rojas, D.; Feng, G.; Zhao, H.; Lai, H.; Yu, H.; Wang, H.; Sun, J.; Zhang, J.; Cheng, J.; Gui, J.; Tang, J.; Zhang, J.; Li, J.; Zhao, L.; Wu, L.; Zhong, L.; Liu, M.; Huang, M.; Zhang, P.; Zheng, Q.; Lu, R.; Duan, S.; Zhang, S.; Cao, S.; Yang, S.; Tam, W. L.; Zhao, W.; Liu, X.; Xia, X.; Zhang, X.; Gu, X.; Lv, X.; Liu, X.; Liu, X.; Yang, Y.; Song, X.; Zhang, X.; An, Y.; Xu, Y.; Niu, Y.; Yang, Y.; Li, Y.; Bai, Y.; Dong, Y.; Qi, Z.; Wang, Z.; Yang, Z.; Du, Z.; Hou, Z.; and Wang, Z. 2024. ChatGLM: A Family of Large Language Models from GLM-130B to GLM-4 All Tools. arXiv:2406.12793.
- Gu, Y.; Wen, J.; Sun, H.; Song, Y.; Ke, P.; Zheng, C.; Zhang, Z.; Yao, J.; Liu, L.; Zhu, X.; et al. 2023. Eva2. 0: Investigating open-domain chinese dialogue systems with large-scale pre-training. *Machine Intelligence Research*, 20(2): 207–219.
- Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.
- Hendrycks, D.; and Gimpel, K. 2016. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.
- Hirose, S.; Kitakoshi, D.; Yamashita, A.; Suzuki, K.; and Suzuki, M. 2020. A Study on Intelligent Dialogue Agent for Older Adults' Preventive Care—Towards Development of a Comprehensive Preventive Care System—. *Advances in Science, Technology and Engineering Systems Journal*, 5(6): 09–21.
- Hurst, A.; Lerer, A.; Goucher, A. P.; Perelman, A.; Ramesh, A.; Clark, A.; Ostrow, A.; Welihinda, A.; Hayes, A.; Radford, A.; et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.
- Jiang, J.; Chen, P.; Chen, L.; Wang, S.; Bao, Q.; Kong, L.; Li, Y.; and Wu, C. 2025a. How Well Do LLMs Handle Cantonese? Benchmarking Cantonese Capabilities of Large Language Models. In Chiruzzo, L.; Ritter, A.; and Wang, L., eds., *Findings of the Association for Computational Linguistics: NAACL 2025*, 4464–4505. Albuquerque, New Mexico: Association for Computational Linguistics. ISBN 979-8-89176-195-7.
- Jiang, J.; Chen, P.; Wang, J.; He, D.; Wei, Z.; Hong, L.; Zong, L.; Wang, S.; Yu, Q.; Ma, Z.; et al. 2025b. Benchmarking large language models on multiple tasks in bioinformatics nlp with prompting. *arXiv preprint arXiv:2503.04013*.
- Jiang, J.; Wang, S.; Li, Q.; Kong, L.; and Wu, C. 2023a. A Cognitive Stimulation Dialogue System with Multi-source Knowledge Fusion for Elders with Cognitive Impairment. arXiv:2305.08200.
- Jiang, J.; Wang, S.; Li, Q.; Kong, L.; and Wu, C. 2023b. A Cognitive Stimulation Dialogue System with Multi-source Knowledge Fusion for Elders with Cognitive Impairment. *arXiv preprint arXiv:2305.08200*.
- Jiang, J.; Wang, Z.; Shan, Y.; Chai, H.; Li, J.; Ma, Z.; Zhang, X.; and Li, Y. 2025c. Biological sequence with language model prompting: A survey. *arXiv preprint arXiv:2503.04135*.
- Kim, J.; Shin, E.; Han, K.; Park, S.; Youn, J. H.; Jin, G.; and Lee, J.-Y. 2021. Efficacy of smart speaker-based metamemory training in older adults: Case-control cohort study. *Journal of medical Internet research*, 23(2): e20177.
- Lee, H.; Ahn, Y.; Lee, H.; Ha, S.; and Lee, S.-g. 2016. Quote recommendation in dialogue using deep neural network. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, 957–960.
- Li, J.; Galley, M.; Brockett, C.; Gao, J.; and Dolan, B. 2015. A diversity-promoting objective function for neural conversation models. *arXiv preprint arXiv:1510.03055*.
- Li, Q.; Chen, H.; Ren, Z.; Ren, P.; Tu, Z.; and Chen, Z. 2020. EmpDG: Multi-resolution Interactive Empathetic Dialogue Generation. In Scott, D.; Bel, N.; and Zong, C., eds., *Proceedings of the 28th International Conference on Computational Linguistics*, 4454–4466. Barcelona, Spain (Online): International Committee on Computational Linguistics.
- Li, Q.; Li, P.; Ren, Z.; Ren, P.; and Chen, Z. 2022. Knowledge bridging for empathetic dialogue generation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, 10993–11001.
- Lin, C.-Y. 2004. Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 74–81.
- Liu, A.; Feng, B.; Xue, B.; Wang, B.; Wu, B.; Lu, C.; Zhao, C.; Deng, C.; Zhang, C.; Ruan, C.; et al. 2024. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*.
- Liu, S.; Zheng, C.; Demasi, O.; Sabour, S.; Li, Y.; Yu, Z.; Jiang, Y.; and Huang, M. 2021. Towards Emotional Support Dialog Systems. In *ACL*.
- Loshchilov, I. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.

- Magyar, J.; Kobayashi, M.; Nishio, S.; Sinčák, P.; and Ishiguro, H. 2019. Autonomous robotic dialogue system with reinforcement learning for elderlies with dementia. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 3416–3421. IEEE.
- Mateos-Sanchez, M.; Melo, A. C.; Blanco, L. S.; and García, A. M. F. 2022. Chatbot, as Educational and Inclusive Tool for People with Intellectual Disabilities. *Sustainability*, 14(3).
- Morley, J. E.; and Cruz-Oliver, D. M. 2014. Cognitive stimulation therapy. *Journal of the American Medical Directors Association*, 15(10): 689–691.
- Ni, J.; Young, T.; Pandelea, V.; Xue, F.; and Cambria, E. 2023. Recent advances in deep learning based dialogue systems: A systematic survey. *Artificial intelligence review*, 56(4): 3055–3155.
- Pan, M.; Kitson, A.; Wan, H.; and Prpa, M. 2025. Ellma-t: an embodied llm-agent for supporting english language learning in social vr. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 576–594.
- Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W.-J. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, 311–318.
- Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8): 9.
- Sánchez Cuadrado, J.; Pérez-Soler, S.; Guerra, E.; and De Lara, J. 2024. Automating the Development of Task-oriented LLM-based Chatbots. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*, 1–10.
- Shao, Z.; Wang, P.; Zhu, Q.; Xu, R.; Song, J.; Bi, X.; Zhang, H.; Zhang, M.; Li, Y.; Wu, Y.; et al. 2024. Deepseekmath: Pushing the limits of mathematical reasoning in open language models. *arXiv preprint arXiv:2402.03300*.
- Spector, A.; Thorgrimsen, L.; Woods, B.; Royan, L.; Davies, S.; Butterworth, M.; and Orrell, M. 2003. Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial. *The British Journal of Psychiatry*, 183(3): 248–254.
- Sun, Y.; Zhang, X.; and Wang, Z. 2022. Comparative effectiveness of 3 settings of cognitive stimulation therapy on cognition and quality of life for people with dementia: A systematic review and network meta-analysis. *Journal of the American Medical Directors Association*, 23(3): 461–467.
- Tokunaga, S.; Tamura, K.; and Otake-Matsuura, M. 2021. A dialogue-based system with photo and storytelling for older adults: toward daily cognitive training. *Frontiers in Robotics and AI*, 8: 644964.
- Wang, Z.; Wang, Z.; Jiang, J.; Chen, P.; Shi, X.; and Li, Y. 2025. Large language models in bioinformatics: A survey. *arXiv preprint arXiv:2503.04490*.
- Waswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.; Kaiser, L.; and Polosukhin, I. 2017. Attention is all you need. In *NIPS*.
- Woods, B.; Thorgrimsen, L.; Spector, A.; Royan, L.; and Orrell, M. 2006. Improved quality of life and cognitive stimulation therapy in dementia. *Aging and Mental Health*, 10(3): 219–226.
- Yang, A.; Yu, B.; Li, C.; Liu, D.; Huang, F.; Huang, H.; Jiang, J.; Tu, J.; Zhang, J.; Zhou, J.; et al. 2025. Qwen2.5-1m technical report. *arXiv preprint arXiv:2501.15383*.
- Yi, Z.; Ouyang, J.; Liu, Y.; Liao, T.; Xu, Z.; and Shen, Y. 2024. A survey on recent advances in llm-based multi-turn dialogue systems. *arXiv preprint arXiv:2402.18013*.
- Zhang, A. Y.; Wong, G. H. Y.; Lum, T. Y. S.; Woods, B.; and Spector, A. 2025. Constructive engagement in cognitive stimulation therapy groups among people with dementia: A mixed-methods study. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 11(2): e70075.
- Zhang, T.; Kishore, V.; Wu, F.; Weinberger, K. Q.; and Artzi, Y. 2019a. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.
- Zhang, Z.; Han, X.; Liu, Z.; Jiang, X.; Sun, M.; and Liu, Q. 2019b. ERNIE: Enhanced language representation with informative entities. *arXiv preprint arXiv:1905.07129*.
- Zhou, H.; Ke, P.; Zhang, Z.; Gu, Y.; Zheng, Y.; Zheng, C.; Wang, Y.; Wu, C. H.; Sun, H.; Yang, X.; et al. 2021. Eva: An open-domain chinese dialogue system with large-scale generative pre-training. *arXiv preprint arXiv:2108.01547*.