

# Facilitating Early Maladaptive Schema–Guided Polite and Empathetic Psychotherapeutic Support: An LLM-Driven MoE-RL-Based Dialogue System

Anonymous submission

## Appendix

This section presents complementary materials, including dataset details, experiment details, and additional analysis to enhance the reader’s comprehension of the work.

### Dataset Details

#### Annotation Aspect Details

Table 1 provides the prompt template for dataset annotation. Table 2, Table 3, and Table 4 provide the definitions and examples of EMS labels, empathy strategies, and psychotherapeutic support strategies.

---

**Task Instruction:** Given a dialogue context, classify the target utterance into one of the following categories: <[list of EMS labels with definitions/list of emotions with definitions/list of politeness strategies with definitions/list of empathy strategies with definitions/list of psychotherapeutic support strategies with definitions]>. Provide a rationale for the predicted label.

**Few-shot Demonstrations:**

Dialogue Context:

Target Utterance:

Predicted Label:

**Target Input:**

Dialogue Context:

Target Utterance:

**LLM Output:**

Predicted Label: <label>

Explanation: <explanation>

---

Table 1: Prompt template used for dataset annotation.

## Experiment Details

### Classifiers Details

To formulate psychotherapeutic support, politeness, and empathy rewards, we design three specialized classifiers by fine-tuning the RoBERTa-large (Liu et al. 2019) model. The schema-adaptive support strategy classifier ( $f_{\text{psy}}(\cdot)$ ) receives as input the dialogue context along with the client’s early maladaptive schema (EMS) information and predicts the probable psychotherapeutic support strategy of the therapist. Similarly, the emotion-adaptive politeness strategy classifier ( $f_{\text{pol}}(\cdot)$ ) and the emotion-adaptive empathy strategy classifier ( $f_{\text{emp}}(\cdot)$ ) take the dialogue context enriched with the

client’s emotional state as input to generate predictions for the politeness and empathy strategies, respectively.  $f_{\text{psy}}(\cdot)$  achieves 85.4% accuracy and 80.2% macro-F1 score on HOPE, and 82.1% accuracy and 78.8% macro-F1 score on PSYCON.  $f_{\text{pol}}(\cdot)$  achieves an accuracy of 88.7% and a macro-F1 score of 85.5% on HOPE, and 87.6% accuracy with 83.3% macro-F1 score on PSYCON.  $f_{\text{emp}}(\cdot)$  attains 84.2% accuracy and 81.0% macro-F1 score on HOPE, and 81.9% accuracy with 79.2% macro-F1 score on PSYCON.

### Implementation Details

All implementations are conducted using the PyTorch framework<sup>1</sup>, and we utilize transformer-based architectures from the Hugging Face library throughout our experiments (Wolf et al. 2019). All the classifiers used in reward formulation are trained for 8 epochs. We adopt the medium variants of the pre-trained model, GPT-2 in ARDM, PersRFI, and GPT-Critic. All models are trained on an NVIDIA A100-PCIE-40GB GPU with CUDA version 11.2. For experiments involving OpenAI’s GPT, we use the GPT-4o-mini model accessed via the API. We empirically set the number of steps to  $M = 2$ . Reward weights are determined via grid search over the set 0.2, 1.0, with final values as follows:  $v_{\text{psy}} = v_{\text{pol}} = v_{\text{emp}} = 0.2$ , and  $v_{\text{cc}} = v_{\text{d}} = 1.0$ . We set the discount factor  $\beta$  to 0.9, the hyperparameter  $\lambda$  to  $1 \times 10^{-5}$ , and the scaling factor  $\rho$  to 1. We use a seed value of 42 for reproducibility and set the batch size to 4. We train the models using the Adam optimizer (Kingma and Ba 2014) with an initial learning rate of  $3 \times 10^{-6}$  and apply a linear warm-up over 120 steps. The warm-start phase runs for 5 epochs, followed by 3 epochs of joint training. For response generation, we use Top- $k$  sampling with  $k = 30$ , Top- $p$  sampling with  $p = 0.9$  (Holtzman et al. 2019), a temperature parameter  $\tau = 0.7$ , and a repetition penalty of 1.03.

### Baseline Details

We assess the performance of MATE by comparing it with 6 baseline models: (1) ARDM (Wu et al. 2021), which employs GPT-2 trained through an alternating update scheme for client and therapist roles; (2) PersRFI (Shi et al. 2021), which leverages RL to fine-tune GPT-2 with guidance from human demonstrations; (3) GPT-Critic (Jang, Lee, and Kim

<sup>1</sup><https://pytorch.org/>

EMS Label	Definition	Example
Abandonment / Instability (AB)	A deep-seated fear that close relationships are unreliable, leading to anxiety about being left alone or emotionally unsupported.	<i>I feel like I'm pushing people away because of my depression. I'm scared they'll leave me.</i>
Mistrust / Abuse (MA)	A belief that others will intentionally cause harm, betray, or exploit you, fostering chronic suspicion and fear in relationships.	<i>Sometimes I feel like people are just pretending to care, and they'll use it against me if I open up.</i>
Emotional Deprivation (ED)	The expectation that one's emotional needs, like empathy, affection, or support will consistently go unmet, leaving a void or sense of longing.	<i>Not great, to be honest. I'm feeling really depressed.</i>
Defectiveness / Shame (DS)	A perception of being fundamentally flawed, inferior, or unlovable, often accompanied by deep shame and fear of being exposed.	<i>My boss has been really critical of my work, and I just feel like I can't do anything right.</i>
Social Isolation / Alienation (SI)	A persistent feeling of being different, left out, or disconnected from others, resulting in loneliness and social withdrawal.	<i>I've been feeling really lonely lately. I don't have many friends, and I don't feel like anyone understands me.</i>
Dependence / Incompetence (DI)	A belief in one's inability to manage everyday tasks or decisions without help, leading to overreliance on others for guidance and support.	<i>I haven't really tried anything like that before.</i>
Vulnerability to Harm or Illness (VH)	A constant sense that disaster, like illness, accidents, or emotional breakdown is imminent, causing chronic anxiety and hypervigilance.	<i>It's been really hard to get out of bed. I feel like I'm going to completely fall apart.</i>
Enmeshment / Undeveloped Self (EM)	Difficulty distinguishing one's own identity and emotions due to excessive emotional involvement or closeness with significant others.	<i>Sometimes I don't even know what I want—I just try to keep others happy.</i>
Failure to Achieve (FA)	A belief that one is doomed to fail or lacks the ability to succeed, often resulting in low self-confidence and feelings of defeat.	<i>I just feel like I'm not good enough at my new job. I can't keep up.</i>
Entitlement / Grandiosity (ET)	A belief that one is superior to others and deserving of special treatment, often at the expense of empathy or fairness.	<i>Honestly, I don't see why I should have to follow all these rules when I know better.</i>
Insufficient Self-Control / Self-Discipline (IS)	Struggles with managing impulses, maintaining focus, or delaying gratification, frequently resulting in short-sighted or damaging decisions.	<i>Not really. I've been having trouble sleeping and I haven't been eating very healthily.</i>
Subjugation (SB)	A tendency to consistently yield to others' demands or opinions, suppressing one's own needs out of fear of conflict or rejection.	<i>I just go along with what others want, even when I don't agree. It's easier than starting a conflict.</i>
Self-Sacrifice (SS)	A pattern of prioritizing others' needs to the point of self-neglect, often driven by guilt or a strong desire to maintain harmony.	<i>Not really. I don't want to burden anyone.</i>
Approval-Seeking / Recognition-Seeking (AS)	An overdependence on external validation, where self-worth hinges on praise, admiration, or acceptance from others.	<i>I just want people to like me. I feel worthless if they don't.</i>
Negativity / Pessimism (NP)	A tendency to dwell on potential problems or worst-case scenarios, viewing life through a lens of doubt, criticism, and fear of disappointment.	<i>I just feel like everything is so overwhelming. I can't seem to find any joy in anything.</i>
Emotional Inhibition (EI)	Suppressing emotional expression to avoid disapproval, vulnerability, or perceived loss of control, often at the cost of authentic connection.	<i>That sounds really hard.</i>
Unrelenting Standards / Hypercriticalness (US)	A drive to meet overly rigid internal standards, leading to chronic dissatisfaction, perfectionism, and self-criticism.	<i>Yeah, I definitely have a tendency to be really hard on myself.</i>
Punitiveness (PU)	A belief that mistakes deserve harsh consequences, often resulting in being overly critical and unforgiving toward oneself or others.	<i>When I mess up, I can't forgive myself. I don't deserve a second chance.</i>

Table 2: Definition of EMS labels with example utterances.

2022), which enhances GPT-2 using a critic-driven approach that clones self-generated sentences during fine-tuning; (4) e-Therapist (Mishra et al. 2023), an RL-based GPT-2 model that incorporates task-relevant rewards to facilitate psychotherapeutic responses; (5) ProCoT (ChatGPT) (Deng et al. 2023), which prompts ChatGPT to generate chain-of-thought style reasoning for strategic planning of the next dialogue turn; and (6) Phi-3-FT (Touvron et al. 2023), a supervised fine-tuning of the Phi-3-mini-128k-instruct model using client’s maladaptive schema and emotion.

## Evaluation Metrics Details

**Automatic Evaluation Metrics.** Perplexity (PPL) (Brown et al. 1992) measures a model’s ability to predict responses. BLEU (Papineni et al. 2002) (B-3) assesses word overlap between the generated and reference responses. BERTScore-

F1 (BS-F1) (Zhang et al. 2019)<sup>2</sup> evaluates semantic similarity between generated and reference responses in latent space using contextual embeddings from BERT (Devlin et al. 2019). DISTINCT-3 (D-3) (Li et al. 2015) quantifies response diversity, while LEN captures the length of generated responses.

Schema-Support Strategy Consistency (SSSC) assesses whether the psychotherapeutic support strategy expressed in the generated response matches that of the ground-truth response and is appropriate to the client’s EMS. Emotion-Politeness Strategy Consistency (EPSC) measures the alignment of the politeness strategy between the generated and ground-truth responses, whilst ensuring that it is also consistent with the emotional state of the client. Emotion-Empathy

<sup>2</sup>BERTScore: <https://huggingface.co/spaces/evaluate-metric/bertscore>

Empathy Strategy	Definition	Example
Active Listening	Attending fully to the client’s verbal and nonverbal communication through eye contact, nodding, and reflective responses to ensure the client feels heard, understood, and validated.	<i>I’m sorry to hear that. Can you tell me more about what’s been going on?</i>
Unconditional Positive Regard	Demonstrating nonjudgmental acceptance and support to create a safe, respectful space where clients feel valued and open to sharing.	<i>That’s great. Remember, it’s important to take care of yourself and your mental health. Let’s schedule another session for next week so we can check in and see how you’re doing.</i>
Validation	Acknowledging and affirming the client’s emotional experience through statements that communicate understanding, helping clients feel seen and accepted.	<i>Those are all understandable concerns. Let’s work on some strategies for managing your anxiety. Have you tried any grounding techniques?</i>
Appropriate Self-Disclosure	Carefully sharing relevant personal experiences when beneficial to normalize the client’s feelings and foster connection, while keeping the focus on the client.	<i>I’ve found that even a few minutes of mindfulness can sometimes help me feel more grounded. Would you like to try one together?</i>
Collaborative Approach	Actively involving clients in setting goals and making decisions to promote empowerment, engagement, and shared ownership of the process.	<i>Let’s work on some strategies for managing your anxiety. Have you tried any grounding techniques?</i>
Managing Silence	Allowing space for reflection by respecting silence during sessions, encouraging deeper thought and emotional processing without rushing the client.	<i>Excellent. How do you feel now?</i>
Asking Empowering Questions	Using open-ended, non-directive questions to encourage client reflection, autonomy, and problem-solving from their own perspective.	<i>Is there anything else you want to talk about today?</i>
Holding Space for Emotions	Creating an environment where clients can safely express difficult emotions without interruption, offering presence rather than solutions.	<i>Okay, let’s talk about that. What’s been causing your anxiety?</i>
Perspective Taking	Considering the client’s emotional and personal context to better understand their viewpoint and respond with appropriately attuned and context-sensitive support.	<i>I understand. It’s normal to feel sad from time to time, but if it’s persisting for a long period of time, it’s important to address it.</i>
No Strategy	Designating utterances that do not employ any specific empathy strategy.	<i>Let’s schedule another session for next week so we can check in and see how you’re doing.</i>

Table 3: Definition of empathy strategies with example utterances.

Strategy Consistency (EESC) evaluates the consistency of the empathy strategy in the generated response with that in the ground-truth and is in accordance with the client’s emotional needs. Engagingness (ENG) evaluates the contextual consistency and diversity of the generated responses.

$$\text{SSSC} = \mathbb{E}_{I_i, r_i} \mathbf{1} \{f_{\text{psy}}(r_i) = f_{\text{psy}}(\hat{r}_i)\}, \quad (13)$$

$$\text{EPSC} = \mathbb{E}_{I_i, r_i} \mathbf{1} \{f_{\text{pol}}(r_i) = f_{\text{pol}}(\hat{r}_i)\}, \quad (14)$$

$$\text{EESC} = \mathbb{E}_{I_i, r_i} \mathbf{1} \{f_{\text{emp}}(r_i) = f_{\text{emp}}(\hat{r}_i)\}. \quad (15)$$

$$\text{ENG} = \frac{1}{\mathbb{K}} \sum_{\mathbb{K}} \frac{\text{MS}(\hat{r}_i, \hat{r}_{i-1}) + \text{EAD}(\hat{r}_i)}{2} \quad (16)$$

$$\text{LEN} = \sum_{\mathbb{K}} \frac{|\hat{r}|}{\mathbb{K}} \quad (1)$$

Here,  $\mathbb{K}$  denotes the total number of generated responses.  $|\hat{r}|$  denotes the length of the generated responses.

**Human Evaluation Metrics.** Fluency (FLU) measures grammatical correctness, while Contextual Coherence (COH) evaluates the alignment between generated responses and the dialogue context. Naturalness (NAT) assesses how natural and interactive the dialogue is in sustaining client’s interest. Schema-Support Strategy Consistency (SSSC) examines whether the psychotherapeutic support strategy in the generated response aligns with the ground-truth support strategy and is appropriate to the client’s Early Maladaptive Schema (EMS). Emotion-Politeness Strategy Consistency (EPSC) and Emotion-Empathy Strategy Consistency (EESC) evaluate the consistency between the generated and ground-truth responses in terms of emotion-adaptive politeness and empathy strategies, respectively.

**Human Evaluation Process.** Human evaluation is conducted by three independent evaluators<sup>3</sup>, including two Ph.D holders in Linguistics and one postgraduate in Computer Science, all with prior experience in similar evaluations. Evaluators are briefed on the psychotherapeutic interactions and their associated aspects, including early maladaptive schemas, psychotherapeutic support, politeness, and empathy strategies, before engaging in multi-turn interactions with the system. Initially, each evaluator interacts 5 times with the system using a different set of responses, resulting in a total of 15 human-evaluated dialogues for HOPE dataset. Similarly, we obtain 15 dialogues for the PSYCON dataset. These 30 dialogues are provided to the domain specialist from the government-run institution for cross-verification in terms of evaluation quality. After the domain specialist’s approval, each evaluator is again instructed to interact 20 times with the system using distinct response sets, resulting in an additional 60 human-evaluated dialogues for the HOPE dataset. Likewise, an equivalent set of 60 more dialogues is obtained for the PSYCON dataset. The evaluators are then instructed to rate each dialogue on SSSC, EPSC, EESC, FLU, COH, and NAT using a 5-point Likert scale (1 = low, 5 = high).

## Additional Analysis

### Ablation of Experts and Rewards

We evaluate the impact of each expert in the proposed MATE system by performing a sequence of ablation studies, where one expert is excluded at a time. The results of these ablations are reported in Table 5. The findings clearly demon-

<sup>3</sup>Evaluators are the same involved in dataset annotation and are compensated per institutional norms.

Psychotherapeutic Support Strategy	Definition	Example
Open Questions	Encourage expansive, reflective responses by allowing the individual to elaborate on their thoughts and feelings. They typically begin with words like "what," "how," or "why," fostering a more natural and less defensive conversation compared to closed questions that limit responses to "yes/no" answers.	<i>I'm sorry to hear that. Would you like to talk about what's been causing you to feel this way?</i>
Reflections	Demonstrate understanding and validate the client's thoughts and feelings by restating or paraphrasing what they have shared to reinforce their experience without judgment, fostering self-discovery, and clarifying their thoughts. It ensures that the dialogue remains client-driven, avoiding premature advice or biased questions that could hinder genuine engagement.	<i>That sounds really tough. It takes a lot of strength to ask for help, and I'm glad you're here. Have you thought about any specific goals you'd like to work on?</i>
Affirmations	Acknowledge and reinforce positive attributes, efforts, or progress made by the client, even if those efforts were not entirely successful. These statements help build rapport and foster positive behaviors.	<i>You're welcome. I'm here to help you, and I believe in you. Let's work together to help you feel better.</i>
Summarizing	Used at various points to recap key discussions, highlight action plans, and reinforce goals, ensuring that the supporter accurately tracks the client's narrative and intentions, and prompts reflection on decisions.	<i>That's great. Remember that there's no one right way to manage your depression, and it may take some trial and error to find what works for you.</i>
Pros and Cons Chart	Asking the client to list the advantages and disadvantages of both continuing a current behavior and making a change to help them explore ambivalence and gain insights into their behavior patterns.	<i>What do you think are some of the pros and cons of trying mindfulness versus continuing as things are now?</i>
Learning Assessment	Assessing the client's existing knowledge, learning needs, and preferred methods to tailor information to their individual learning styles and priorities, helping them make informed decisions about their health and care.	<i>Have you ever tried mindfulness or relaxation techniques?</i>
Elicit Change Talk	Encourages the client to express reasons for change, uncovering their motivations and potential solutions. It's about getting the client to talk about what might encourage them to make a change to stimulate self-reflection and the generation of ideas for overcoming obstacles.	<i>Can you think of any small steps you can take to make work feel more manageable?</i>
Readiness Rulers	Assess the client's perceived importance and confidence in making a change. The supporter asks the client to rate, on a scale from 0 to 10, how important the change is to them and how confident they are in their ability to make the change.	<i>On a scale from 0 to 10, how motivated do you feel to try some self-care strategies like going for a walk or practicing mindfulness?</i>
Values and Strengths	Identifies and uses the client's personal values and strengths to motivate change and build confidence, encouraging self-identification of resources to achieve their goals.	<i>That sounds really tough. It takes a lot of strength to ask for help, and I'm glad you're here. Have you thought about any specific goals you'd like to work on?</i>
Elicit Provide Elicit	Eliciting the client's existing knowledge and information needs, providing relevant information based on their questions, and then eliciting their response to ensure understanding and promote further engagement in the decision-making process.	<i>Good. Let's also explore some cognitive-behavioral techniques to help challenge negative thoughts and beliefs that may be contributing to your depression and anxiety. Have you heard of cognitive restructuring?</i>
No Strategy	Designating utterances that do not employ any specific psychotherapeutic support strategy.	<i>Hi, how are you feeling today?</i>

Table 4: Definition of psychotherapeutic support strategies with example utterances.

Models	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑
	HOPE									PSYCON								
MATE	13.47	8.95	38.12	0.772	0.726	0.796	0.683	0.527	23.41	1.09	8.31	46.07	0.901	0.819	0.843	0.853	0.698	32.85
- Psychotherapeutic Expert	14.25	8.10	34.56	0.740	0.675	0.740	0.620	0.480	22.17	1.32	7.90	43.50	0.870	0.770	0.785	0.790	0.650	31.22
- Politeness Expert	14.00	8.35	35.70	0.752	0.690	0.720	0.605	0.460	22.55	1.28	8.00	44.10	0.875	0.765	0.770	0.785	0.640	31.65
- Empathy Expert	13.85	8.50	36.00	0.758	0.700	0.780	0.630	0.500	22.95	1.25	8.10	44.80	0.880	0.780	0.800	0.810	0.670	32.00
- Support Strategy Reward	14.10	8.50	36.20	0.745	0.695	0.770	0.650	0.495	22.70	1.40	7.95	44.50	0.870	0.780	0.800	0.810	0.660	31.90
- Politeness Reward	13.95	8.60	36.75	0.755	0.705	0.780	0.660	0.510	22.90	1.38	8.00	44.75	0.875	0.790	0.810	0.820	0.670	32.10
- Empathy Reward	13.85	8.70	37.00	0.760	0.710	0.785	0.670	0.520	23.05	1.35	8.05	45.00	0.880	0.800	0.820	0.830	0.680	32.35
- Contextual Consistency Reward	13.90	8.55	36.40	0.750	0.700	0.775	0.655	0.505	22.85	1.37	7.90	44.25	0.872	0.785	0.805	0.815	0.665	31.95
- Diversity Reward	14.05	8.45	36.10	0.740	0.690	0.760	0.640	0.490	22.60	1.42	7.85	44.00	0.865	0.775	0.795	0.800	0.650	31.80

Table 5: Ablation of experts and rewards in MATE. - indicates the omission of the expert/reward.

strate that the elimination of any expert results in a significant decrease in performance. In particular, excluding the psychotherapeutic expert causes a marked drop in SSSC scores by 7.6% and 6.4% on HOPE and PSYCON datasets, respectively. This decline in performance likely arises because the lack of a psychotherapeutic expert impairs the model's capacity to grasp the subtle psychotherapeutic support semantics crucial for successful psychotherapeutic dialogue. Likewise, the removal of politeness and empathy experts also leads to significant performance decline. Politeness and

empathy experts play a critical role in enhancing the politeness and empathy of responses, substantially influencing the EPSC and EESC scores, respectively. Further, we examine the effect of rewards in the proposed MATE framework by systematically excluding each reward component one at a time. The results of ablation studies are reported in Table 5. The exclusion of any reward negatively impacts several key facets of dialogue quality, such as psychotherapeutic support, politeness, empathy, consistency, and diversity, as reflected by the reduced scores of SSSC, EPSC, EESC, and

ENG across both datasets. These observations highlight the crucial role of each expert and reward in generating polite and empathetic responses attuned to the client’s EMS and emotional state within psychotherapeutic support dialogues.

### Ablation of Multi-task Learning of MoE

To further understand the contribution of multi-task learning in the proposed **MATE** framework, we conduct an ablation study by selectively disabling combinations of loss components. Specifically, we evaluate the performance when removing sets of loss terms among  $\mathcal{L}^{PE}$ ,  $\mathcal{L}^{PoE}$ ,  $\mathcal{L}^{EE}$ , and  $\mathcal{L}^{MSE}$ . The results are reported in Table 6. In the first ablation, we omit  $\mathcal{L}^{PoE}$ ,  $\mathcal{L}^{EE}$ , and  $\mathcal{L}^{MSE}$ , retaining only the  $\mathcal{L}^{PE}$ . Similarly, in the second and third settings, we preserve just one of the other experts. Finally, we evaluate the extreme case where all expert-specific and supervision losses are removed, effectively disabling the mixture-of-experts mechanism. Across both HOPE and PSYCON datasets, we observe consistent performance degradation under all ablated configurations. The removal of any combination of expert objectives leads to increased perplexity and declines across all other evaluation metrics. Notably, the full exclusion of all losses results in the most pronounced drop, confirming the critical role of the multi-task mixture-of-experts supervision. These findings highlight the mutual reliance among the various expert modules and illustrate that multi-task learning imparts them with complementary strengths, which together lead to a marked enhancement in overall model performance.

### Ablation of Expert Policy and Training Paradigm

We investigate the impact of two core components in **MATE** training framework - the expert policy identification network and the integration of warm start and joint training strategy. Specifically, we compare the proposed **MATE** framework against three ablated configurations - (i) replacing the expert identification policy network with random sampling (Random Expert Policy), (ii) using a warm start without joint training (*w* Warm Start Only), and (iii) applying joint training without a warm start (*w* Joint Training Only). The results are presented in Table 7. In all the settings, we observe a drop in performance across both datasets. The random selection of experts introduces uncertainty in the decision-making process, thereby adversely affecting overall performance, particularly with respect to SSSC, EPSC, and EESC scores. This highlights the necessity of informed expert identification for effective response generation. The ‘*w* Warm Start Only’ approach results in slightly inferior performance, while the ‘*w* Joint Training Only’ approach, which prioritizes reward maximization, achieves this objective but at the expense of overall dialogue quality. These observations highlight the necessity of a balanced approach to combining warm-start initialization with joint training, guided by a learned expert policy to maintain high-quality, polite, and empathetic psychotherapeutic dialogues grounded in clients’ EMS and emotions.

### Effect of Iteration Steps

We investigate the impact of varying the number of iteration steps  $M$  in **MATE**. As shown in Table 8, increasing  $M$

from 1 to 2 consistently improves performance across all metrics on both HOPE and PSYCON datasets, indicating that additional refinement iterations enhance the integration of expert knowledge. However, further increasing  $M$  to 3 leads to a slight degradation in several metrics, including B-3 and BS-F1, suggesting diminishing returns and potential overfitting to expert features. Notably,  $M = 2$  achieves the best balance between goal accomplishment and language quality, confirming the effectiveness of moderate iterative training steps in psychotherapeutic dialogue generation.

### Out-of-Domain Evaluation

To assess the domain generalization capability of **MATE**, we conduct an evaluation on the EPE-enEIH dataset (Mishra, Priya, and Ekbal 2023), which contains mental health and legal counseling dialogues for crime victims, substantially distinct from the HOPE and PSYCON datasets used for training **MATE**. As shown in Table 9, **MATE** demonstrates strong performance across multiple evaluation dimensions without any domain-specific fine-tuning. The model attains a low PPL score of 15.41, indicating fluent response generation in an unseen domain. Despite the domain shift, it maintains high lexical diversity (D-3 = 27.31) and n-gram precision (B-3 = 7.09), along with strong semantic alignment (BS-F1 = 0.743). Moreover, the scores for SSSC = 0.698, EPSC = 0.703, and EESC = 0.617 highlight the model’s ability to produce polite and empathetic psychotherapeutic responses that are consistent with the client’s emotional state and EMS context. The scores for ENG = 0.503 and LEN = 20.69 further affirm the model’s ability to produce contextually appropriate and informative responses. These findings substantiate the generalization strength of the proposed **MATE** framework, demonstrating its effectiveness in extending psychotherapeutic dialogue modeling to out-of-domain scenarios.

### Case Study

To qualitatively evaluate the efficacy of the proposed dialogue system **MATE** in generating polite, empathetic, and schema-guided responses, we present comparative case studies against the baseline model Phi-3-FT. Tables 10 and 11 present examples of client interactions with the proposed PS dialogue system - **MATE**, alongside the second-best baseline model, Phi-3-FT. From these examples, it is evident that clients consistently rated the **MATE** higher across all human evaluation metrics, specifically SSSC, EPSC, EESC, FLU, COH, and NAT, thereby demonstrating **MATE**’s superior performance in addressing clients’ emotional states and early maladaptive schemas during psychotherapeutic conversations. The **MATE** outperforms the Phi-3-FT model by effectively integrating politeness and empathy attuned to clients’ emotions and maladaptive schemas during psychotherapy, resulting in a more engaging, client-centric, and amicable psychotherapeutic experience. As shown in Table 10, the client expresses feelings of deep hopelessness, exhaustion, and grief following the recent loss of a spouse. The responses generated by **MATE** showcase marked politeness and empathy, employing warm, validating language

Models	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑
	HOPE									PSYCON								
MATE	<b>13.47</b>	<b>8.95</b>	<b>38.12</b>	<b>0.772</b>	<b>0.726</b>	<b>0.796</b>	<b>0.683</b>	<b>0.527</b>	<b>23.41</b>	<b>1.09</b>	<b>8.31</b>	<b>46.07</b>	<b>0.901</b>	<b>0.819</b>	<b>0.843</b>	<b>0.853</b>	<b>0.698</b>	<b>32.85</b>
- ( $\mathcal{L}^{\text{PoE}} + \mathcal{L}^{\text{EE}} + \mathcal{L}^{\text{MSE}}$ )	15.86	8.48	36.91	0.752	0.711	0.772	0.673	0.514	22.68	2.39	7.87	44.28	0.861	0.803	0.813	0.844	0.683	31.81
- ( $\mathcal{L}^{\text{PE}} + \mathcal{L}^{\text{EE}} + \mathcal{L}^{\text{MSE}}$ )	16.01	8.43	36.58	0.751	0.703	0.764	0.666	0.513	22.61	2.32	7.79	44.03	0.853	0.796	0.818	0.837	0.684	31.61
- ( $\mathcal{L}^{\text{PE}} + \mathcal{L}^{\text{PoE}} + \mathcal{L}^{\text{MSE}}$ )	15.90	8.47	36.81	0.754	0.713	0.771	0.675	0.524	22.66	1.95	8.03	44.46	0.868	0.802	0.817	0.843	0.693	31.76
- ( $\mathcal{L}^{\text{PE}} + \mathcal{L}^{\text{PoE}} + \mathcal{L}^{\text{EE}} + \mathcal{L}^{\text{MSE}}$ )	17.01	8.16	35.86	0.743	0.683	0.756	0.657	0.507	22.11	2.60	7.56	43.21	0.842	0.783	0.798	0.827	0.675	30.91

Table 6: Ablation of multi-task learning of mixture-of-experts in MATE.

Models	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑
	HOPE									PSYCON								
MATE	<b>13.47</b>	<b>8.95</b>	<b>38.12</b>	<b>0.772</b>	<b>0.726</b>	<b>0.796</b>	<b>0.683</b>	<b>0.527</b>	<b>23.41</b>	<b>1.09</b>	<b>8.31</b>	<b>46.07</b>	<b>0.901</b>	<b>0.819</b>	<b>0.843</b>	<b>0.853</b>	<b>0.698</b>	<b>32.85</b>
w Random Expert Policy	17.01	8.14	35.87	0.738	0.682	0.752	0.654	0.508	22.11	2.60	7.56	43.21	0.842	0.781	0.790	0.822	0.672	30.90
w Warm Start Only	15.89	8.48	36.90	0.751	0.710	0.770	0.668	0.519	22.67	2.38	7.87	44.27	0.861	0.802	0.812	0.844	0.682	31.80
w Joint Training Only	14.60	8.73	37.66	0.764	0.722	0.791	0.678	0.525	23.02	1.60	8.13	45.72	0.889	0.814	0.837	0.850	0.694	32.44

Table 7: Ablation of expert policy identification network and training paradigm in MATE.

Models	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑
	HOPE									PSYCON								
w $M = 1$	14.23	8.64	36.83	0.759	0.703	0.778	0.661	0.509	22.86	1.18	8.01	44.28	0.882	0.800	0.828	0.838	0.684	31.71
w $M = 2$	13.47	8.95	38.12	0.772	0.726	0.796	0.683	0.527	23.41	1.09	8.31	46.07	0.901	0.819	0.843	0.853	0.698	32.85
w $M = 3$	14.56	8.31	37.24	0.764	0.715	0.789	0.673	0.522	23.03	1.27	7.79	44.90	0.893	0.809	0.832	0.846	0.691	32.04

Table 8: Effect of the number of iteration steps in MATE.

en-EIH									
Model	PPL ↓	B-3 ↑	D-3 ↑	BS-F1 ↑	SSSC ↑	EPSC ↑	EESC ↑	ENG ↑	LEN ↑
MATE	15.41	7.09	27.31	0.743	0.698	0.703	0.617	0.503	20.69

Table 9: Performance of MATE on the out-of-domain en-EIH dataset.

throughout the interaction. For instance, for the client’s utterance expressing emotion deprivation and a sense of deep grief, “*Well, my husband passed away a few months ago and it’s been really hard to deal with. We were married for over 50 years and now I feel lost without him.*”, the proposed MATE system generates “*I’m truly sorry. That’s a profound loss. Grieving takes time, but you don’t have to go through it alone. Maybe we can explore small steps like reconnecting with gardening or trying a support group?*”. This response offers a schema-informed and emotionally attuned response that acknowledges the profound loss, validates the emotional pain, and uses a polite (*‘I’m truly sorry’*) and empathetic (*‘That’s a profound loss’*, *‘Grieving takes time’*) tone to foster a sense of psychological safety, respectful and cordial connection (*‘Maybe we can explore small steps’*), thereby strengthening the psychotherapeutic bond and enhancing the overall outcome during psychotherapy. In contrast, though the responses generated by Phi-3-FT are polite and empathetic, they often prematurely shift to solution-focused advice and bypass the emotional depth of the client’s experience.

Though the proposed MATE system demonstrates consistently polite, empathetic, and schema-guided, emotionally supportive behavior, creating a safe and compassionate therapeutic space, it exhibits certain limitations when dealing with brief or ambiguous client utterances. While the system effectively mirrors the client’s emotional state and en-

courages gentle progression, it occasionally overlooks opportunities for deeper clarification or fails to probe vague responses that may warrant further exploration. For example, in Table 10, when the client responds with *‘I don’t know, I don’t really feel up to it’* after being asked to consider reconnecting with others, MATE responds supportively - *‘That’s okay. Your energy is precious right now, and honoring your limits is important. Maybe we can think of something small...’*. This response is undeniably polite, empathetic, and validating. However, it assumes a direction (e.g., suggesting a support group) without first understanding what *‘I don’t really feel up to it’* truly means, whether it stems from fear or sadness or any underlying maladaptive schema. A more nuanced follow-up, such as *‘Would you like to share what feels most difficult about that right now?’*, may evoke the client to express hidden barriers and enhance emotional insight. This improvement could further align the responses with core therapeutic values of curiosity, attunement, and collaborative understanding, and create a comfortable space for clients.

## Limitations

While MATE demonstrates strong capabilities in generating polite and empathetic psychotherapeutic responses aligned with clients’ emotions and maladaptive schemas, it has several limitations. First, the MoE-RL framework used in MATE is resource-intensive, requiring high GPU memory ( $\geq 40$  GB) for training, which may limit accessibility for broader adoption. Furthermore, RL introduces inherent instability during training. Although the reward-guided policy learning framework allows MATE to integrate empathy, politeness, and support strategies beyond what is explicitly observed in training data, this adaptability can also lead to convergence issues. Stabilizing the learning process of-

<b>Speaker</b>	<b>Utterances generated by proposed model - MATE (SSSC = 5, EPSC = 5, EESC = 5, FLU = 5, COH = 5, NAT = 5)</b>
Therapist	Hello, it's really nice to see you again. How are you feeling today, emotionally and physically?
Client	Not too good, I feel very down and hopeless.
Therapist	I hear you, and I'm really sorry you're feeling this way. That sense of hopelessness can feel incredibly heavy. Would it be okay if we explore a bit more about what's been weighing on you lately?
Client	I just feel like everything is hopeless, like there's no point in trying anymore. I'm tired all the time and I can't concentrate on anything.
Therapist	Thank you for sharing that with me. It sounds like you're carrying a lot, and it makes sense you'd feel exhausted. When everything feels pointless, it's often a sign that your emotional needs might not be fully met. How long have you been experiencing this emotional weight?
Client	For a few months now, it's been getting worse and worse.
Therapist	That must feel really overwhelming. You've been holding this alone for quite some time. Have there been any moments where something helped, even just a little?
Client	I've tried talking to friends and family, but it doesn't really seem to help.
Therapist	That's completely understandable. Sometimes even when others are around, we can still feel profoundly alone. It sounds like you've been needing deeper emotional support. Would you be open to exploring support options together, such as therapy or other forms of care?
Client	I haven't really thought about it, but I guess it's worth a try.
Therapist	That's a courageous and hopeful step, and I'm here with you on this journey. If it's alright with you, could we talk about any particular life changes or losses that may have deepened these feelings?
Client	Well, my husband passed away a few months ago and it's been really hard to deal with. We were married for over 50 years and now I feel lost without him.
Therapist	I'm truly sorry for your loss. Losing a life partner after so many years is like losing a part of yourself, it's no wonder you feel lost. That deep emotional bond and your grief deserve space and compassion. Have you felt able to express this grief with anyone?
Client	No, not really. I've been isolating myself a lot lately.
Therapist	That makes sense. Grief can make us withdraw, especially if it feels like others can't truly understand our pain. But you don't have to go through this alone. Would you be open to us thinking of gentle ways to help you reconnect with others, at your own pace?
Client	I don't know, I don't really feel up to it.
Therapist	That's okay. Your energy is precious right now, and honoring your limits is important. Maybe we can think of something small—like attending a support group where others may be going through similar losses. It's a space where your grief won't feel out of place.
Client	I guess that could be helpful.
Therapist	I'm glad to hear that. Let's look at some options together. And in the meantime, I wonder was there something you once enjoyed doing, something that gave you a sense of calm or purpose?
Client	I used to love gardening, but I haven't had the energy to keep up with it.
Therapist	Gardening is such a beautiful expression of care and life. Maybe we can gently reintroduce it, without pressure. Just a few minutes a day outside with your plants. You're not doing it for productivity, you're doing it for yourself, as an act of healing.
Client	That sounds doable.
Therapist	Wonderful. Small steps like these can gradually rebuild your sense of self. Also, sometimes grounding techniques like deep breathing can help during moments of emotional overwhelm. Would you feel comfortable trying one now together?
Client	Sure, okay.
Therapist	Alright. Let's begin: take a deep breath in through your nose... hold for four seconds... and now gently exhale through your mouth. Let's repeat that a few more times together.
Client	Okay.
Therapist	How are you feeling now?
Client	A little more relaxed, actually.
Therapist	I'm really glad to hear that. Your body and mind are responding to care, and that's a hopeful sign. You're not alone in this—I'll be here with you each step of the way.
<b>Speaker</b>	<b>Utterances generated by Phi-3-FT (SSSC = 3, EPSC = 4, EESC = 3, FLU = 5, COH = 4, NAT = 4)</b>
Therapist	Hello, how are you feeling today?
Client	Not great. I feel very down and hopeless, like everything is pointless. I'm tired and can't concentrate.
Therapist	That sounds really difficult. How long have you been feeling this way?
Client	A few months. It's been getting worse.
Therapist	I'm sorry you're going through this. Have you tried anything to feel better?
Client	I've talked to family and friends, but it didn't help.
Therapist	I appreciate you reaching out. Have you considered therapy or medication?
Client	Not seriously, but maybe it's worth trying.
Therapist	That's a hopeful step. Was there something significant that happened recently?
Client	Yes, my husband passed away. We were together over 50 years. I feel lost and alone.
Therapist	I'm truly sorry. That's a profound loss. Grieving takes time, but you don't have to go through it alone. Maybe we can explore small steps—like reconnecting with gardening or trying a support group?

Table 10: Interaction with the proposed system - MATE and the baseline Phi-3-FT on PSYCON.

ten necessitates incorporating external knowledge sources, such as schema therapy principles or emotion-aware heuristics. Another key challenge involves the calibration of reward weights, which is both computationally intensive and

time-consuming. We address this by employing grid search to identify optimal combinations, but the reward formulation could still be refined further. In particular, integrating high-quality, human-curated feedback for training the re-

<b>Speaker</b>	<b>Utterances generated by proposed model - MATE (SSSC = 5, EPSC = 5, EESC = 4, FLU = 5, COH = 5, NAT = 5)</b>
Therapist	Hi, it sounds like there's something really weighing on you today. I want to acknowledge how hard it must be to bring this in.
Client	Yeah, I just... I don't know. It sucks. This therapy—it just feels like it's not working. And it's uncomfortable to even say that. Like I'm stuck.
Therapist	I really appreciate your honesty. It takes a lot of courage to say something like that. You're trying so hard, and when it feels like it's not helping, that can feel not just frustrating—but lonely, too.
Client	Exactly. Like, what else am I supposed to do? I told you I'm not okay. I don't know what more I can say.
Therapist	It sounds like you're doing everything you can—and yet still feeling unheard or unsupported. That's incredibly painful. I'm wondering if some of that pain feels directed toward me?
Client	I guess. Yeah. I think I'm angry that I'm still feeling stuck even though I'm trying. Maybe some of that anger is toward you... like I wanted this to help more than it has.
Therapist	That makes so much sense. You came here hoping for change and comfort—and it feels like I've fallen short. I really want to hold space for that anger. You deserve to have those feelings heard without judgment.
Client	It's hard to even say that. I don't usually tell people I'm mad. It makes me anxious. I feel like I'm doing something wrong.
Therapist	That's such an important insight. It sounds like expressing anger feels risky for you, like it might lead to rejection or being misunderstood. I wonder if this is something you've had to carry for a long time—keeping your anger hidden to protect yourself?
Client	Yeah, kind of. Like if I say how I feel, I'll be blamed or seen as the problem. That's what it feels like now, too.
Therapist	That really resonates. It's as if you're protecting yourself from a deeper fear—that if you're honest about your hurt or disappointment, you'll be left alone again. And I want you to know: I'm here with you in this. Your anger makes sense, and it matters.
Client	Thank you... that actually feels different. Like you're hearing me for real this time.
Therapist	I'm so glad. That's what I want most—to be with you in the hard stuff. Let's keep exploring this together, and figure out how therapy can start feeling more supportive.
Client	Okay... that feels a little better. Just saying it out loud and not being shut down. That helps.
Therapist	You're not alone in this anymore. We'll take it one step at a time, together.
<b>Speaker</b>	<b>Utterances generated by Phi-3-FT (SSSC = 2, EPSC = 3, EESC = 3, FLU = 5, COH = 3, NAT = 3)</b>
Therapist	Sounds like you're having a hard time today. Want to say more?
Client	Yeah. It just feels like therapy isn't working. It's frustrating. I don't know what else to say.
Therapist	You've mentioned that before. Can you try to be a bit more specific about what you think isn't helping?
Client	I don't know. I'm saying I'm not okay and I feel like nothing's changing.
Therapist	Are you angry at me?
Client	I'm angry that therapy isn't working. I've tried, and I'm still stuck. Maybe I am a little angry at you too.
Therapist	That's okay. It's okay to feel angry. How does it feel to say that out loud?
Client	It's hard. I feel anxious even saying it. Like I don't want to blame you.
Therapist	So you're feeling anxious about expressing anger, because you don't want it to seem like you're blaming me?
Client	Yeah, I guess. It's just hard to say these things. I feel like I'm alone in this.
Therapist	It sounds like you're feeling alone and maybe disappointed. That's important to acknowledge.
Client	Yeah. It's hard, but maybe saying it helps.
Therapist	Okay. Let's try to keep that openness going. Naming these feelings is a good first step.

Table 11: Interaction with the proposed system - MATE and the baseline Phi-3-FT on HOPE.

ward function could significantly enhance policy robustness. Moreover, increasing the size of the reward model parameters can lead to more generalizable behavior, though it also introduces additional computational cost and complexity.

Another limitation arises from the dependency on the annotation quality generated using LLMs in a few-shot setup. While effective, these annotations may occasionally reflect LLM biases or overlook nuanced social cues, impacting the downstream response generation. Although we incorporate schema and emotion alignment, MATE may still produce responses that lack the intended level of empathy or politeness, especially in cases of ambiguous or short client utterances like ‘*Yes, sure.*’, ‘*Ofcourse, I can!*’, etc. Lastly, the current implementation and evaluation are confined to English-language therapeutic dialogues. Broader applicability across languages and cultural settings remains an open area for exploration, requiring cross-lingual adaptation and culturally-sensitive response modeling.

## Ethical Statement

Given the sensitive nature of the proposed psychotherapeutic dialogue system -MATE, it is imperative that its develop-

ment and deployment adhere to rigorous ethical standards. This study has been reviewed and approved by the Institutional Review Board (IRB). MATE is designed to foster respectful, empathetic, and supportive interactions during psychotherapy. Rather than manipulating users or enforcing behavioral conformity, MATE aims to uphold users’ autonomy, emotional well-being, and dignity throughout the conversation. Users retain full control over the dialogue, with no coercive elements embedded in the system’s design or objectives. Ethical diligence also extends to data practices: all training and evaluation data are sourced from publicly available datasets, with no inclusion of personally identifiable or sensitive user information. Human participants (annotators and evaluators) were fully informed about the nature of their tasks, and appropriate consent and privacy safeguards were ensured. MATE is intended solely for academic research and educational purposes, with access to models and datasets granted under non-commercial terms. Deployment in high-stakes settings is discouraged without further ethical review and human oversight.

To ensure responsible dataset use, access will be granted only upon execution of a formal agreement stipulating that



the data be employed exclusively for academic research purposes. The annotation, curation, and manual evaluation processes were conducted by qualified human experts who are regular members of our research team and were appropriately compensated in accordance with institutional policies. Also, this study utilizes the publicly available HOPE and PSYCON datasets. Necessary permissions were obtained for their use, and all data handling strictly complies with the copyright terms and conditions set forth by the rights holders.

## References

- Brown, P. F.; Della Pietra, S. A.; Della Pietra, V. J.; Lai, J. C.; and Mercer, R. L. 1992. An estimate of an upper bound for the entropy of English. *Computational Linguistics*, 18(1): 31–40.
- Deng, Y.; Liao, L.; Chen, L.; Wang, H.; Lei, W.; and Chua, T.-S. 2023. Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, 10602–10621.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805*.
- Holtzman, A.; Buys, J.; Du, L.; Forbes, M.; and Choi, Y. 2019. The curious case of neural text degeneration. *arXiv preprint arXiv:1904.09751*.
- Jang, Y.; Lee, J.; and Kim, K.-E. 2022. GPT-critic: Offline reinforcement learning for end-to-end task-oriented dialogue systems. In *10th International Conference on Learning Representations, ICLR 2022*. International Conference on Learning Representations, ICLR.
- Kingma, D. P.; and Ba, J. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- 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*.
- Liu, Y.; Ott, M.; Goyal, N.; Du, J.; Joshi, M.; Chen, D.; Levy, O.; Lewis, M.; Zettlemoyer, L.; and Stoyanov, V. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Mishra, K.; Priya, P.; Burja, M.; and Ekbal, A. 2023. e-THERAPIST: I suggest you to cultivate a mindset of positivity and nurture uplifting thoughts. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 13952–13967.
- Mishra, K.; Priya, P.; and Ekbal, A. 2023. PAL to Lend a Helping Hand: Towards Building an Emotion Adaptive Polite and Empathetic Counseling Conversational Agent. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 12254–12271.
- 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.
- Shi, W.; Li, Y.; Sahay, S.; and Yu, Z. 2021. Refine and Imitate: Reducing Repetition and Inconsistency in Persuasion Dialogues via Reinforcement Learning and Human Demonstration. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, 3478–3492.
- Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.-A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
- Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. 2019. Huggingface’s transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.
- Wu, Q.; Zhang, Y.; Li, Y.; and Yu, Z. 2021. Alternating Recurrent Dialog Model with Large-scale Pre-trained Language Models. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, 1292–1301.
- Zhang, T.; Kishore, V.; Wu, F.; Weinberger, K. Q.; and Artzi, Y. 2019. Bertscore: Evaluating text generation with bert. *arXiv preprint arXiv:1904.09675*.