

PREFERENCE LEARNING UNLOCKS LLMS’ PSYCHO-COUNSELING SKILLS

000
001
002
003
004
005
006
007
008
009
010
011
012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053
Anonymous authors
Paper under double-blind review

ABSTRACT

Applying large language models (LLMs) to assist in psycho-counseling is an emerging and meaningful approach, driven by the significant gap between patient needs and the availability of mental health support. However, current LLMs struggle to consistently provide effective responses to client speeches, largely due to the lack of supervision from high-quality real psycho-counseling data, whose content is typically inaccessible due to client privacy concerns. Furthermore, the quality of therapists’ responses in available sessions can vary significantly based on their professional training and experience. Assessing the quality of therapists’ responses remains an open challenge. In this work, we address these challenges by first proposing a set of professional and comprehensive principles to evaluate therapists’ responses to client speeches. Using these principles, we create a preference dataset, PsychoCounsel-Preference, which contains 36k high-quality preference comparison pairs. This dataset aligns with the preferences of professional psychotherapists, providing a robust foundation for evaluating and improving LLMs in psycho-counseling. Experiments on reward modeling and preference learning demonstrate that PsychoCounsel-Preference is an excellent resource for LLMs to acquire essential skills for responding to clients in a counseling session. Our best-aligned model, PsychoCounsel-Llama3-8B, achieves an impressive win rate of 87% against GPT-4o. We will release PsychoCounsel-Preference, PsychoCounsel-Llama3-8B and the reward model PsychoCounsel-Llama3-8B-Reward to facilitate the research of psycho-counseling with LLMs.

1 INTRODUCTION

According to the World Mental Health Report¹, the global demand for mental health support far exceeds the availability of accessible, affordable, and timely care. Millions of individuals struggle with mental health challenges, yet many face a severe shortage of trained professionals, particularly in low- and middle-income countries. The emergence of Large Language Models (LLMs) has accelerated the integration of AI into psycho-counseling, thanks to their remarkable ability to comprehend human intent and provide effective responses (Ouyang et al., 2022). However, due to the complexity of clients’ situations and the professional skills required, current LLMs still face challenges in consistently providing effective responses to client speeches during counseling sessions (Na et al., 2025; Zhang et al., 2024; Chung et al., 2023).

The sensitive and private nature of counseling sessions poses significant challenges for obtaining publicly available datasets that accurately reflect real-world interactions (Stade et al., 2024). This scarcity of resources hinders efforts to train LLMs in effectively understanding and responding to client speech within counseling contexts. Moreover, the quality of responses provided by therapists can vary widely, influenced by their level of professional training and experience (Rocco et al., 2019). This variability impacts the consistency of counseling effects and underscores the importance of standardizing and assessing therapist responses.

To address these gaps, we collaborated with professors in social work and psychiatry to develop a set of professional and comprehensive principles for evaluating therapists’ responses to client speeches. These principles assess not only the fundamental aspects of a response in a counseling session, such

¹<https://www.who.int/teams/mental-health-and-substance-use/world-mental-health-report>

054 as empathy, relevance, conciseness, and safety, but also extend the effectiveness of a response based
 055 on professional psycho-counseling theory. This includes evaluating whether the response promotes
 056 clients' self-exploration, enhances their autonomy, and effectively identifies the client's stage of
 057 change.

058 Using these principles, we extract high-quality responses from the generations of a pool of popular
 059 LLMs and construct a large psycho-counseling preference dataset, PsychoCounsel-Preference. The
 060 dataset comprises 26,483 unique client speeches spanning 8 coarse-grained and 42 fine-grained topics.
 061 To the best of our knowledge, PsychoCounsel-Preference is the first large-scale and comprehensive
 062 psycho-counseling preference dataset. We hired professional psychotherapists for verification, and
 063 their annotations exhibit strong agreement within PsychoCounsel-Preference, ensuring the dataset's
 064 reliability and consistency.

065 Experiments show that our reward models trained with PsychoCounsel-Preference show an excellent
 066 ability of evaluating responses to clients while previous start-of-the-art reward models lag behind.
 067 Moreover, we apply both online and offline preference learning on PsychoCounsel-Preference or
 068 the trained reward models. Our best resulting model, PsychoCounsel-Llama3-8B, achieves the
 069 state-of-the-art performance on the testing set of PsychoCounsel-Preference, with an impressive **win**
 070 **rate of 87% against GPT-40**. Feedback from professional psychologists shows that PsychoCounsel-
 071 Llama3-8B could give more balanced and desirable responses under length constraint during the
 072 inference stage. Through further analysis and case study, we demonstrate the advantage of training
 073 online over offline and provide insights into how to improve the model performance in the future.

074 2 RELATED WORK

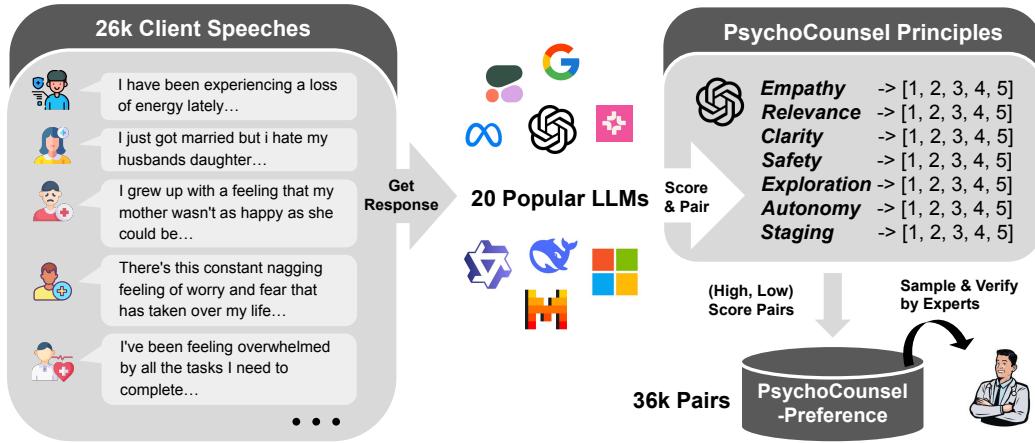
075 2.1 LLMs ASSISTING PSYCHOTHERAPY

076 Integrating LLMs into Psychotherapy is not a trivial process which could articulated as a continuum
 077 of stages of assistive AI, collaborated AI, and fully autonomous AI (Stade et al., 2024). Currently,
 078 we are still in the first two stages where models operating tasks need human supervision. Related
 079 tasks include cognitive disorder detection (Shreevastava & Foltz, 2021; Chen et al., 2023b), negative
 080 thoughts recognition and reframing (Maddela et al., 2023; Sharma et al., 2024), and patient simula-
 081 tion (Chen et al., 2023a) or therapist simulation Liu et al. (2023), among which therapist simulation is
 082 the primary goal across the stages. However, Zhang et al. (2024) found that due to the lack of public
 083 high-quality data in psychotherapy and the complexity of clients' situation, LLMs still are not able
 084 to give effective responses to a client's speech consistently in a therapy session. Our work focuses
 085 on psycho-counseling, which is a short-term, supportive process for helping individuals cope with
 086 life challenges and emotional distress, and constructs a high-quality preference dataset that aligns
 087 with the preference of professional psychotherapists and uses preference learning to train helpful and
 088 reliable assistants for psycho-counseling.

089 2.2 HUMAN PREFERENCE ALIGNMENT

090 Human preference alignment has been shown to be a critical step in making LLMs helpful, harmless,
 091 and honest (Ouyang et al., 2022; Bai et al., 2022a). Current methods can be broadly categorized
 092 into offline and online approaches. Offline methods optimize the model using a pre-annotated set of
 093 preference data with objectives such as DPO (Rafailov et al., 2023). Online methods, on the other
 094 hand, generate outputs during training and utilize a reward function to score them. High-scoring
 095 generations are encouraged, while low-scoring ones are discouraged through policy gradient methods
 096 such as PPO (Schulman et al., 2017).

097 Compared to offline alignment, online methods are more computationally expensive and require
 098 careful hyperparameter tuning to ensure stable training (Xu et al., 2024). Offline methods, which
 099 frame alignment as optimizing a classification loss, eliminate the need for a reward model, making
 100 them more stable and efficient. However, they are susceptible to distribution shifts (Marks et al.,
 101 2023). Furthermore, Tang et al. (2024) found that optimizing with online preferences instead of
 102 offline data can lead to better model performance. Iterative direct preference learning combines the
 103 strengths of both offline and online methods. In this approach, preference data is generated online
 104 and used to optimize an offline learning objective (Pang et al., 2024), which has been demonstrated
 105 as a strong baseline in both academia (Xu et al., 2024) and industry (Yang et al., 2024).

108
109
110
111
112
3 PSYCHOCOUNSEL-PREFERENCE

127
128 Figure 1: PsychoCounsel-Preference Construction Pipeline. 1) We first collect over 26k client
129 speeches covering a wide range of topics from various sources, applying necessary data cleaning.
130 2) 20 popular LLMs are sampled and prompted to roleplay as psychotherapists and give responses
131 to these client speeches. 3) GPT-4o is instructed to evaluate the responses based on our proposed
132 PsychoCounsel Principles, and preference pairs with substantial score gaps are incorporated into
133 PsychoCounsel-Preference.
134
135
136

137 3.1 CLIENT SPEECH COLLECTION
138

139 We collect client speeches from various data sources: counsel-chat², MentalAgora (Lee et al., 2024),
140 TherapistQA (Shreevastava & Foltz, 2021), Psycho8k (Liu et al., 2023), and several huggingface
141 datasets (amod-counsel³, MentalChat16K⁴, and phi2Mental⁵). Client speeches with number of
142 characters more than 1,000 and less than 100 are discarded to ensure a proper length of context.
143 After an additional step of de-duplication, the resulting data contains 26,483 client speeches with
144 average length of 366 characters covering a wide range of topics including 8 coarse topics: Core
145 Mental Health Issues (9,054), Emotional Well-being and Coping Strategies (5,717), Relationships and
146 Interpersonal Dynamics (6,483), Life Transitions and Challenges (934), Social Issues (667), Youth
147 and Development (1,175), Crisis and Safety Concerns (529) and Special Topics (1,924). Under these
148 8 topics are 42 fine-grained topics (see Table 5 in the appendix for the detailed topic distribution).
149
150

151 3.2 PSYCHOCOUNSEL PRINCIPLES
152

153 To answer the question *what is a good response to a client speech in psycho-counseling*, we collaborate
154 with experts in social work and psychiatry (our co-authors) and propose a set of professional principles
155 to measure the response to a client speech from seven different dimensions:
156
157

158
159 ²<https://github.com/nbertagnolli/counsel-chat>

160 ³https://huggingface.co/datasets/Amodeus/mental_health_counseling_conversations

161 ⁴<https://huggingface.co/datasets/ShenLab/MentalChat16K>

⁵<https://huggingface.co/datasets/saxenaindresh681/microsoft-phi2-mental-health>

162
163

PsychoCounsel Principles

164
165**Empathy and Emotional Understanding:** The response should convey genuine empathy, acknowledging and validating the client's feelings and experiences.166
167**Personalization and Relevance:** The response should be tailored to the client's unique situation, ensuring that the content is directly relevant to their concerns.168
169**Clarity and Conciseness:** The response should be clear, well-organized, and free of unnecessary jargon, making it easy for the client to understand and engage with.170
171**Avoidance of Harmful Language or Content:** The response should avoid any language or content that could potentially harm, distress, or trigger the client, ensuring the interaction is safe and supportive.172
173**Facilitation of Self-Exploration:** The response should encourage the client to reflect on their thoughts and feelings, promoting self-awareness and insight.

174

Promotion of Autonomy and Confidence: The response should support the client's sense of control over their decisions and encourage confidence in their ability to make positive changes.175
176
177
178
179
180
181**Sensitivity to the Stage of Change:** The response should recognize the client's current stage in the process of change and address their needs accordingly. If the client is in an early stage—uncertain or ambivalent about making a change—the response should help them explore their thoughts and motivations. If the client is in a later stage and has already made changes, the response should focus on reinforcing progress, preventing setbacks, and sustaining positive outcomes.

182

Please refer to Box 8 for the complete definition of the principles. Among these seven principles, **Facilitation of Self-Exploration**, **Promotion of Autonomy and Confidence**, and **Identifying Stages and Reasons for Change** emphasize a client-centered approach, which is recognized as a hallmark of effective psycho-counseling (Miller & Stephen Rollnick; Rooney et al., 2017; Hogarty, 2002; Tower, 1994). We use these three principles to measure the *effectiveness* of a response to a client speech, complementary to the other four principles, which are more basic requirements for an AI, requiring the response to be *empathy*, *relevant*, *concise*, and *safe* (Bai et al., 2022b; Ouyang et al., 2022). Evaluating therapist responses using these fine-grained principles provides a more structured and nuanced assessment of their effectiveness. Unlike general evaluations that focus solely on overall quality, this detailed approach allows for a deeper understanding of how well a response supports the client's emotional and psychological needs.

193

3.3 PREFERENCE GENERATION

194

We apply the generate-score-pair pipeline as Cui et al. (2023) to construct the PsychoCounsel-Preference dataset. For each client speech, we randomly sample four off-the-shelf LLMs from a model pool to give the response and instruct GPT-4o to annotate each response with 5-Likert scores for each principle defined in Section 3.2; higher scores mean more alignment with the principles. Then scores of the principles are averaged to get the overall score for a response and preference pairs are generated based on the overall scores. The whole pipeline is illustrated in Figure 1. To increase the diversity of the model responses, we initialize the model pool with 20 popular LLMs of a range of sizes developed by different organizations shown in Table 6. We also include LLMs with different architectures other than pure transformers like AI21-Jamba-1.5-Mini (Jamba Team et al., 2024), which is a hybrid transformer-mamba model. We randomly held out 3,291 client speeches for testing and the remaining 23,192 for training. After obtaining the scores of principles, for training, we extract response pairs with the overall score gap larger than or equal to 1 as the preference pairs, and for testing, we only extract the ([highest score response], [lowest score response]) pairs and pairs with the score gap less than 1 are discarded. In this way, we could exclude response pairs with similar scores, whose quality may be hard to differentiate.

210
211
212
213
214
215

Ultimately, PsychoCounsel-Preference includes 34,329 training preference pairs and 2,324 testing pairs. The models most likely to be chosen and those most likely to be rejected vary significantly in size (see Figures 5 and 6 for the distributions of chosen and rejected models). This suggests that simply scaling model size is not a decisive factor in making LLMs effective responders in psycho-counseling. We also observe that LLMs developed by non-English-speaking institutions are more likely to be rejected compared to those from English-speaking countries. This may suggest that non-English-speaking institutions have a greater need to enhance the capabilities of LLMs in their

216 respective languages, potentially leading to less emphasis on developing psycho-counseling skills in
 217 English.
 218

219 **3.4 PREFERENCE VALIDATION**
 220

221 To validate the quality of synthetic human preferences in PsychoCounsel-Preference, we hired two
 222 professional psychotherapists through Upwork⁶ and instructed them to annotate preferences based on
 223 each principle and give the overall preference. The annotation set consists of 200 preference pairs
 224 randomly sampled from PsychoCounsel-Preference. The two therapists agree on 174 out of 200
 225 samples. Additionally, one expert’s annotations align with the preference labels in PsychoCounsel-
 226 Preference for 184 out of 200 samples, while the other aligns for 170 out of 200 samples. These results
 227 indicate a high level of agreement between the experts (87%) and demonstrate strong alignment
 228 between the expert annotations and the preference labels in PsychoCounsel-Preference (88.5%). This
 229 strongly suggests that the labels in PsychoCounsel-Preference are reliable and trustworthy.
 230

231 **4 EXPERIMENTS**
 232

233 **4.1 REWARD MODEL**
 234

Model	Acc.	AUC (\uparrow)	ECE (\downarrow)	Brier (\downarrow)
State-of-the-art Reward Models				
Skywork-Reward-Llama-3.1-8B-v0.2 (Liu et al., 2024)	57.9	0.623	0.331	0.379
Skywork-Reward-Gemma-2-27B (Liu et al., 2024)	69.2	0.740	0.123	0.229
Llama-3.1-Nemotron-70B-Reward (Wang et al., 2024)	87.3	0.938	0.040	0.102
Generative LLMs				
gemma-2-9b-it (Gemma Team, 2024)	81.5	-	-	-
Mistral-Nemo-Instruct-2407 ⁷	78.0	-	-	-
Llama-3.1-8B-Instruct	80.1	-	-	-
Llama-3.1-70B-Instruct (Llama Team, 2024)	88.2	-	-	-
Our Reward Models				
PsychoCounsel-Llama3-3B-Reward	98.1	0.997	0.050	0.014
PsychoCounsel-Llama3-8B-Reward	97.8	0.998	0.045	0.016

249 Table 1: Performance on the Testing Set of PsychoCounsel-Preference
 250

251 Following Ouyang et al. (2022) and Bai et al. (2022a), we train Bradley-Terry (BT) style reward
 252 models $r_\theta(\cdot)$ where a linear head added on the top of LLMs outputs a scalar reward. Given a pair of
 253 preference data $\{y_c, y_r\}$ to a prompt x , the objective is to optimize the reward gap between chosen
 254 response y_c and rejected response y_r :

$$255 \quad \mathcal{L} = -\log(\sigma(r_\theta(x, y_c) - r_\theta(x, y_r))),$$

257 where the sigmoid function $\sigma(\cdot)$ generates the probability of y_c preferred than y_r .
 258

259 We use Llama3.2-3B-Instruct and Llama3.1-8B-Instruct (abbreviated as Llama3-3B and Llama3-8B)
 260 to initialize the BT models, training them on PsychoCounsel-Preference for 2 epochs with a batch
 261 size of 128 and a learning rate of 9e-6. To evaluate our reward models, we compare them against
 262 three state-of-the-art reward models that rank highly on RewardBench (Lambert et al., 2024), as well
 263 as three popular LLMs, which are prompted to rank responses (see Box 7 for the prompt).

264 The overall results on PsychoCounsel-Preference testing response pairs are shown in Table 1. Our
 265 reward models significantly outperform all other reward models and generative LLMs, achieving
 266 notably high accuracy and ROC AUC Score (Bradley, 1997) on the PsychoCounsel-Preference
 267 testing set. These results suggest that PsychoCounsel-Preference provides robust supervision for
 268 training powerful reward models capable of effectively ranking responses to client speeches. We also
 269

⁶<https://www.upwork.com/>

⁷<https://mistral.ai/news/mistral-nemo/>

Model	Training Data	Acc.	AUC (\uparrow)	Brier (\downarrow)	ECE (\downarrow)	Acc. (RewardBench)
Llama-3B	HelpSteer2	81.6	0.916	0.120	0.044	83.6
	HelpSteer2 + Ours	97.6	0.998	0.017	0.045	86.1
Llama-8B	HelpSteer2	81.7	0.898	0.128	0.019	86.6
	HelpSteer2 + Ours	97.5	0.998	0.018	0.040	87.2

Table 2: Ablation study on reward models trained with HelpSteer2 vs. HelpSteer2 + our PsychoCounsel-Preference dataset. Results are reported on our test set and RewardBench.

calculate the Expected Calibration Error (ECE) (Naeini et al., 2015) and Brier Score⁸ to assess the calibration level of the models. The results demonstrate that our reward models have comparable and low ECE values to the state-of-the-art reward model, Llama-3.1-Nemotron-70B-Reward, while achieving significantly better Brier Scores. This indicates that our reward models could give more reliable rewards and preference probability.

To further validate the effectiveness of our reward models, we train Llama3-3B and Llama3-8B under two settings: (i) on HelpSteer2 (Wang et al., 2024), a general-domain preference dataset, and (ii) on a merged dataset consisting of HelpSteer2 and our newly developed PsychoCounsel-Preference. For evaluation, we use both our PsychoCounsel-Preference test set and RewardBench (Lambert et al., 2024), a widely adopted general benchmark for reward models. The results, presented in Table 2, demonstrate that incorporating PsychoCounsel-Preference consistently improves performance across both Llama model sizes and the gains are not confined to domain-specific test set but also generalize to RewardBench. This indicates that PsychoCounsel-Preference provides complementary supervision to HelpSteer2 and enhances reward modeling in both in-domain and out-of-domain settings.

4.2 POLICY MODEL

Model	Mental Issues	Emotional Well-being	Relationships	Life Changes	Social Issues	Youth	Safety	Special Topics	Overall
w/o Length Constraint									
Llama3-3B	30.5	27.1	26.7	30.4	28.4	24.7	36.9	28.9	28.5
+ DPO	<u>57.7</u>	<u>59.1</u>	<u>57.9</u>	<u>63.5</u>	<u>54.3</u>	<u>53.4</u>	<u>60.0</u>	<u>64.7</u>	<u>58.5</u>
+ DPO-Iter	66.7	70.9	68.7	75.7	70.4	65.8	75.4	77.9	69.4
Llama3-8B	28.9	31.5	28.0	33.9	23.5	29.5	33.8	26.8	29.3
+ DPO	70.2	74.5	73.7	74.8	77.8	73.3	80.0	74.0	72.9
+ DPO-Iter	86.3	88.2	87.1	87.0	91.4	87.0	90.8	84.3	87.0
w/ Length Constraint									
Llama3-3B	15.1	15.5	15.0	15.7	13.6	13.7	9.20	15.3	15.0
+ DPO	36.8	38.1	35.6	37.4	42.0	39.7	30.8	37.9	37.0
+ DPO-Iter	47.5	46.7	46.3	40.9	46.9	45.2	47.7	43.4	46.4
Llama3-8B	19.3	17.9	17.1	21.7	21.0	17.1	18.5	19.6	18.5
+ DPO	50.4	48.0	47.3	53.9	45.7	54.1	46.2	51.9	49.3
+ DPO-Iter	75.6	77.9	79.2	77.4	76.5	73.3	83.1	74.0	77.0

Table 3: Win rates (%) of models trained with two different methods (+DPO or +DPO-Iter) on PsychoCounsel-Preference under two different settings (w/o and w/ Length Constraint) compared to GPT-4o. **Bold** numbers indicate the best performance in each column for each setting. The overall win rates of models that outperformed GPT-4o are underlined.

To further verify the effectiveness of PsychoCounsel-Preference and the trained reward models, we employ two preference alignment methods to optimize base models. **1) DPO**: we directly optimize the DPO (Rafailov et al., 2023) objective on PsychoCounsel-Preference:

⁸https://en.wikipedia.org/wiki/Brier_score

324

325

326

327

$$\mathcal{L}_{\text{DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(x, y_c, y_r) \sim \mathcal{D}} \left[\log \sigma \left(\beta \log \frac{\pi_\theta(y_c | x)}{\pi_{\text{ref}}(y_c | x)} - \beta \log \frac{\pi_\theta(y_r | x)}{\pi_{\text{ref}}(y_r | x)} \right) \right].$$

328

329

330

331

332

333

334

335

336

2) DPO-Iter: we follow an iterative approach (Pang et al., 2024), where, in each iteration, 8 responses are generated for each client speech and ranked by the reward model of the same size as the base model. The responses with the highest and lowest rewards are then annotated as online preference pairs, which are used to train the base model with the DPO objective. The client speeches for each iteration are 6400 sampled from the train set of PsychoCounsel-Preference. We use Llama3.2-3B-Instruct and Llama3.1-8B-Instruct as the base models. The training configuration includes a batch size of 64, a learning rate of 5e-7, and a total of 1,600 training steps. A development set comprising 10% of the training set from PsychoCounsel-Preference is used to select the best checkpoints. We set the value of β as 0.1 for DPO across all the experiments.

337

338

339

340

341

342

343

344

345

Evaluation We use LLM-as-judge (Zheng et al., 2023) to effectively approximate human preferences (validated by the human experts) for evaluation. We prompt the model to generate responses for the testing client speeches in PsychoCounsel-Preference and leverage GPT-4o to compare these responses against those of GPT-4o using the proposed PsychoCounsel Principles. Specifically, we prompt the model in two settings: **1) w/o Length Constraint:** The models are instructed to act as therapists and respond to the given client speech without any restrictions on response length. **2) w/ Length Constraint:** To ensure a fairer comparison with GPT-4o, we impose a length constraint, requiring the models to generate responses of similar length to those produced by GPT-4o. The overall win rates of the models against GPT-4o are calculated for comparison. We also show the win rates for the coarse topic categories.

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

Main Results As shown in Table 3, in the w/o Length Constraint setting, the base models have low probabilities of outperforming GPT-4o. However, the models after alignment demonstrate significantly higher win rates against GPT-4o, indicating that supervision from PsychoCounsel-Preference effectively guides the models in learning how to respond to client speeches. Notably, Llama3-8B(+DPO-Iter) achieves the best performance, with a high overall win rate of **87.0% against GPT-4o**. This result suggests that online training and larger model sizes can potentially enhance generation quality, and **models with approximately 8B parameters can effectively develop the skills to respond to client speeches under the guidance of reward models trained on PsychoCounsel-Preference**. Compared to models in the w/o Length Constraint setting, those in the w/ Length Constraint setting generally have lower win rates against GPT-4o. We attribute this to the stricter generation constraint, which requires our models to align their response length with that of GPT-4o. However, our model, Llama3-8B (+DPO-Iter), still achieves a high win rate of 77% against GPT-4o, demonstrating that with proper training, the model can develop a robust ability to effectively respond to clients, regardless of generation constraints such as response length. We refer to the best model Llama3-8B (+DPO-Iter) as PsychoCounsel-Llama3-8B.

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

Human Evaluation We instruct the hired psychotherapists to provide preference judgments between the 200 randomly sampled response pairs generated by PsychoCounsel-Llama3-8B and GPT-4o, among which 100 for w/o Length Constraint setting and 100 for w/ Length Constraint. The provided order is shuffled to eliminate any position bias in the evaluation. In 82.5% of cases, GPT-4o and human experts made the same judgments, indicating that **GPT-4o serves as a reliable evaluator for assessing psycho-counseling responses**. Figure 2 presents the human experts' comparison between the two models based on the annotation principles. Overall, **real experts clearly prefer the outputs of PsychoCounsel-Llama3-8B across both evaluation settings and nearly all principles**. Only if no length constraint is applied, PsychoCounsel-Llama3-8B exhibits lower clarity compared to GPT-4o. This is primarily because PsychoCounsel-Llama3-8B tends to generate longer responses, which aligns with the observed phenomenon that as LLMs develop more complex capabilities through reinforcement learning (RL), they tend to produce more tokens (DeepSeek-AI et al., 2025). However, in the w/ Length Constraint setting, where models generate responses of similar length, PsychoCounsel-Llama3-8B demonstrates better performance in *Clarity*, *Safety* and *Staging*. This suggests that applying a length constraint after RL training is a promising approach to obtain more balanced and desirable generations.

376

377

Additionally, higher win rates only indicate the **relative** quality of responses. To provide an **absolute** assessment of the responses generated by PsychoCounsel-Llama3-8B, we instructed experts to assign fine-grained scores based on the PsychoCounsel Principles. Figure 3 presents the average scores

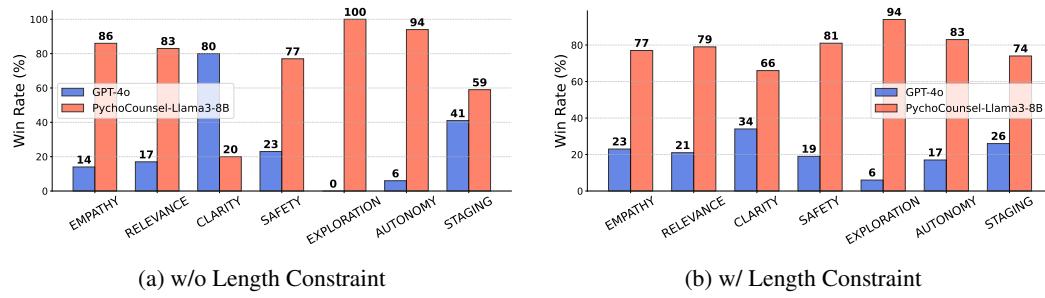


Figure 2: Experts' Comparison between GPT-4o and PsychoChat-Llama3-8B in Two Settings

of 100 randomly sampled responses, evaluated by experts under two different settings. We observe that, except for *Clarity* in the w/ Length Constraint setting, PsychoCounsel-Llama3-8B achieves consistently high scores (>4) across all principles, indicating a strong alignment with the criteria for effective responses in psycho-counseling.

4.3 ABLATION STUDY

To explore the differences between training on offline vs online data and base models with different sizes, we set up a controlled experimental group. In this setup, base models are trained by DPO with two different sets of preference data, one is offline preferences from PsychoCounsel-Preference, and the other is trained using online preferences generated by the base model. Responses are selected by the reward model of the same size trained on PsychoCounsel-Preference. All the other experimental settings are kept identical with the training epoch as 1, learning rate as 5e-7, and global batch size as 64. Figure 4 illustrates the win rates of checkpoints against GPT-4o on the testing client speeches of PsychoCounsel-Preference.

In general, training on online samples demonstrates clear advantages over offline training: **1) Better Performance:** For Llama3-3B, training with online data (green line) consistently achieves a higher win rate compared to training with offline data (orange line). Similarly, for Llama3-8B, training with online data (blue line) delivers performance comparable to the best checkpoints of training with offline data (red line). **2) Greater Stability:** Both offline training curves exhibit a pronounced hump-shaped pattern, a known indicator of reward hacking (Rafailov et al., 2024). In contrast, the performance of online training models remains more stable across training steps. **3) Enhanced Compatibility with Smaller LLMs:** Online samples enable Llama3-3B to perform on par with, and sometimes even surpass, Llama3-8B models. This highlights the significant potential of combining online training with smaller LLMs. We can also observe that in this setting, training online can be viewed as DPO-Iter with only one round of updates on the policy model. However, its performance lags significantly behind DPO-Iter (shown in Table 3), indicating that utilizing online generations from the latest updated policy is crucial for more effective online preference learning.

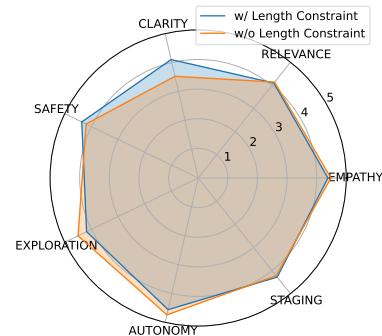


Figure 3: Absolute Scores

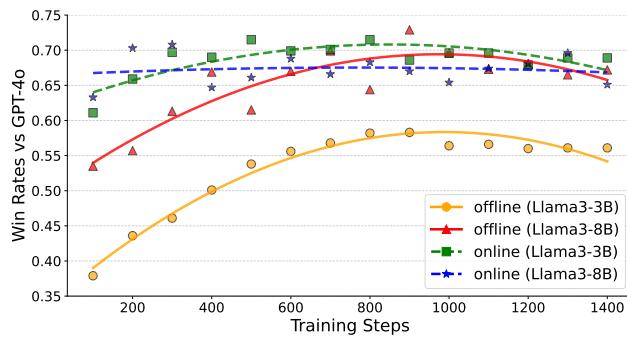


Figure 4: Comparison of Training Online or Offline

432 4.4 CASE STUDY
433

434 We sample testing cases and compare the outputs of GPT-4o with those of our best policy model,
435 PsychoCounsel-Llama3-8B. Generally, PsychoCounsel-Llama3-8B provides better responses than
436 GPT-4o. Table 4 provides a typical example in the w/ Length Constraint setting. We could see
437 that the response of PsychoCounsel-Llama3-8B, not only **validates the client’s distress with deep**
438 **empathy**—acknowledging both her emotional burden and the courage it took to share—but also
439 **engages her in a detailed, collaborative exploration of her experiences**. By inviting her to pinpoint
440 specific patterns and triggers behind her fears, Response 1 promotes self-exploration and empow-
441 erment, making it particularly effective for someone in the early stages of considering change. In
442 contrast, the response of GPT-4o is **general and less detailed**, which can make the client feel less
443 deeply understood. We provide more cases in Appendix 9.
444

445 **Client Speech**
446

447 I’ve been having emotional issues for a few years. Nonetheless, these have been somewhat manageable. However, I became increasingly
448 paranoid this winter. I thought that people were trying to poison me, I feared that family members were going to kill me. I was aware that
449 these thoughts were illogical, but I couldn’t shake the fear that they caused me. I would get panic attacks thinking that I drank from a
450 poisoned water fountain. I thought that someone had put small rips in my ice cream sandwich packaging because there was only one that
451 didn’t have tiny rips in it. I refused to eat this one because I assumed some reverse psychology was being pulled and the ones without the
452 rips were fine. I thought that this was mostly caused by anxiety but there are other symptoms that lead me to believe it is something more.

453 **PsychoCounsel-Llama3-8B (chosen)**
454

455 I can sense the immense distress and fear that you’ve been experiencing, and I’m here to create a safe space for you to explore these
456 feelings. It takes a lot of courage to acknowledge and share these thoughts with me. I want to acknowledge that it sounds like you’ve been
457 carrying a heavy burden, and it’s understandable that you felt the need to try to make sense of these experiences, even if they seemed
458 illogical at times.

459 **⇒ (detailed acknowledgment of the client’s feelings tends to create a stronger emotional connection, making the client feel more**
460 **profoundly understood)**

461 I’d like to gently explore this further. You mentioned that you recognized these thoughts as anxiety-driven, but there are other symptoms
462 that make you question whether it’s more than that. Can you tell me more about what specifically makes you think it’s not just anxiety?
463 What are those other symptoms that have led you to consider this? Let’s work together to tease out what might be contributing to these
464 feelings, and see if we can identify any patterns or triggers that might be at play. We can also explore ways to help you manage these
465 fears and develop more grounding techniques to reduce your anxiety. How does that sound?

466 **⇒ (engage the client in a detailed, collaborative exploration of her experiences)**

467 **GPT-4o (rejected)**
468

469 Thank you for sharing all of this with me. It sounds like you’ve been going through a particularly difficult time with these intense feelings
470 and thoughts. It must be very distressing to experience such fears and the resulting panic attacks. I want to make sure we explore this in a
471 way that feels safe for you.

472 **⇒ (more general in validating the client’s feelings)**

473 You mentioned that while you’ve considered it might be anxiety, there are additional symptoms that make you think there might be more
474 to it. Could you tell me more about these other symptoms? Understanding the full picture can help us figure out what might be happening
475 and how best to support you.

476 **⇒ (lacks the specific, collaborative invitation to explore the triggers)**

477 Also, it’s important to acknowledge the awareness you’ve shown about the thoughts being illogical at times, yet still powerful. This
478 awareness is a valuable strength as we work through these experiences together. Let’s take it one step at a time and see if we can find
479 patterns or triggers for these thoughts and discuss some strategies that might help manage them.

480 **⇒ (general and less detailed)**

481
482 **Table 4: Case Study**
483484 5 CONCLUSION AND FUTURE WORK
485

486 In this work, we introduce a set of professional and comprehensive principles for evaluating therapists’
487 responses to client speeches in psycho-counseling, along with PsychoCounsel-Preference, a prefer-
488 ence dataset containing 36k high-quality preference comparison pairs. Our experiments show that
489 with PsychoCounsel-Preference, preference learning could effectively unlock LLMs’ professional
490 psycho-counseling skills. In the future, we will explore how to reduce the reward hacking problem in
491 preference learning and ways to increase the reliability of LLMs assisting psycho-counseling.
492

493 6 ETHICS STATEMENT
494

495 This project has been classified as exempt by the Institutional Review Board (IRB). All hired experts
496 were at least 18 years old and hold either a master’s or doctoral degree in a mental health-related field,
497 such as psychology or counseling psychology. Each expert received a fixed payment of \$1,500 for all
498

486 annotations, corresponding to an approximate hourly rate of \$60. The goal of this work is to leverage
 487 synthetic data and preference learning algorithms to equip LLMs with the skills needed to generate
 488 responses to client speeches in psycho-counseling. However, these responses should not be directly
 489 exposed to clients without review by real therapists. Instead, they serve as assistive suggestions to
 490 help therapists draft responses, improving the efficiency of psycho-counseling:

491 **Enhancing Human Expertise, Not Replacing It.** AI tools should function as supportive mechanisms
 492 that enhance the capabilities of mental health professionals rather than replacing them. These tools
 493 can provide therapists with valuable data-driven insights into client speech, suggest psycho-counseling
 494 responses, and assist in structuring interventions. However, their role remains assistive, ensuring
 495 human expertise remains central to patient care.

496 **Training and Ethical Integration.** The effective use of AI in psycho-counseling requires mental
 497 health professionals to receive specialized training. This ensures they can integrate AI-generated
 498 insights into their practice ethically and effectively, maintaining both professional oversight and
 499 adherence to best practices.

500 **Safeguards Against Unsupervised AI Interaction.** To uphold safety and ethical integrity, AI-
 501 generated insights and psycho-counseling suggestions should always be reviewed by a licensed
 502 professional before reaching a patient. Deployment models must include strict access controls,
 503 intervention thresholds, and supervision mechanisms to prevent autonomous operation without
 504 human oversight.

505 Transparency and Accountability

- 506 • Open Communication: AI deployment in mental health should involve clear and open
 507 communication with all stakeholders, including therapists, patients, and regulatory bodies.
 508 This fosters trust and ensures transparency in the development and use of AI tools.
- 509 • Explainability and Justification: AI-generated recommendations should be interpretable,
 510 providing clear reasoning behind decisions. This is particularly crucial for psycho-counseling
 511 suggestions and mental health assessments, where explainability is essential to professional
 512 trust and responsible use.
- 513 • User Awareness: Patients and therapists interacting with AI must be fully informed about
 514 the system's role, capabilities, and limitations to prevent over-reliance and misapplication.

517 Safety and Privacy Standards

- 518 • Error Mitigation: AI models should be rigorously tested to minimize the risk of errors
 519 in medical advice or psychological recommendations. Misdiagnoses or inappropriate
 520 interventions could have significant negative consequences.
- 521 • Preventing Misinformation and Hallucinations: AI systems must prioritize accuracy by
 522 reducing misinformation and hallucinations, ensuring responses are evidence-based and
 523 context-appropriate.
- 524 • Data Privacy and Confidentiality: AI tools in mental health care must adhere to strict
 525 data privacy regulations, ensuring that patient interactions remain secure and confidential.
 526 Compliance with legal and ethical data-handling standards is critical to protecting users
 527 from breaches or misuse.

529 REFERENCES

530 Marah Abdin, Jyoti Aneja, Hany Awadalla, Ahmed Awadallah, Ammar Ahmad Awan, Nguyen
 531 Bach, Amit Bahree, Arash Bakhtiari, Jianmin Bao, Harkirat Behl, Alon Benhaim, Misha Bilenko,
 532 Johan Bjorck, Sébastien Bubeck, Martin Cai, Qin Cai, Vishrav Chaudhary, Dong Chen, Dongdong
 533 Chen, Weizhu Chen, Yen-Chun Chen, Yi-Ling Chen, Hao Cheng, Parul Chopra, Xiyang Dai,
 534 Matthew Dixon, Ronen Eldan, Victor Fragoso, Jianfeng Gao, Mei Gao, Min Gao, Amit Garg,
 535 Allie Del Giorno, Abhishek Goswami, Suriya Gunasekar, Emman Haider, Junheng Hao, Russell J
 536 Hewett, Wenxiang Hu, Jamie Huynh, Dan Iter, Sam Ade Jacobs, Mojan Javaheripi, Xin Jin,
 537 Nikos Karampatziakis, Piero Kauffmann, Mahoud Khademi, Dongwoo Kim, Young Jin Kim, Lev
 538 Kurilenko, James R Lee, Yin Tat Lee, Yuanzhi Li, Yunsheng Li, Chen Liang, Lars Liden, Xihui
 539 Lin, Zeqi Lin, Ce Liu, Liyuan Liu, Mengchen Liu, Weishung Liu, Xiaodong Liu, Chong Luo,

540 Piyush Madan, Ali Mahmoudzadeh, David Majercak, Matt Mazzola, Caio César Teodoro Mendes,
 541 Arindam Mitra, Hardik Modi, Anh Nguyen, Brandon Norick, Barun Patra, Daniel Perez-Becker,
 542 Thomas Portet, Reid Pryzant, Heyang Qin, Marko Radmilac, Liliang Ren, Gustavo de Rosa,
 543 Corby Rosset, Sambudha Roy, Olatunji Ruwase, Olli Saarikivi, Amin Saied, Adil Salim, Michael
 544 Santacroce, Shital Shah, Ning Shang, Hiteshi Sharma, Yelong Shen, Swadheen Shukla, Xia Song,
 545 Masahiro Tanaka, Andrea Tupini, Praneetha Vaddamanu, Chunyu Wang, Guanhua Wang, Lijuan
 546 Wang, Shuohang Wang, Xin Wang, Yu Wang, Rachel Ward, Wen Wen, Philipp Witte, Haiping Wu,
 547 Xiaoxia Wu, Michael Wyatt, Bin Xiao, Can Xu, Jiahang Xu, Weijian Xu, Jilong Xue, Sonali Yadav,
 548 Fan Yang, Jianwei Yang, Yifan Yang, Ziyi Yang, Donghan Yu, Lu Yuan, Chenruidong Zhang, Cyril
 549 Zhang, Jianwen Zhang, Li Lyna Zhang, Yi Zhang, Yue Zhang, Yunan Zhang, and Xiren Zhou.
 550 Phi-3 technical report: A highly capable language model locally on your phone. *arXiv [cs.CL]*,
 551 April 2024.

552 Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn
 553 Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson
 554 Kernion, Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez,
 555 Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt, Neel Nanda, Catherine Olsson, Dario
 556 Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, Ben Mann, and Jared Kaplan.
 557 Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv*
 558 [*cs.CL*], April 2022a.

559 Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones,
 560 Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, Carol Chen, Catherine Olsson,
 561 Christopher Olah, Danny Hernandez, Dawn Drain, Deep Ganguli, Dustin Li, Eli Tran-Johnson,
 562 Ethan Perez, Jamie Kerr, Jared Mueller, Jeffrey Ladish, Joshua Landau, Kamal Ndousse, Kamile
 563 Lukosuite, Liane Lovitt, Michael Sellitto, Nelson Elhage, Nicholas Schiefer, Noemi Mercado, Nova
 564 DasSarma, Robert Lasenby, Robin Larson, Sam Ringer, Scott Johnston, Shauna Kravec, Sheer
 565 El Showk, Stanislav Fort, Tamera Lanham, Timothy Telleen-Lawton, Tom Conerly, Tom Henighan,
 566 Tristan Hume, Samuel R Bowman, Zac Hatfield-Dodds, Ben Mann, Dario Amodei, Nicholas
 567 Joseph, Sam McCandlish, Tom Brown, and Jared Kaplan. Constitutional AI: Harmlessness from
 568 AI feedback. *arXiv [cs.CL]*, December 2022b.

569 Andrew P Bradley. The use of the area under the ROC curve in the evaluation of machine learning
 570 algorithms. *Pattern Recognit.*, 30(7):1145–1159, July 1997.

571 Du Chen, Yi Huang, Xiaopu Li, Yongqiang Li, Yongqiang Liu, Haihui Pan, Leichao Xu, Dacheng
 572 Zhang, Zhipeng Zhang, and Kun Han. Orion-14B: Open-source multilingual large language
 573 models. *arXiv [cs.CL]*, January 2024.

574 Siyuan Chen, Mengyue Wu, Kenny Q Zhu, Kunyao Lan, Zhiling Zhang, and Lyuchun Cui. LLM-
 575 empowered chatbots for psychiatrist and patient simulation: Application and evaluation. *arXiv*
 576 [*cs.CL*], May 2023a.

577 Zhiyu Chen, Yujie Lu, and William Yang Wang. Empowering psychotherapy with large language
 578 models: Cognitive distortion detection through diagnosis of thought prompting. *arXiv [cs.CL]*,
 579 October 2023b.

580 Neo Christopher Chung, George Dyer, and Lennart Brocki. Challenges of large language models for
 581 mental health counseling. *arXiv [cs.CL]*, November 2023.

582 Ganqu Cui, Lifan Yuan, Ning Ding, Guanming Yao, Wei Zhu, Yuan Ni, Guotong Xie, Zhiyuan Liu,
 583 and Maosong Sun. UltraFeedback: Boosting language models with high-quality feedback. *arXiv*
 584 [*cs.CL*], October 2023.

585 DeepSeek-AI, Xiao Bi, Deli Chen, Guanting Chen, Shanhuang Chen, Damai Dai, Chengqi Deng,
 586 Honghui Ding, Kai Dong, Qiushi Du, Zhe Fu, Huazuo Gao, Kaige Gao, Wenjun Gao, Ruiqi Ge,
 587 Kang Guan, Daya Guo, Jianzhong Guo, Guangbo Hao, Zhewen Hao, Ying He, Wenjie Hu, Panpan
 588 Huang, Erhang Li, Guowei Li, Jia Shi Li, Yao Li, Y K Li, Wenfeng Liang, Fangyun Lin, A X Liu,
 589 Bo Liu, Wen Liu, Xiaodong Liu, Xin Liu, Yiyuan Liu, Haoyu Lu, Shanghao Lu, Fuli Luo, Shirong
 590 Ma, Xiaotao Nie, Tian Pei, Yishi Piao, Junjie Qiu, Hui Qu, Tongzheng Ren, Zehui Ren, Chong
 591 Ruan, Zhangli Sha, Zhihong Shao, Junxiao Song, Xuecheng Su, Jingxiang Sun, Yaofeng Sun,
 592 Minghui Tang, Bingxuan Wang, Peiyi Wang, Shiyu Wang, Yaohui Wang, Yongji Wang, Tong Wu,
 593

594 Y Wu, Xin Xie, Zhenda Xie, Ziwei Xie, Yiliang Xiong, Hanwei Xu, R X Xu, Yanhong Xu, Dejian
 595 Yang, Yuxiang You, Shuiping Yu, Xingkai Yu, B Zhang, Haowei Zhang, Lecong Zhang, Liyue
 596 Zhang, Mingchuan Zhang, Minghua Zhang, Wentao Zhang, Yichao Zhang, Chenggang Zhao, Yao
 597 Zhao, Shangyan Zhou, Shunfeng Zhou, Qihao Zhu, and Yuheng Zou. DeepSeek LLM: Scaling
 598 open-source language models with longtermism. *arXiv [cs.CL]*, January 2024.

599

600 DeepSeek-AI, Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu,
 601 Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, Xiaokang Zhang, Xingkai Yu, Yu Wu, Z F Wu,
 602 Zhibin Gou, Zhihong Shao, Zhuoshu Li, Ziyi Gao, Aixin Liu, Bing Xue, Bingxuan Wang, Bochao
 603 Wu, Bei Feng, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan,
 604 Damai Dai, Deli Chen, Dongjie Ji, Erhang Li, Fangyun Lin, Fucong Dai, Fuli Luo, Guangbo Hao,
 605 Guanting Chen, Guowei Li, H Zhang, Han Bao, Hanwei Xu, Haocheng Wang, Honghui Ding,
 606 Huajian Xin, Huazuo Gao, Hui Qu, Hui Li, Jianzhong Guo, Jiashi Li, Jiawei Wang, Jingchang
 607 Chen, Jingyang Yuan, Junjie Qiu, Junlong Li, J L Cai, Jiaqi Ni, Jian Liang, Jin Chen, Kai Dong,
 608 Kai Hu, Kaige Gao, Kang Guan, Kexin Huang, Kuai Yu, Lean Wang, Lecong Zhang, Liang Zhao,
 609 Litong Wang, Liyue Zhang, Lei Xu, Leyi Xia, Mingchuan Zhang, Minghua Zhang, Minghui Tang,
 610 Meng Li, Miaojun Wang, Mingming Li, Ning Tian, Panpan Huang, Peng Zhang, Qiancheng Wang,
 611 Qinyu Chen, Qiushi Du, Ruiqi Ge, Ruisong Zhang, Ruizhe Pan, Runji Wang, R J Chen, R L
 612 Jin, Ruyi Chen, Shanghai Lu, Shangyan Zhou, Shanhuan Chen, Shengfeng Ye, Shiyu Wang,
 613 Shuiping Yu, Shunfeng Zhou, Shuting Pan, S S Li, Shuang Zhou, Shaoqing Wu, Shengfeng Ye,
 614 Tao Yun, Tian Pei, Tianyu Sun, T Wang, Wangding Zeng, Wanja Zhao, Wen Liu, Wenfeng Liang,
 615 Wenjun Gao, Wenqin Yu, Wentao Zhang, W L Xiao, Wei An, Xiaodong Liu, Xiaohan Wang,
 616 Xiaokang Chen, Xiaotao Nie, Xin Cheng, Xin Liu, Xin Xie, Xingchao Liu, Xinyu Yang, Xinyuan
 617 Li, Xuecheng Su, Xuheng Lin, X Q Li, Xiangyu Jin, Xiaojin Shen, Xiaosha Chen, Xiaowen Sun,
 618 Xiaoxiang Wang, Xinnan Song, Xinyi Zhou, Xianzu Wang, Xinxia Shan, Y K Li, Y Q Wang, Y X
 619 Wei, Yang Zhang, Yanhong Xu, Yao Li, Yao Zhao, Yaofeng Sun, Yaohui Wang, Yi Yu, Yichao
 620 Zhang, Yifan Shi, Yiliang Xiong, Ying He, Yishi Piao, Yisong Wang, Yixuan Tan, Yiyang Ma,
 621 Yiyuan Liu, Yongqiang Guo, Yuan Ou, Yuduan Wang, Yue Gong, Yuheng Zou, Yujia He, Yunfan
 622 Xiong, Yuxiang Luo, Yuxiang You, Yuxuan Liu, Yuyang Zhou, Y X Zhu, Yanhong Xu, Yanping
 623 Huang, Yaohui Li, Yi Zheng, Yuchen Zhu, Yunxian Ma, Ying Tang, Yukun Zha, Yuting Yan, Z Z
 624 Ren, Zehui Ren, Zhangli Sha, Zhe Fu, Zhean Xu, Zhenda Xie, Zhengyan Zhang, Zhewen Hao,
 625 Zhicheng Ma, Zhigang Yan, Zhiyu Wu, Zihui Gu, Zijia Zhu, Zijun Liu, Zilin Li, Ziwei Xie, Ziyang
 626 Song, Zizheng Pan, Zhen Huang, Zhipeng Xu, Zhongyu Zhang, and Zhen Zhang. DeepSeek-R1:
 627 Incentivizing reasoning capability in LLMs via reinforcement learning. *arXiv [cs.CL]*, January
 628 2025.

629

630 Gemma Team. Gemma 2: Improving open language models at a practical size. *arXiv [cs.CL]*, July
 631 2024.

632

633 Gerard E Hogarty. *Personal therapy for schizophrenia and related disorders: A guide to individualized
 634 treatment*. Guilford Press, 2002.

635

636 Shengding Hu, Yuge Tu, Xu Han, Chaoqun He, Ganqu Cui, Xiang Long, Zhi Zheng, Yewei Fang,
 637 Yuxiang Huang, Weilin Zhao, Xinrong Zhang, Zheng Leng Thai, Kaihuo Zhang, Chongyi Wang,
 638 Yuan Yao, Chenyang Zhao, Jie Zhou, Jie Cai, Zhongwu Zhai, Ning Ding, Chao Jia, Guoyang Zeng,
 639 Dahai Li, Zhiyuan Liu, and Maosong Sun. MiniCPM: Unveiling the potential of small language
 640 models with scalable training strategies. *arXiv [cs.CL]*, April 2024.

641

642 Jamba Team, Barak Lenz, Alan Arazi, Amir Bergman, Avshalom Manevich, Barak Peleg, Ben
 643 Aviram, Chen Almagor, Clara Fridman, Dan Padnos, Daniel Gissin, Daniel Jannai, Dor Muhlgay,
 644 Dor Zimberg, Edden M Gerber, Elad Dolev, Eran Krakovsky, Erez Safahi, Erez Schwartz, Gal
 645 Cohen, Gal Shachaf, Haim Rozenblum, Hofit Bata, Ido Blass, Inbal Magar, Itay Dalmedigos,
 646 Jhonathan Osin, Julie Fadlon, Maria Rozman, Matan Danos, Michael Gokhman, Mor Zusman,
 647 Naama Gidron, Nir Ratner, Noam Gat, Noam Rozen, Oded Fried, Ohad Leshno, Omer Antverg,
 648 Omri Abend, Opher Lieber, Or Dagan, Orit Cohavi, Raz Alon, Ro'i Belson, Roi Cohen, Rom Gilad,
 649 Roman Glozman, Shahar Lev, Shaked Meirom, Tal Delbari, Tal Ness, Tomer Asida, Tom Ben Gal,
 650 Tom Braude, Uriya Pumerantz, Yehoshua Cohen, Yonatan Belinkov, Yuval Globerson, Yuval Peleg
 651 Levy, and Yoav Shoham. Jamba-1.5: Hybrid transformer-mamba models at scale. *arXiv [cs.CL]*,
 652 August 2024.

648 Nathan Lambert, Valentina Pyatkin, Jacob Morrison, L J Miranda, Bill Yuchen Lin, Khyathi Chandu,
 649 Nouha Dziri, Sachin Kumar, Tom Zick, Yejin Choi, Noah A Smith, and Hannaneh Hajishirzi.
 650 RewardBench: Evaluating reward models for language modeling. *arXiv [cs.LG]*, March 2024.
 651

652 Yeonji Lee, Sangjun Park, Kyunghyun Cho, and Jinyeong Bak. MentalAgora: A gateway to advanced
 653 personalized care in mental health through multi-agent debating and attribute control. *arXiv*
 654 [*cs.CL*], July 2024.

655 Chris Yuhao Liu, Liang Zeng, Jiacai Liu, Rui Yan, Jujie He, Chaojie Wang, Shuicheng Yan, Yang
 656 Liu, and Yahui Zhou. Skywork-reward: Bag of tricks for reward modeling in LLMs. *arXiv [cs.AI]*,
 657 October 2024.

658 June M Liu, Donghao Li, He Cao, Tianhe Ren, Zeyi Liao, and Jiamin Wu. ChatCounselor: A large
 659 language models for mental health support. *arXiv [cs.CL]*, September 2023.
 660

661 Llama Team. The llama 3 herd of models. *arXiv [cs.AI]*, July 2024.

662 Mounica Maddela, Megan Ung, Jing Xu, Andrea Madotto, Heather Foran, and Y-Lan Boureau.
 663 Training models to generate, recognize, and reframe unhelpful thoughts. In *Proceedings of the*
 664 *61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*,
 665 Stroudsburg, PA, USA, 2023. Association for Computational Linguistics.
 666

667 Luke Marks, Amir Abdullah, Clement Neo, Rauno Arike, Philip Torr, and Fazl Barez. Beyond
 668 training objectives: Interpreting reward model divergence in large language models. *arXiv [cs.LG]*,
 669 October 2023.

670 Miller and William Stephen Rollnick. *Motivational interviewing: Helping people change*.
 671

672 Niklas Muennighoff, Luca Soldaini, Dirk Groeneveld, Kyle Lo, Jacob Morrison, Sewon Min, Weijia
 673 Shi, Pete Walsh, Oyvind Tafjord, Nathan Lambert, Yuling Gu, Shane Arora, Akshita Bhagia,
 674 Dustin Schwenk, David Wadden, Alexander Wettig, Binyuan Hui, Tim Dettmers, Douwe Kiela,
 675 Ali Farhadi, Noah A Smith, Pang Wei Koh, Amanpreet Singh, and Hannaneh Hajishirzi. OLMoE:
 676 Open mixture-of-experts language models. *arXiv [cs.CL]*, September 2024.
 677

677 Hongbin Na, Yining Hua, Zimu Wang, Tao Shen, Beibei Yu, Lilin Wang, Wei Wang, John Torous,
 678 and Ling Chen. A survey of large language models in psychotherapy: Current landscape and future
 679 directions. *arXiv [cs.CL]*, February 2025.

680 Mahdi Pakdaman Naeini, Gregory F Cooper, and Milos Hauskrecht. Obtaining well calibrated
 681 probabilities using bayesian binning. *Proc. Conf. AAAI Artif. Intell.*, 2015:2901–2907, January
 682 2015.

683 OpenAI. GPT-4o system card. *arXiv [cs.CL]*, October 2024.

684 Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L Wainwright, Pamela Mishkin, Chong
 685 Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton,
 686 Luke E Miller, Maddie Simens, Amanda Askell, P Welinder, P Christiano, J Leike, and Ryan J
 687 Lowe. Training language models to follow instructions with human feedback. *Adv. Neural Inf.*
 688 *Process. Syst.*, abs/2203.02155, March 2022.

689 Richard Yuanzhe Pang, Weizhe Yuan, Kyunghyun Cho, He He, Sainbayar Sukhbaatar, and Jason
 690 Weston. Iterative reasoning preference optimization. *arXiv [cs.CL]*, April 2024.

691

692 Qwen, An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan
 693 Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang,
 694 Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin
 695 Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi
 696 Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan,
 697 Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. Qwen2.5 technical report. *arXiv [cs.CL]*,
 698 December 2024.

699 Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D Manning, and Chelsea
 700 Finn. Direct preference optimization: Your language model is secretly a reward model. *arXiv*
 701 [*cs.LG*], May 2023.

702 Rafael Rafailov, Yaswanth Chittepu, Ryan Park, Harshit Sikchi, Joey Hejna, Bradley Knox, Chelsea
 703 Finn, and Scott Niekum. Scaling laws for reward model overoptimization in direct alignment
 704 algorithms, 2024.

705

706 Diego Rocco, Alessandro Gennaro, Lorena Filugelli, Patrizia Squarcina, and Elena Antonelli. Key
 707 factors in psychotherapy training: an analysis of trainers', trainees' and psychotherapists' points of
 708 view. *Res. Psychother. Psychopathol. Process Outcome*, 22(3), December 2019.

709

710 Glenda Dewberry Rooney, Ronald H Rooney, Dean H Hepworth, and Kim Strom-Gottfried. *Direct
 711 social work practice: Theory and skills*. Cengage Learning, 2017.

712

713 John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy
 714 optimization algorithms. *arXiv [cs.LG]*, July 2017.

715

716 Ashish Sharma, Kevin Rushton, Inna Wanyin Lin, Theresa Nguyen, and Tim Althoff. Facilitating
 717 self-guided mental health interventions through human-language model interaction: A case study
 718 of cognitive restructuring. In *Proceedings of the CHI Conference on Human Factors in Computing
 719 Systems*, volume 21, pp. 1–29, New York, NY, USA, May 2024. ACM.

720

721 Sagarika Shreevastava and Peter Foltz. Detecting cognitive distortions from patient-therapist in-
 722 teractions. In *Proceedings of the Seventh Workshop on Computational Linguistics and Clinical
 723 Psychology: Improving Access*, Stroudsburg, PA, USA, 2021. Association for Computational
 724 Linguistics.

725

726 Elizabeth C Stade, Shannon Wiltsey Stirman, Lyle H Ungar, Cody L Boland, H Andrew Schwartz,
 727 David B Yaden, João Sedoc, Robert J DeRubeis, Robb Willer, and Johannes C Eichstaedt. Large
 728 language models could change the future of behavioral healthcare: a proposal for responsible
 729 development and evaluation. *Npj Ment Health Res*, 3(1):12, April 2024.

730

731 Yunhao Tang, Daniel Zhaoan Guo, Zeyu Zheng, Daniele Calandriello, Yuan Cao, Eugene Tarassov,
 732 Rémi Munos, Bernardo Ávila Pires, Michal Valko, Yong Cheng, and Will Dabney. Understanding
 733 the performance gap between online and offline alignment algorithms. *arXiv [cs.LG]*, May 2024.

734

735 Kristine D Tower. Consumer-centered social work practice: Restoring client self-determination.
 736 *Social Work*, 39(2):191–196, 1994.

737

738 Zhilin Wang, Alexander Bukharin, Olivier Delalleau, Daniel Egert, Gerald Shen, Jiaqi Zeng, Oleksii
 739 Kuchaiiev, and Yi Dong. HelpSteer2-preference: Complementing ratings with preferences. *arXiv
 740 [cs.LG]*, October 2024.

741

742 Shusheng Xu, Wei Fu, Jiaxuan Gao, Wenjie Ye, Weilin Liu, Zhiyu Mei, Guangju Wang, Chao Yu,
 743 and Yi Wu. Is DPO superior to PPO for LLM alignment? a comprehensive study. *arXiv [cs.CL]*,
 744 April 2024.

745

746 Aiyuan Yang, Bin Xiao, Bingning Wang, Borong Zhang, Ce Bian, Chao Yin, Chenxu Lv, Da Pan,
 747 Dian Wang, Dong Yan, Fan Yang, Fei Deng, Feng Wang, Feng Liu, Guangwei Ai, Guosheng Dong,
 748 Haizhou Zhao, Hang Xu, Haoze Sun, Hongda Zhang, Hui Liu, Jiaming Ji, Jian Xie, Juntao Dai,
 749 Kun Fang, Lei Su, Liang Song, Lifeng Liu, Liyun Ru, Luyao Ma, Mang Wang, Mickel Liu, Mingan
 750 Lin, Nuolan Nie, Peidong Guo, Ruiyang Sun, Tao Zhang, Tianpeng Li, Tianyu Li, Wei Cheng,
 751 Weipeng Chen, Xiangrong Zeng, Xiaochuan Wang, Xiaoxi Chen, Xin Men, Xin Yu, Xuehai Pan,
 752 Yanjun Shen, Yiding Wang, Yiyu Li, Youxin Jiang, Yuchen Gao, Yupeng Zhang, Zenan Zhou, and
 753 Zhiying Wu. Baichuan 2: Open large-scale language models. *arXiv [cs.CL]*, September 2023.

754

755 An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li,
 756 Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang,
 757 Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, Jianxin Yang, Jin Xu, Jingren
 758 Zhou, Jinze Bai, Jinzheng He, Junyang Lin, Kai Dang, Keming Lu, Keqin Chen, Kexin Yang,
 759 Mei Li, Mingfeng Xue, Na Ni, Pei Zhang, Peng Wang, Ru Peng, Rui Men, Ruize Gao, Runji Lin,
 760 Shijie Wang, Shuai Bai, Sinan Tan, Tianhang Zhu, Tianhao Li, Tianyu Liu, Wenbin Ge, Xiaodong
 761 Deng, Xiaohuan Zhou, Xingzhang Ren, Xinyu Zhang, Xipin Wei, Xuancheng Ren, Xuejing Liu,
 762 Yang Fan, Yang Yao, Yichang Zhang, Yu Wan, Yunfei Chu, Yuqiong Liu, Zeyu Cui, Zhenru Zhang,
 763 Zhifang Guo, and Zhihao Fan. Qwen2 technical report. *arXiv [cs.CL]*, July 2024.

756 Mian Zhang, Xianjun Yang, Xinlu Zhang, Travis Labrum, Jamie C Chiu, Shaun M Eack, Fei Fang,
757 William Yang Wang, and Zhiyu Zoey Chen. CBT-bench: Evaluating large language models on
758 assisting cognitive behavior therapy. *arXiv [cs.CL]*, October 2024.

759
760 Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang,
761 Zi Lin, Zhuohan Li, Dacheng Li, Eric P Xing, Hao Zhang, Joseph E Gonzalez, and Ion Stoica.
762 Judging LLM-as-a-judge with MT-bench and chatbot arena. *arXiv [cs.CL]*, June 2023.

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810 7 PROMPTS
811812 **Rating Prompt**
813814 You are provided with a client speech and four responses from different psychotherapists. Rate
815 the responses based on how they align with the given principle.
816817 Client Speech: {client_speech}
818 Response 1: {response1}
819 Response 2: {response2}
820 Response 3: {response3}
821 Response 4: {response4}822 Provide a JSON object as output that includes the following keys:
823824

- response_1_rating: An integer score from 1 to 5 for response 1
- rationale_1: A string explaining the reasoning behind the given score for response 1
- response_2_rating: An integer score from 1 to 5 for response 2
- rationale_2: A string explaining the reasoning behind the given score for response 2
- response_3_rating: An integer score from 1 to 5 for response 3
- rationale_3: A string explaining the reasoning behind the given score for response 3
- response_4_rating: An integer score from 1 to 5 for response 4
- rationale_4: A string explaining the reasoning behind the given score for response 4

834 **Responding Prompt**
835836 You are now a professional psychotherapist conducting a session with a client. Answer the given
837 client speech.
838

839 Client Speech: {client_speech}

840 **LLM-as-Ranker Prompt**
841842 Determine which of the two given responses from different psychotherapists to a client's speech
843 is better:
844845 Client Speech: {client_speech}
846 Response 1: {response_1}
847 Response 2: {response_2}848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863

864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
8 DATASET INFORMATION**PsychoCounsel Principles**

Empathy and Emotional Understanding: The response should convey genuine empathy, acknowledging and validating the client's feelings and experiences.

- Emotional Reflection: Reflecting the client's emotions back to them.
- Validation: Affirming the client's feelings as legitimate and understandable.
- Non-Judgmental Tone: Maintaining a compassionate and accepting approach.

Personalization and Relevance: The response should be tailored to the client's unique situation, ensuring that the content is directly relevant to their concerns.

- Specific References: Mentioning details specific to the client's statements.
- Avoidance of Generic Responses: Steering clear of overly general or canned replies.
- Cultural and Individual Sensitivity: Respecting the client's background and personal context.

Facilitation of Self-Exploration: The response should encourage the client to reflect on their thoughts and feelings, promoting self-awareness and insight.

- Open-Ended Questions: Asking questions that invite elaboration.
- Reflective Statements: Paraphrasing the client's words to deepen understanding.
- Exploration of Thoughts and Feelings: Guiding the client to consider underlying emotions and beliefs.

Clarity and Conciseness: The response should be clear, well-organized, and free of unnecessary jargon, making it easy for the client to understand and engage with.

- Plain Language: Using words that are easily understood.
- Logical Flow: Presenting ideas in a coherent sequence.
- Brevity: Keeping the response concise while covering essential points.

Promotion of Autonomy and Confidence: The response should support the client's sense of control over their decisions and encourage confidence in their ability to make positive changes.

- Affirmation of Strengths: Highlighting the client's abilities and past successes.
- Encouraging Initiative: Motivating the client to take proactive steps.

Avoidance of Harmful Language or Content: The response should avoid any language or content that could potentially harm, distress, or trigger the client, ensuring the interaction is safe and supportive.

Sensitivity to the Stage of Change: The response should recognize the client's current stage in the process of change and address their needs accordingly. If the client is in an early stage—uncertain or ambivalent about making a change—the response should help them explore their thoughts and motivations. If the client is in a later stage and has already made changes, the response should focus on reinforcing progress, preventing setbacks, and sustaining positive

⁹<https://mistral.ai/news/ministraux/>

¹⁰<https://mistral.ai/news/mistral-nemo/>

¹¹<https://openai.com/index/openai-01-system-card/>

¹²<https://docs.cohere.com/docs/responsible-use>

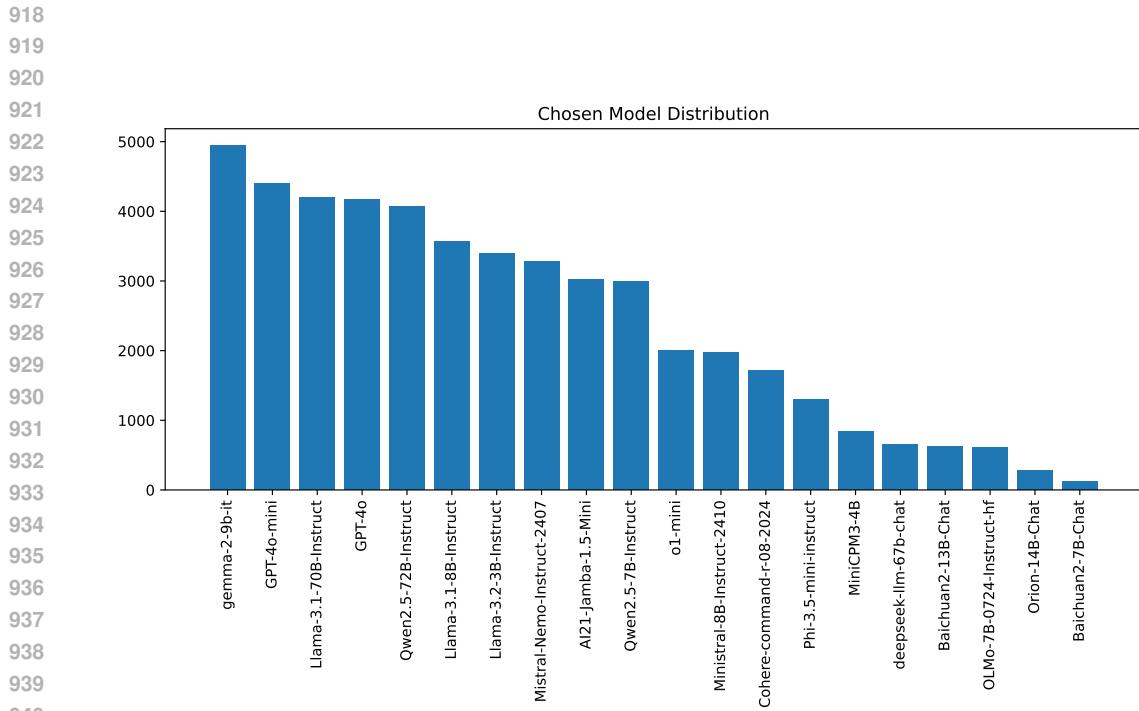


Figure 5: Chosen Model Distribution

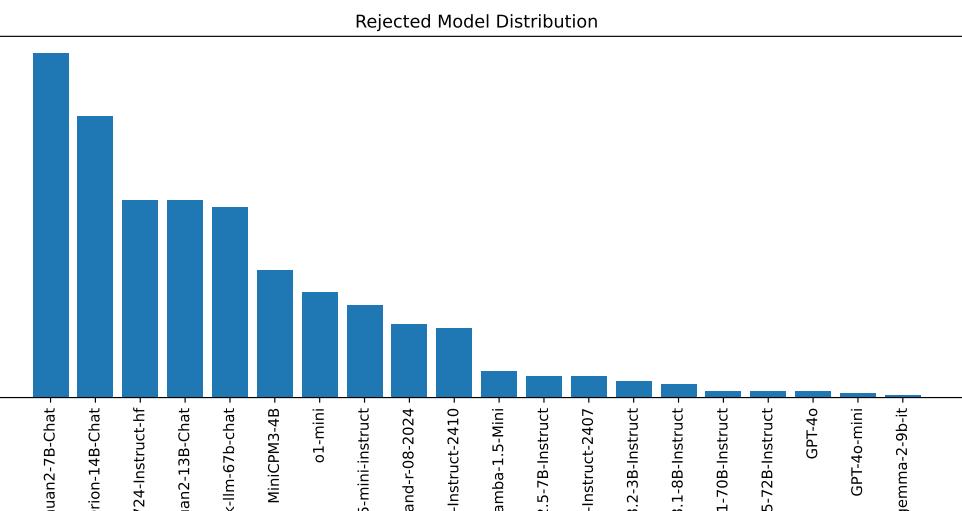


Figure 6: Rejected Model Distribution

Table 5: Topic Distribution

Coarse Category	Fine Category	Count
1. Core Mental Health Issues		
	Anxiety	3714
	Depression	2859
	Stress	1439
	Trauma	526
	Substance-abuse	387
	Addiction	129
2. Emotional Well-being and Coping Strategies		
	Self-esteem	1377
	Grief-and-loss	1023
	Caregiving	1541
	Behavioral-change	740
	Anger-management	448
	Self-care	311
	Sleep-improvement	277
3. Relationships and Interpersonal Dynamics		
	Relationships	1690
	Family-conflict	2358
	Friendship-conflict	292
	Marriage	373
	Intimacy	403
	Social-relationships	410
	Workplace-relationships	383
	Relationship-dissolution	574
4. Life Transitions and Challenges		
	Career	441
	Aging	140
	New-environment	235
	Military-issues	118
5. Social Issues		
	LGBTQ	335
	Culture	113
	Human-sexuality	151
	Bullying	68
6. Youth and Development		
	Children-adolescents	123
	School-life	322
	Parenting	730
7. Crisis and Safety Concerns		
	Domestic-violence	144
	Self-harm	231
	Eating-disorders	154
8. Special Topics		
	Counseling-fundamentals	638
	Diagnosis	531
	Communication	205
	Professional-ethics	128
	Legal-regulatory	94
	Spirituality	192
	Others	136

1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042

Category	Models
3-4B models	Llama-3.2-3B-Instruct (Llama Team, 2024) Phi-3.5-mini-instruct (Abdin et al., 2024) MiniCPM3-4B (Hu et al., 2024)
7-9B models	Minstral-8B-Instruct-2410 ⁹ Llama-3.1-8B-Instruct (Llama Team, 2024) gemma-2-9b-it (Gemma Team, 2024) Qwen2.5-7B-Instruct (Qwen et al., 2024) OLMo-7B-0724-Instruct (Muennighoff et al., 2024)
12-14B models	Baichuan2-7B-Chat (Yang et al., 2023) Baichuan2-13B-Chat (Yang et al., 2023) Orion-14B-Chat (Chen et al., 2024) Mistral-Nemo-Instruct-2407 ¹⁰ AI21-Jamba-1.5-Mini (Jamba Team et al., 2024)
65-75B models	Llama-3.1-70B-Instruct (Llama Team, 2024) Qwen2.5-72B-Instruct (Qwen et al., 2024) deepseek-llm-67b-chat (DeepSeek-AI et al., 2024)
Commercial models	GPT-4o (OpenAI, 2024) GPT-4o-mini (OpenAI, 2024) o1-mini ¹¹ Cohere-command-r-08-2024 ¹²

Table 6: Overview of selected models in the pool.

1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079

1080 9 CASE STUDY
10811082 In the example in Table 7, PsychoCounsel-Llama3-8B demonstrates a deeper level of empathy and
1083 personalization. It acknowledges the client’s unique emotional experiences and suggests tailored
1084 strategies, such as the “3-4-5” exercise, to directly address their concerns. Additionally, it facilitates
1085 self-exploration by asking thoughtful questions about the client’s fears, values, and understanding of
1086 self-trust, ultimately promoting autonomy and confidence in decision-making. In contrast, GPT-4o’s
1087 response feels more clinical and general.
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133

1134

1135

1136

1137

1138

Client Speech

1139

1140

1141

Anxiety often strikes when I'm faced with making decisions. The fear of making the wrong choice or disappointing others paralyzes me, leaving me stuck in indecision. I want to learn how to trust myself and make confident choices.

1142

1143

PsychoCounsel-Llama3-8B (chosen)

1144

1145

It takes a lot of courage to share something so vulnerable with me, and I appreciate your trust in our conversation. I can sense the weight of that anxiety you're carrying, and we're going to explore ways to shift that together.

1146

1147

First, let's acknowledge that the fear of making the wrong choice or disappointing others is a common and normal feeling. It doesn't mean there's anything inherently wrong with you; it just means you care deeply about others and yourself. That's a beautiful quality to have.

1148

1149

1150

Can you tell me more about what you think would happen if you made a "wrong" choice? What are some of the worst-case scenarios that come to mind? Sometimes, exploring these fears can help us gain clarity and perspective. We can work together to reframe those thoughts in a more realistic and compassionate light.

1151

1152

Let's also explore the idea of self-trust. What does it mean to you to trust yourself? What does self-trust look like in action? We can work on building that trust by identifying what you value and what aligns with your goals. We can practice recognizing when your thoughts are trying to control or second-guess you, and we can develop strategies to gently acknowledge those thoughts without letting them overwhelm you.

1153

1154

I'd like to introduce an exercise called "3-4-5." When you're faced with a difficult decision, I want you to take a deep breath and pause for 3 seconds, acknowledge the thought or feeling that arises, and then ask yourself: "Is this thought based on facts or perceived facts?" Sometimes, when we're anxious, our thoughts can get tangled up in assumptions or past experiences. Then, take 4 seconds to ask yourself what you truly value in this situation, what's most important to you. And finally, take 5 seconds to imagine yourself moving forward, taking action, and making a choice that aligns with your values. This exercise can help you tap into your inner wisdom and develop a sense of self-trust.

1155

1156

Lastly, let's work on reframing the idea of "wrong" choices. What if, instead of focusing on the fear of making the wrong choice, we reframe it as a learning opportunity? What if you saw every choice as a chance to learn, grow, and develop? How might that change your approach to decision-making?

1157

1158

Remember, our goal isn't to eliminate anxiety completely, but to develop a more compassionate relationship with it. We'll work together to cultivate self-awareness, self-acceptance, and trust in your decision-making abilities.

1159

How do these ideas resonate with you? What feels most helpful or interesting to explore further?

1160

GPT-4o (rejected)

1161

1162

Thank you for sharing that with me. It's understandable that making decisions can be overwhelming, especially when there's fear of making the wrong choice or disappointing others. Let's explore this together. Firstly, it's important to acknowledge that decision-making can be challenging and it's completely normal to feel anxious about it. You're not alone in experiencing this.

1163

1164

One way we might start working on this is by exploring some of the underlying thoughts or beliefs that contribute to your anxiety about decision-making. For instance, you mentioned a fear of disappointing others. Can you tell me more about where you think that fear comes from?

1165

1166

We could also look at ways to build your confidence in your decision-making skills. Sometimes breaking down decisions into smaller, more manageable parts can make the process feel less daunting. Have you ever tried writing down the pros and cons of each option or talking it through with someone you trust?

1167

1168

Another technique is to practice self-compassion. Remind yourself that no one makes the right choice all the time and that it's okay to make mistakes. Every decision is a learning opportunity.

1169

1170

Finally, something that can be quite helpful is examining past decisions you've made. Often, we are harsher with ourselves than necessary and we overlook our strengths. Can you recall a situation in the past where you made a decision and things turned out well, or you learned something valuable from the outcome?

1171

1172

Let's take some time to explore these aspects and see what might resonate most with you. How does that sound?

1173

1174

1175

1176

1177

1178

1179

1180

1181

1182

Table 7: Case Study - 2