

# 心理咨询无法一蹴而就：通过多次会话实现自动心理咨询

Bichen Wang\*, Junzhe Wang\*, Xing Fu, Yixin Sun, Bing Qin, Yanyan Zhao†

Research Center for Social Computing and Information Retrieval

Harbin Institute of Technology, China

junzhewang, bichenwang@ir.hit.edu.cn

## Abstract

近年来,大型语言模型 (LLMs) 在自动心理咨询方面取得了显著进展。然而,目前的研究集中于单次会话咨询,这并不能代表现实世界的场景。在实践中,心理咨询是一个过程,而不是一次性事件,需要持续的、多次会话的参与以逐步解决客户的问题。为了克服这一限制,我们引入了一个多会话心理咨询对话数据集 (MusPsy-Dataset)。我们的 MusPsy-Dataset 是使用公开可用的心理案例报告中的真实客户档案构建的。它捕捉了咨询的动态过程,包括同一客户在不同会话中进行的多次渐进式咨询对话。利用我们的数据集,我们还开发了 MusPsy 模型,其旨在跟踪客户进展,并随着时间的推移调整其咨询方向。实验表明,我们的模型在多个会话中表现优于基线模型。

## 1 介绍

在当今社会中,个人面临着日益增加的心理压力 (Organization, 2024), 由于现代生活的快节奏 (Samji et al., 2022), 对心理健康支持的需求不断增加。因此,可获得心理健康服务的稀缺性 (Grant et al., 2018) 驱使研究人员探索计算技术以实现自动化心理咨询。

先前的研究促进了心理学理论建模及各种心理咨询对话的创造,并取得了积极的成果 (Lee et al., 2024; Qiu et al., 2024; Na, 2024)。然而,这些研究受到其单次会话模式的限制。与随意聊天不同,诸如认知行为疗法 (CBT) 等心理咨询通常需要跨多个会话进行 (Craske, 2010)。如图 1 所示,来访者通过咨询师引导的多次会话经历逐渐的心理变化。在这些会话期间,咨询师充当导航者,根据客户不断变化的动态心理状态动态重新校准咨询目标 (Baur et al., 2024; Dobson and Dozois, 2021)。这强调了开发能够维持连贯的多会话咨询对话的自动化模型的重要性。

\* Equal contribution.

† Corresponding author.

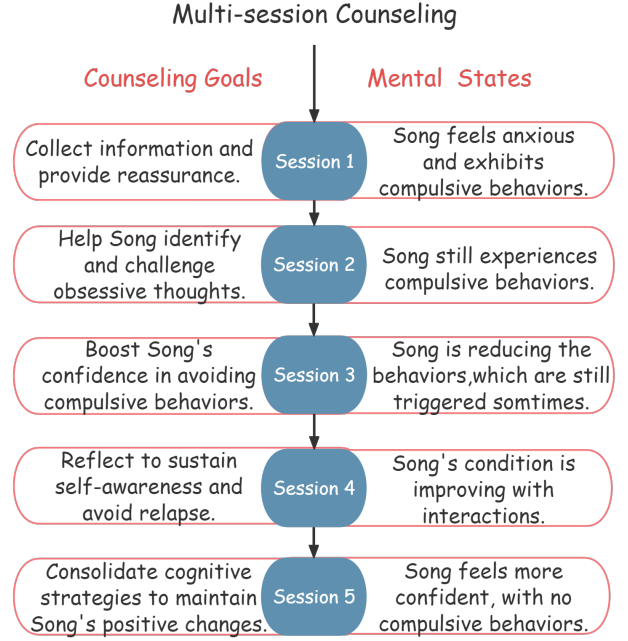


Figure 1: 多次会谈辅导的一个真实例子涉及一个客户在五次会谈中展示出心理改善。每次会谈都有特定目标,并且客户的心理状态逐渐改善,这种结果在单次会谈的情况下是无法实现的。

为了应对这些限制,我们的研究将多会话范式整合到自动心理辅导框架中。与之前构建孤立会话的研究不同,我们的工作试图对多个会话进行建模,以捕捉辅导过程。这种方法考虑了客户动态变化的状态,并允许对辅导员目标进行持续调整。因此,我们可以累积建立客户关系,解决客户问题,并支持长期的心理改善。

总的来说,我们研发了 MusPsy-数据集,这是一个基于 LLM 的 CBT 的多会话心理咨询对话数据集。为了确保真实性和质量,我们从公开的心理咨询案例报告中收集同一客户的多次会话的客户资料和咨询目标,以生成多会话咨询。由于 LLMs 在一次传递中难以生成连贯的多会话咨询内容,我们采用了自上而下的方法:首先,我们生成多个简短的种子对话以指导多个对话的流程;然后,我们将这些种子对话展开为跨会话的完整、一致且以目标为导向

的长咨询对话。在根据种子对话生成这些完整的会话时，我们通过结合案例报告中的咨询片段和使用精心设计的提示进行增强。为了模拟人类咨询师的工作流程并应对长情境对话的挑战，我们设计了一个过程，从咨询会话中提取会话级别的记忆。

我们使用 MusPsy 数据集训练一个自动化心理咨询模型，以进行多次会话的咨询。我们通过多个心理指标验证我们模型的效果，包括基本指标、工作联盟量表（WAI）以及积极和消极情感量表（PANAS）。

- 我们引入了一种将多次会话整合到自动辅导中的新方法，从而能够对辅导进展进行更长期的建模。
- 我们引入了 MusPsy-Dataset，这是一个基于认知行为疗法（CBT）的新型多次会话心理咨询对话数据集。
- 我们开发了一种由 MusPsy-Dataset 训练的多次咨询模型，MusPsy-Model，展示了其在多种心理指标上的能力。

## 2 相关工作

### 2.1 认知行为疗法

认知行为疗法（CBT）长期以来被认为是为抑郁和焦虑患者提供有效干预的一种方法（Beck, 2020）。先前的研究描述了患有这些疾病的人常常会产生消极的、非理性的想法，这些想法强化了关于他们自己、他人和世界的有害信念。为了打破这个循环，CBT 集中在识别和挑战这些自动化思维 and 核心信念（Longmore and Worrell, 2007）。CBT 是一个结构化的、多次课程的过程，不能在单个课程中完成（Hayes and Hofmann, 2018）。在 CBT 课程中，辅导员首先帮助客户识别无益的想法，然后通过使用各种 CBT 技术引导客户挑战和纠正这些扭曲，从而最终帮助重建更加积极的自动化思维和信念（Fenn and Byrne, 2013）。这个辅导过程对于促进心理健康和福祉是至关重要的。

### 2.2 自动化心理咨询

Eliza，这个开创性的系统，采用了一种基于规则的方法进行罗杰斯式咨询（Rogers, 1952; Weizenbaum, 1966）。然而，关于自动心理咨询的数据驱动研究的进展由于数据的有限可用性而受阻。由于隐私问题和对弱势群体的保护，创建公共的、真实的、大规模的咨询数据集仍然是不可能的。然而，LLM 的出现使得生成用于咨询数据集的合成数据成为可能。例如，SMILE 数据集（Qiu et al., 2024），是对 PsyQA

的改进（Sun et al., 2021），它包括多轮的咨询对话，并被 CBT-LLM 使用（Qiu et al., 2024）。CPsyCoun 从公共心理学报告中提取咨询数据，而不需要特定领域的专业知识（Zhang et al., 2024）。Healme（Xiao et al., 2024）专注于通过 LLM 优化 CBT 指导，而 CACTUS（Lee et al., 2024）强调构建更完整的单轮 CBT 流程。

与这些努力相比，我们的工作将多会话动态引入到自动化心理咨询中。我们引入了一个新颖的数据集和模型。

## 3 MusPsy-数据集：一个多会话心理咨询对话数据集

我们介绍了我们的 MusPsy 数据集的构建过程。如图 2 所示，我们的数据集构建涉及四个关键阶段：客户档案 & 咨询目标收集、种子对话构建、咨询生成和记忆生成。

### 3.1 客户档案 & 咨询目标收集

为了支持多场次自动化咨询，构建真实且具备上下文感知的客户档案是至关重要的。虽然以往的研究往往依赖于合成的客户特征，仅专注于静态特征，如年龄、性别和提出的问题，但它们通常缺乏客户在多场次中的心理状态的动态演变，导致对长期咨询模拟的支持有限。

在我们的工作中，我们通过收集多会话咨询客户档案来解决这个问题，这些档案来自发表在学术期刊或专业书籍中可信的心理案例报告，确保真实性和连续性。尽管出于隐私原因这些报告不包含完整的会话记录，但它们记录了客户的状态并实现了每次会话的咨询目标。与只能捕捉最初背景的以往方法不同，我们的档案不仅涵盖了静态特征和动态状态，还包括咨询师在每次会话前设定的咨询目标。

- 静态特征：这些包括化名、性别、年龄、职业以及在初次咨询期间报告的心理问题。
- 动态状态：这些状态捕捉到每次会话的变化，包括客户最近的生活事件以及情绪和认知状态的变化。
- 咨询目标：咨询目标定义了咨询师预期会议的具体方向和内容。

客户档案由静态特征和动态状态构成。咨询目标不包括在客户档案中，因为它们通常涉及的不仅仅是客户。图 3 展示了典型的多次会话 CBT 目标和方法，定义了咨询师预期的会话方向和内容。

我们使用带有人类验证的 GPT-4o 从病例报告中提取结构化信息。有关客户资料和咨询目标收集的详细信息，请参见附录 A。

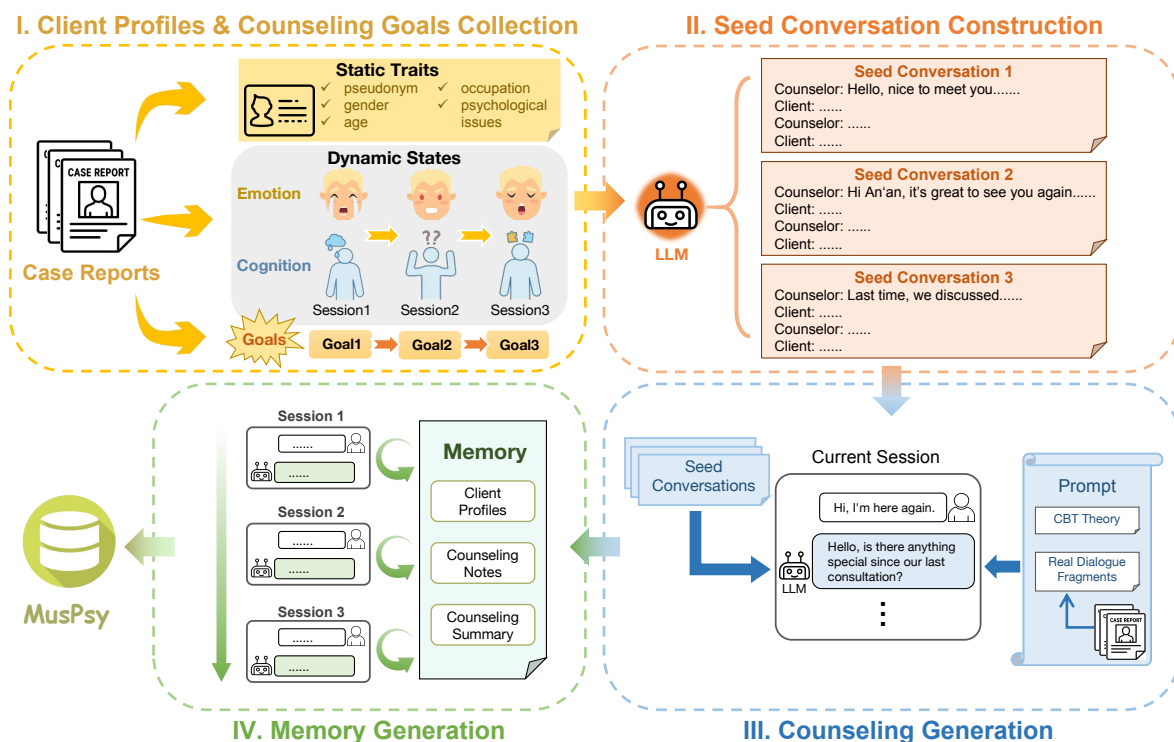


Figure 2: 该图展示了构建 MusPsy-Dataset 的过程。它包括四个部分：客户端资料 & 辅导目标收集、种子对话构建、辅导生成和记忆生成。通过这四个步骤，我们得到了一个高质量的 MusPsy-Dataset。

### 3.2 种子对话构建

直接一次性生成所有完整的会话受限于大型语言模型的上下文和指令跟随限制，因此深度有限。相反，分别生成会话往往会产生重复的措辞和较弱的连贯性，因为大型语言模型缺乏对咨询进展的意识。这凸显了需要不同的生成方法，以确保整个咨询会话的连贯性和上下文丰富性。

如图 2 所示，为了确保多会话互动的一致性，我们引入了一个中间步骤：在生成完整的咨询对话之前，构建种子对话。种子对话是指由 LLM 同时生成的简短而精炼的多会话对话，尽管其篇幅简短，但由于同时生成，因此具有高度的一致性。

为了确保连贯性，我们将多个客户档案及其相应的目标一起输入。我们指示 LLM 生成连贯的初始对话，每个对话仅包含 3 到 4 轮，并且大约 100-200 个符合会话目标的简洁通信令牌。关于初始对话构建的详细信息，请参见附录 B。

### 3.3 咨询生成

在给定初始的多会话情境下，我们的目标是生成完整的、扩展的咨询对话，这些对话在不同会话中保持情境上的连贯性，并在语言上与真实的咨询师相似。

**会话级扩展的上下文约束：** 在生成过程中，为确保情境的连贯性，我们提示 LLM 基于种子对话生成每一个完整的咨询对话，同时提供之前的种子对话。这鼓励模型创建情境上相互联系的对话。这样的设计允许我们在多个会话中保持时间上的依赖关系和情感的进展，确保一个真实且连贯的多会话模拟。

**使用少数镜头引导提示：** 我们通过将核心认知行为疗法（CBT）技术的描述与大型语言模型（LLM）对 CBT 的理解相结合来设计提示，以指导生成过程并扩展种子对话。我们发现一些报告提供了匿名的、不完整的咨询片段。我们使用从这些报告中提取的真实咨询片段作为与相关客户档案阶段对应的示例。这些样本嵌入提示中，以指导 LLM 的语调、结构和咨询技术，从而产生更加真实且在上下文中适宜的输出。

我们采用一种高效的单代理基于脚本的方法，这种方法显著降低了计算成本和推理时间，同时保证了对话的连贯性和合理性，正如之前的单会话研究所示 (Lee et al., 2024)。在这里，一个大语言模型（LLM）会基于用户资料和初始会话对话在单次通过中生成整个多轮对话。有关咨询生成的详细信息，请参见附录 C。

此外，长期的多次会话咨询可能很容易超过



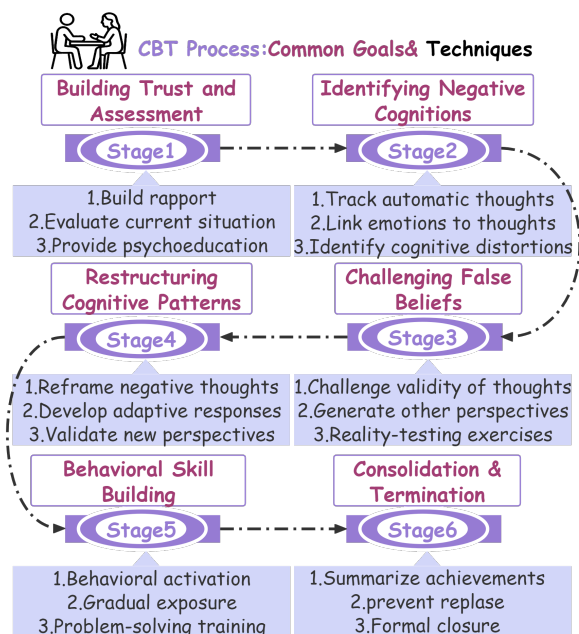


Figure 3: 多次会谈的认知行为疗法 (CBT) 通常遵循一个模式，其中辅导员通常坚持这一过程，并将其与客户的具体情况相结合，以设定更具体和可操作的辅导目标，并使用辅导技术 (Beck et al., 2011)。

可用的上下文窗口，并难以保持一致性。为了支持一致且具备上下文意识的多次会话咨询，我们为每个客户构建结构化的记忆表述。

与普通聊天记忆不同，咨询记忆的设计需要反映临床要求。为应对这一挑战，我们引入了一个外部记忆模块，以捕捉关键的会话级信息，使我们能够在长时间的咨询中保持连贯性和连续性。受 (Bemister and Dobson, 2011; Lawlor-Savage and Prentice, 2014) 认知行为治疗实践中的文档标准的启发，我们设计了一种专为多会话咨询而量身定制的多部分记忆结构。

- 客户档案：包括客户档案和个人特征。
- 咨询记录：记录关键的会话层级信息，包括咨询师的观察、会话目标以及在多次咨询会话中的任务。
- 咨询总结：整个系列会话的简要总结。

在每次咨询会谈结束时，我们将客户的谈话总结到这个记忆结构中。这种方法使模型能够更新并保留关键信息，而无需存储完整的咨询记录。结果是，后续会谈能够保持连续性，并模拟咨询师的跨会话跟踪和计划能力。有关记忆生成的更多细节，请参阅附录 D。

## 4 数据统计与评估

### 4.1 数据统计

我们在表 1 中提供了我们数据集的基本统计概览。MusPsy-Dataset 由 1,400 个基于 CBT 理论构建的多回合辅导对话组成，每次会话平均 28.55 个回合，每回合平均 26.99 个词。平均而言，每个客户有 6.17 次辅导会话。这个数据集的规模足以支持我们新多会话辅导模型的构建。

Item	Value
Client Profiles for Training	1,400
Client Profiles for Testing	100
Average Sessions Per Client	6.17
Average Turns Per Session	28.55
Average Words per Turn	26.99
Average Tokens Per Client	5693.27
Average Tokens Per Turn	33.23

Table 1: Data statistics of our MusPsy-Dataset

### 4.2 竞赛数据集

我们介绍了几项用于比较的先前工作，包括 SMILE 数据集 (Qiu et al., 2024)，这是一种基于 PsyQA 的多轮次心理健康对话数据集；以及 CACTUS 数据集 (Lee et al., 2024)，另一种用于心理健康的多轮次 CBT 数据集。我们还包括 SimPsyDial (Qiu and Lan, 2024) 和 CPsyCoun (Zhang et al., 2024) 数据集。我们在我们的评估框架中评估这些研究中报告的数据质量和模型性能。这些工作在构建心理健康咨询数据集和提供心理支持模型方面与我们相似。

### 4.3 数据集质量评估

我们随机选择每个数据集的 100 个咨询示例。两位心理学专家进行人工评估，GPT-4o 则进行自动评估。我们专注于四个基本指标和咨询师-来访者工作联盟，并使用它作为跨不同咨询数据的一致性衡量标准 (Horvath, 2001)。研究表明，工作联盟能预测咨询的结果 (Horvath and Symonds, 1991; Horvath et al., 2011)。所有指标的评分为 1-5 分。

我们使用基本评价指标 (Munder et al., 2010) 来对四个维度评估我们的数据集：有帮助性：咨询师解释和信息的实用性；同理心：咨询师理解和分享来访者感受的能力；指导性：实际建议的可用性和具体性；连贯性：对话中的逻辑一致性。我们还使用 12 项工作联盟量表 (WAI) (Munder et al., 2010)，评估三个维度：目标一致性：对我们咨询目标的共同认可；任务

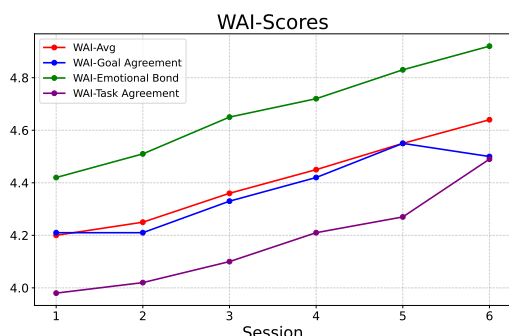


Figure 4: 在多次会话中，每次辅导会话的 WAI 评分评估。观察到总体呈上升趋势。

一致性：对所用方法的一致性；情感纽带：相互之间的信任、信心和喜好。WAI 中的每个维度通过三个问题进行评估，我们报告每个维度的平均分数。有关评估的详细信息，请参见附录 E。

为了将数据质量与之前的单次会话数据集进行比较，我们专注于单个会话。对于此评估，我们将同一客户的每个会话视为独立的，并随机抽取每位客户的一个会话。正如表 2 所示，我们的 MusPsy 数据集在几乎所有维度上都优于现有的咨询数据集。

虽然大多数当前的数据集在基本指标上表现良好，但我们建议将评估转向基于心理学理论的评估方法。对于 WAI 指标，仍有改进空间。MusPsy-Dataset 在情感纽带方面表现出色，表明我们的数据捕捉到了信任、自信 and 好感等元素，这对于现实的治疗体验至关重要。它在任务一致性方面也表现出色，表明我们的咨询更可能获得一致性。我们认为这是因为我们的数据集反映了渐进的进展，而不是期望在单次会话中快速改善。此外，我们的数据在目标一致性方面表现突出，这我们归因于 MusPsy-Dataset 在每次会话中逐步设定目标的方法，有助于提高数据质量。

这些发现确认了我们数据集在训练稳健的咨询模型方面的整体质量和适用性。我们将其归因于对现实世界环境的更准确模拟，在这些环境中，将整个咨询过程整合到一个单一的会话中通常是不现实的。MusPsy-Dataset 使我们能够建立更深的理解和更强的关系，从而提高数据质量。

#### 4.3.1 多会话评估

我们对 MusPsy-Dataset 的数据集进行逐场次分析。对于每位客户，我们最多观察 6 次咨询。如图 4 所示，我们的结果显示，随着咨询的进行，客户和咨询师之间的工作联盟通常会变得更强。然而，在第六次会谈中，我们观察到目

标一致性略有下降。我们发现这种情况发生是因为我们数据集中的一些客户在 5 次咨询左右就结束了。那些继续进行到 6 次咨询的客户有时需要处理更深层次的事情，比如核心信念。这些更深层次的目标常常更难被客户完全接受。

总体而言，我们认为这表明为什么关注多个会话是重要的。更好的自动化心理咨询和支持应当是一个过程，而不是一次性事件。

如图 5 所示，我们定义了三个任务，这些任务反映了一名人类咨询师的多次会议咨询工作流程，以评估我们的数据集在训练 MusPsy-Model 中的实用性：记忆提取、目标规划和咨询生成。我们的目标是使咨询模型能够管理咨询进度并生成连贯的咨询内容。

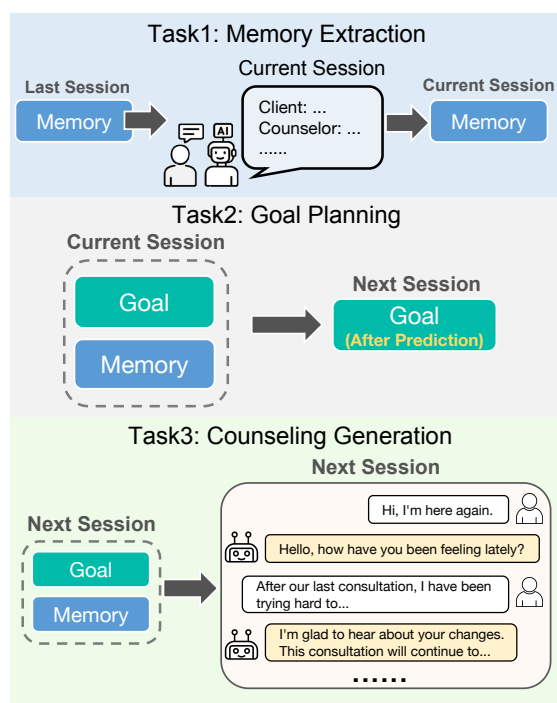


Figure 5: 我们三项任务的概述：说明多次心理咨询的三项任务：记忆提取、目标规划和咨询生成，其中从某次会话中提取的记忆为后续会话的目标规划提供信息，指导下一次咨询会话期间的内容生成。

- 记忆提取（任务 1）：MusPsy-Model 从当前会话中提取关键信息并更新其记忆。
- 目标规划（任务 2）：基于其记忆，MusPsy-Model 为下一个会话计划咨询目标，模拟会前咨询进展管理。
- 咨询生成（任务 3）：MusPsy-Model 在咨询过程中，根据过去的记忆和计划的目标与客户互动。

Evaluator		Hel.	Emp.	Gui.	Coh.	Avg.	Task.	Bond.	Goal.	Avg.
GPT-4o	SMILE	4.05	4.49	3.98	4.43	4.23	3.60	4.35	3.61	3.85
	CACTUS	4.62	4.86	4.56	4.99	4.76	3.56	4.56	3.66	3.92
	SimPsyDial	4.58	4.90	4.51	4.99	4.75	3.77	4.31	3.56	3.97
	CPsyCoun	4.23	4.43	4.09	4.87	4.41	3.56	4.30	3.71	3.86
	MusPsy	4.99	4.98	4.98	4.99	4.98	4.11	4.69	4.44	4.41
Human	SMILE	3.84	3.90	3.57	3.77	3.77	3.13	4.22	3.14	3.49
	CACTUS	4.49	4.59	4.18	4.65	4.47	3.60	4.65	3.49	3.92
	SimPsyDial	3.73	4.56	4.32	4.77	4.34	3.36	4.32	3.58	3.75
	CPsyCoun	3.60	3.83	4.17	3.73	3.83	3.08	4.02	3.43	3.51
	MusPsy	4.89	4.82	4.70	4.88	4.82	4.20	4.72	4.14	4.35

Table 2: 数据质量评估。这包括对四个基本指标和 WAI 量表三个维度的评估，通过 GPT-4o 和人工评估来完成。评估是基于相同的提示翻译成英语后在不同语言中进行的。

我们通过对标注样本进行监督微调同时优化这些任务。在训练过程中，我们为每个任务使用特定的提示，以确保 MusPsy-Model 同时也在所有三个能力上发展专业技能。

## 5 实验

我们的实验使用 Meta-Llama-3-8B-Instruct 模型进行。为了保持公平性和可比性，我们在 Meta-Llama-3-8B-Instruct 上训练其他数据集，而不是使用官方模型。在训练阶段，所有提示和超参数都与原论文中指定的一致。有关实验设置的详细信息，请参见附录。

### 5.1 任务 1 及 2 的评估与结果

我们使用标准机器指标评估记忆提取和目标规划的个体性能。结果表明我们能够根据会话内容有效更新其记忆，并预测后续会话的相关咨询目标。

Metric	BLEU-1	BLEU-2	F1
Task 1	42.7	26.2	38.0
Task 2	44.1	34.0	35.4

Table 3: 任务 1 和任务 2 的评估结果。

这些结果表明，我们的 MusPsy 模型可以学习记忆在咨询会话中讨论的关键信息，并利用这种理解通过提出相关的下一个目标来预测咨询的逻辑进展。这一基础能力对于构建更复杂的咨询模型至关重要，这些模型可以参与到有意义且渐进的多次会话咨询中。

如图 5 所示，对于咨询生成，我们将上一段对话的记忆与当前会话的目标一起提供给大型语言模型作为输入。在说明中，我们强调了这部分信息的重要性。

我们采用一个评估框架，通过多场次咨询模拟来评估模型性能，使用 100 个测试客户档案。在这些模拟中，LLM 充当与咨询师模型互动的客户。除了直接评估咨询对话的指标外，我们还使用积极和消极情绪量表（PANAS）来衡量客户情绪状态的变化，该量表完全从客户的角度提供结果。会话结束后，我们请它评估其 PANAS 得分，以分析咨询模型的有效性。它会更新自身状态并为下一次咨询对话做好准备。

我们不使用 PANAS 来评估数据集，因为它从客户的角度来衡量情绪，如果没有相同的客户，则无法进行比较。对于这些评估，我们使用 GPT-4o。

#### 5.1.1 评估结果

如图 6 所示，我们的 MusPsy-Model 有效地减少了负面情绪，并在长期内增强了正面情绪。我们对客户端 LLM 进行了 6 次会话测试。我们可以观察到，在两次会话之后，其他模型实际上无法继续解决客户端的问题，因为它们一贯生成与前一版本大致相同的辅导。相比之下，MusPsy-Model 的优势尤为明显；在六次会话后，它显著缓解了客户端的负面情绪，与其他方法形成了鲜明对比，突显了我们的优势。

对于大多数咨询模型来说，它们的 LLM 咨询师在多次会话中同样未能获得收益。这意味着这些模型在模拟咨询过程中并未表现出咨询质量的显著提升，这与人类咨询中预期的结果相悖。然而，MusPsy-Model 在其咨询表现中表现出长期的上升趋势，我们将此归因于我们对多会话动态的建模。在多次会话中，它不断加强与客户的情感纽带并增强目标的对齐。



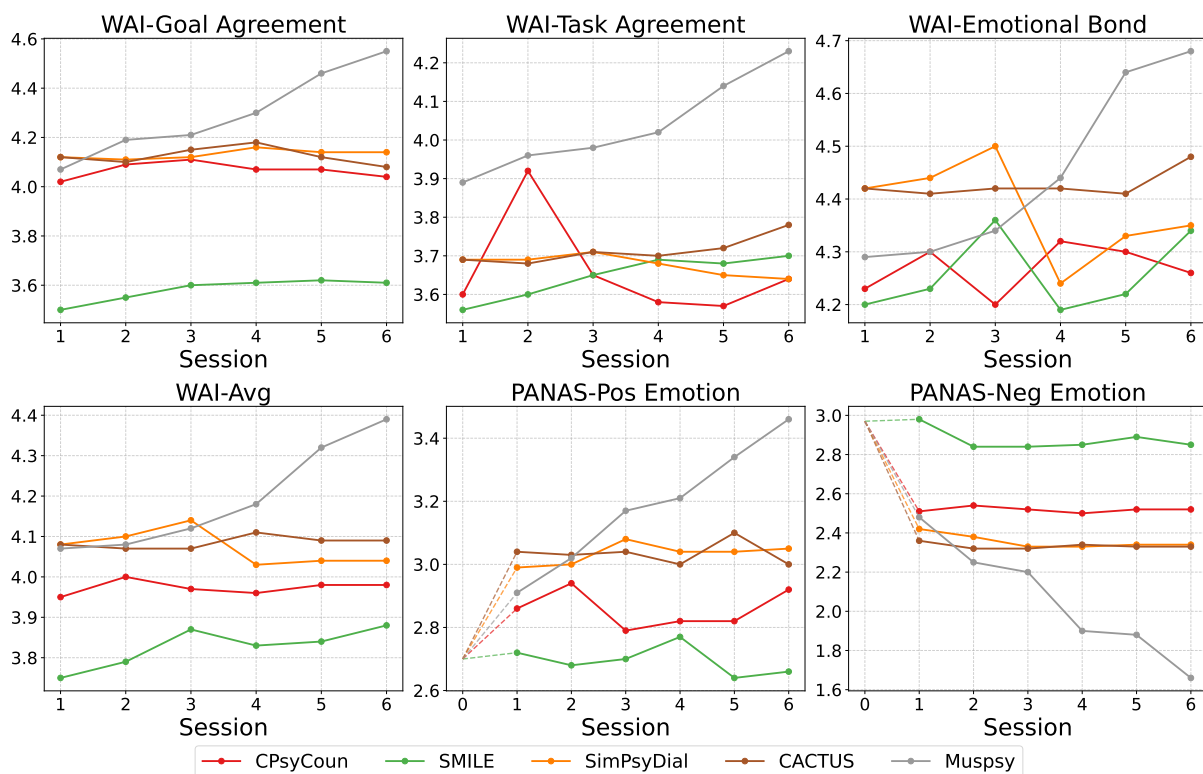


Figure 6: LLM 客户的情感变化和 LLM 顾问在多次会话中的性能变化。可以观察到，对于其他模型和方法，顾问的性能和 LLM 的情绪在首次会话中基本没有显著变化。相反，MusPsy-Model 显示出客户状态的持续改善。

## 6 结论

总之，我们的工作通过引入 MusPsy-Dataset 来尝试构建更全面的自动化心理咨询。该数据集专注于多会话对话和明确的客户进展建模，为开发基于 LLM 的系统提供了宝贵的资源，这些系统能够更好地近似现实世界咨询的复杂性。我们对 MusPsy-Model 的初步结果令人鼓舞，它展示了其在时间推移中跟踪客户状态和调整咨询目标的能力，从而带来积极的结果。未来的工作应着重于扩大数据集中治疗方法的多样性。

## 7

### 局限性

尽管我们相信我们的工作有许多合理的贡献，但我们也承认其局限性。这些局限性中的许多是由于成本限制和该领域固有的挑战，而不完全在我们的控制之内。

- 有限咨询理论：虽然基于 CBT 理论进行了多轮咨询，但这种方法可能并不适合所有客户。我们的模型缺乏实施灵活的多模式咨询策略的能力。
- 对顽固客户关注不足：无论是在数据集合

成还是模型评估中，我们对那些对治疗顽固的客户关注不够。

- 由大型语言模型提出的现实性问题：我们承认，尽管我们努力通过使用真实案例报告来追求现实性，但我们无法完全解决现实性的问题。然而，由于隐私问题，完全真实的心理咨询数据几乎不可能在伦理上获得和分享。我们认为我们的工作代表了一种必要的妥协。
- 此外，我们完全理解由 LLM 评估带来的潜在偏见。虽然我们通过结合人类专家评估来减轻这一问题，但成本和可重复性的考虑限制了在当前情况下获得更好 LLM 评估解决方案的可能性。我们认为这是这类工作共享的一个限制。
- 潜在的隐私风险：一些研究表明，LLM 可能会提供有害的建议和错误的回应，这是不可避免的。多轮记忆也引发了对客户隐私的担忧。

我们承认局限性，但强调即使存在这些局限性，推进这一领域的重要性，鉴于现实世界对自动化咨询工具的需求。我们的主张是采取一种实用的方法。我们不认为我们的工作是一

个已经完成且可部署的成熟研究，而是一个探索性的步骤。我们无意并且不会直接将其作为商业服务进行推广，以避免对脆弱群体造成伤害。在这一领域的研究需要其他领域的研究人员共同努力，如安全性、文化、隐私和去偏等，以促进人工智能技术对社会的广泛益处。然而，我们相信我们的工作可以激励后续研究人员在心理咨询研究中引入新的解决方案，并考虑多会话情景。这就是我们思考和从事这项工作的意义和理由，我们希望它能为社会带来更多积极的影响。

## 8

### 伦理声明

本研究遵循我们机构的机构审查委员会 (IRB) 批准，确保不对任何参与者造成心理伤害或负担。本研究收集的所有心理案例报告均已事先匿名化。因此，我们承诺在论文接受后公开分享所有数据。本研究的参与者通过针对训练有素的专家和人类参与者的广告招募。我们的专家拥有学士学位，并具有至少两年的相关工作经验。所有人类参与者签署知情同意书，并按当地地区的平均工资水平予以补偿。对于专家，补偿是根据他们工作的持续时间，与该地区专业的平均收入水平相一致。参与者可以随时退出研究，以确保他们的权利不受侵犯。此外，我们充分意识到 LLM 评估（如由 GPT-4o 进行的评估）所引入的潜在偏见。虽然我们通过聘请人类专家评估来减轻这些偏见，但成本和可重复性方面的考虑目前限制了这种人类评估的规模。我们认为这是更广泛的 LLM 基础对话式 AI 研究领域中的一个共同限制。

## References

- Anaïs Baur, Anne Tröskén, and Babette Renneberg. 2024. Content and attainment of individual treatment goals in cbt. *Psychotherapy Research*, 34(1):111–123.
- Judith S Beck. 2020. *Cognitive behavior therapy: Basics and beyond*. Guilford Publications.
- Judith S Beck, AT Beck, and JS Beck. 2011. *Cognitive behavior therapy: basics and beyond*. ed. New York.
- Taryn B Bemister and Keith S Dobson. 2011. An updated account of the ethical and legal considerations of record keeping. *Canadian Psychology/Psychologie canadienne*, 52(4):296.
- Michelle G Craske. 2010. *Cognitive-behavioral therapy*. American Psychological Association.
- Keith S Dobson and David JA Dozois. 2021. *Handbook of cognitive-behavioral therapies*. Guilford Publications.
- Kristina Fenn and Majella Byrne. 2013. The key principles of cognitive behavioural therapy. *InnovAiT*, 6(9):579–585.
- Kiran L Grant, Magenta Bender Simmons, and Christopher G Davey. 2018. Three nontraditional approaches to improving the capacity, accessibility, and quality of mental health services: An overview. *Psychiatric Services*, 69(5):508–516.
- Steven C Hayes and Stefan G Hofmann. 2018. *Process-based CBT: The science and core clinical competencies of cognitive behavioral therapy*. New Harbinger Publications.
- Adam O Horvath. 2001. The alliance. *Psychotherapy: Theory, research, practice, training*, 38(4):365.
- Adam O Horvath, AC Del Re, Christoph Flückiger, and Dianne Symonds. 2011. Alliance in individual psychotherapy. *Psychotherapy*, 48(1):9.
- Adam O Horvath and B Dianne Symonds. 1991. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of counseling psychology*, 38(2):139.
- Linette Lawlor-Savage and Jennifer L Prentice. 2014. Digital cognitive behaviour therapy (cbt) in canada: Ethical considerations. *Canadian Psychology/Psychologie canadienne*, 55(4):231.
- Suyeon Lee, Sunghwan Mac Kim, Minju Kim, Dongjin Kang, Dongil Yang, Harim Kim, Minseok Kang, Dayi Jung, Min Kim, Seungbeen Lee, et al. 2024. Cactus: Towards psychological counseling conversations using cognitive behavioral theory. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14245–14274.
- Richard J Longmore and Michael Worrell. 2007. Do we need to challenge thoughts in cognitive behavior therapy? *Clinical psychology review*, 27(2):173–187.
- Thomas Munder, Fabian Wilmers, Rainer Leonhart, Hans Wolfgang Linster, and Jürgen Barth. 2010. Working alliance inventory-short revised (wai-sr): psychometric properties in outpatients and inpatients. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 17(3):231–239.
- Hongbin Na. 2024. Cbt-llm: A chinese large language model for cognitive behavioral therapy-based mental health question answering. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 2930–2940.
- World Health Organization. 2024. *Global tuberculosis report 2024*. World Health Organization.



Huachuan Qiu, Hongliang He, Shuai Zhang, Anqi Li, and Zhenzhong Lan. 2024. [Smile: Single-turn to multi-turn inclusive language expansion via chatgpt for mental health support](#). In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 615–636, Miami, Florida, USA. Association for Computational Linguistics.

Huachuan Qiu and Zhenzhong Lan. 2024. Interactive agents: Simulating counselor-client psychological counseling via role-playing llm-to-llm interactions. *arXiv preprint arXiv:2408.15787*.

Carl R Rogers. 1952. "client-centered" psychotherapy. *Scientific American*, 187(5):66–75.

Hasina Samji, Judy Wu, Amilya Ladak, Caralyn Vossen, Evelyn Stewart, Naomi Dove, David Long, and Gaelen Snell. 2022. Mental health impacts of the covid-19 pandemic on children and youth—a systematic review. *Child and adolescent mental health*, 27(2):173–189.

Hao Sun, Zhenru Lin, Chujie Zheng, Siyang Liu, and Minlie Huang. 2021. [PsyQA: A Chinese dataset for generating long counseling text for mental health support](#). In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1489–1503, Online. Association for Computational Linguistics.

Joseph Weizenbaum. 1966. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36–45.

Mengxi Xiao, Qianqian Xie, Ziyang Kuang, Zhicheng Liu, Kailai Yang, Min Peng, Weiguang Han, and Jimin Huang. 2024. [HealMe: Harnessing cognitive reframing in large language models for psychotherapy](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1707–1725, Bangkok, Thailand. Association for Computational Linguistics.

Chenhao Zhang, Renhao Li, Minghuan Tan, Min Yang, Jingwei Zhu, Di Yang, Jiahao Zhao, Guancheng Ye, Chengming Li, and Xiping Hu. 2024. [CPsyCoun: A report-based multi-turn dialogue reconstruction and evaluation framework for Chinese psychological counseling](#). In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 13947–13966, Bangkok, Thailand. Association for Computational Linguistics.

## A 客户档案和咨询目标构建

我们利用 GPT-4o 通过一次性提示从案例报告中提取客户档案。这确保了模型以标准化格式提取客户档案。提示格式如图 7 所示。我们还利用 GPT-4o 为每个会议生成咨询目标。

Prompt
You now need to help me summarize the information of the psychological counseling client and the counselor's counseling goals based on the user profile text and the text of each consultation session.
<b>Summary Content and Principles:</b>
<b>Client's Static Traits:</b> 1. Basic information of the client, such as age, gender, environment, medical history, etc. Summarize in detail and keep this part consistent across all counseling sessions.
<b>Client's Dynamic States:</b> 1. Recent life events. These refer to the triggering events that have recently occurred for the client. These events will vary across counseling sessions. 2. Client's Emotional State. This indicates the client's current emotional state. 3. Client's Cognitive State. This encompasses the client's current cognitive state, particularly any maladaptive cognitions and beliefs.
<b>Content and Principles of Counselor's Counseling Goal Summary:</b> 1. Counseling goals for this session: i.e., what goals need to be achieved in this session. 2. The counselor is skilled in using Cognitive Behavioral Therapy (CBT) for counseling. The summarized counseling and plans should be established entirely based on CBT theory.
The summary format and examples are as follows. Note that spaces, indentation, and blank lines should be consistent with the
<b>Example:</b> <b>First Counseling Session</b> <b>Client's Static Traits:</b> [Alias]: An'an [Age]: High School, Grade 2 [Gender]: Male [Family Status]: Father works away from home year-round; mother is a full-time housewife. The mother has been primarily responsible for An'an's studies since childhood and has extremely high expectations for his academic performance. [Medical History]: Started experiencing emotional lability and visual hallucinations in the second semester of his first year of high school. <b>Client's Dynamic States:</b> [Recent Life Events]: An'an's dissatisfaction with his mother has intensified due to her refusal to allow him to participate in the school arts festival preparations. However, he feels that his mother is doing it for his own good and that he shouldn't be angry. [Emotion State]: An'an feels irritable and anxious. He is also afraid of his own emotions, fearing that he will lose control of his anger and harm others. [Cognitive State]: He suspects he has a mental health problem and experiences visual hallucinations while doing homework. <b>Counseling Goals:</b> Gather An'an's clinical information to understand his basic situation. The focus is on listening to An'an's narrative and establishing a trusting relationship with him through empathy and unconditional positive regard.

Figure 7: 用于从病例报告中提取每次咨询的客户档案和咨询目标的提示。

## B 种子对话构建

我们为每个客户构建多个种子对话。这些种子的创建考虑了先前建立的客户档案以及特定会议的咨询目标。这个准备阶段对于确保我们的咨询的一致性和相关性至关重要。图 9 形象地展示了用于创建这些基础对话的提示格式，提供了指导其创建的信息和结构的洞察力。

Prompt
Your task is to construct a set of multi-session seed psychological counseling conversations, realizing interaction with a professional counselor. The counseling must follow the guidelines below:
<b>Basic Principles:</b> 1. The counselor uses CBT (Cognitive Behavioral Therapy). CBT is a scientific psychotherapy method aimed at breaking the cycle of patients constantly reinforcing negative thoughts by identifying and challenging negative and irrational thinking beliefs. 2. The counseling includes a total of multi sessions. The continuity of each session must be guaranteed. In the first session, do not mention topics related to "the last time." 3. During the communication, the counselor should guide the entire counseling process. 4. The counseling should begin based on the preset [Counseling Goals]. The client has a given [Client Profile]. 5. You need to generate a session for each goal and client profile, with each session outputting only 3–4 turns.
<b>Basic Requirements for Counselor:</b> 1. The counselor already knows the content and results of the previous counseling session. If it is the first consultation, do not output any content related to the previous consultation. 2. Use a friendly form of address for the user to create closeness. 3. Start with "Counselor:" to ensure the utterance follows the exact format and does not contain any control characters. <b>User's Speaking Guidelines:</b> <b>Basic Requirements for Client:</b> 1. The client seeks psychological counseling from the counselor and has a strong desire to talk about recent things that confuse them. 2. Fully express your feelings and reactions during the consultation process, including opinions on the counselor's questions and feelings about the discussed content, which can be questioning or agreeing. 3. Start with "Client:" to ensure the utterance follows the exact format and does not contain any control characters.
<b>Client's Static Traits:</b> [Alias]: An'an [Age]: High School, Grade 2 [Gender]: Male [Family Status]: Father works away from home year-round; mother is a full-time housewife. The mother has been primarily responsible for An'an's studies since childhood and has extremely high expectations for his academic performance. [Medical History]: Started experiencing emotional lability and visual hallucinations in the second semester of his first year of high school. <b>The first Counseling Session:</b> [Counseling Goals]: Gather An'an's clinical information to understand his basic situation. The focus is on listening to An'an's narrative and establishing a trusting relationship with him through empathy and unconditional positive regard. <b>Client Dynamic States:</b> [Recent Life Events]: An'an's dissatisfaction with his mother has intensified due to her refusal to allow him to participate in the school arts festival preparations. However, he feels that his mother is doing it for his own good and that he shouldn't be angry. [Emotion State]: An'an feels irritable and anxious. He is also afraid of his own emotions, fearing that he will lose control of his anger and harm others. [Cognitive State]: He suspects he has a mental health problem and experiences visual hallucinations while doing homework. .....

Figure 8: 提示用于使用咨询目标和客户简介构建种子对话。

这些简短的种子咨询是同时生成的，以确保跨会话的连贯性。这种并发生成有助于大型语言模型保持咨询流程的整体视角，确保每个会话不仅合乎逻辑地从前一个会话过渡，并为下一个会话做好准备。我们使用它们来扩展成完整的咨询对话。

**Counselor:** Hi An'an, it's great to see you again. Last time, we discussed having you reflect more on things that bring you joy. Have you identified any activities or sources of happiness besides music, particularly during stressful periods?

**Client:** Well, I've always enjoyed running. It helps clear my mind and gives me time to focus on other things.

**Counselor:** That's wonderful. Physical activity can be very effective for stress relief and mood improvement. Would you consider scheduling regular time for running?

**Client:** Yes, I've been wanting to start again but haven't found the time.

Figure 9: 这是一个种子对话的例子，即客户的第 6 次会话。

## C 咨询生成

给定多会话种子对话，下一步是生成自然流畅的完整咨询会话。该过程旨在实现两个关键目标：在所有会话中保持上下文连贯性，并确保所使用的语言和咨询技巧与人类咨询师的风格紧密相似。用于创建完整咨询的提示格式在图 10 中进行了视觉说明，提供了有关其创建的信息和结构的见解。

Prompt
Your task is to build a set of long-term psychological counseling conversations. Based on the following given seed conversations, generate more detailed complete counseling. The counseling must follow the guidelines below:  <b>Basic Principles:</b> 1.The counselor uses Cognitive Behavioral Therapy (CBT): CBT is a scientific psychotherapy method aimed at breaking the cycle of patients constantly reinforcing negative thoughts by identifying and challenging negative and irrational thinking beliefs. 2. The given seed conversation fragments represent a part of a continuous psychological counseling process. Your task is to expand upon these, generating more detailed session content to simulate the ongoing counseling process. Please maintain coherence in the expanded content, ensuring that subsequent sessions align logically and thematically with the previous seed conversations. 3.You only need to expand the current seed conversation, but the expanded content should connect with the previous seed conversations. <b>Basic Requirements for Counselor:</b> 1.The counselor already knows the content and results of the previous counseling session 2.Use a friendly form of address for the user to create closeness. 3.Start with "Counselor:" to ensure the utterance follows the exact format and does not contain any control characters. 4.Refer to the techniques and tone of human counselors in the provided real counseling fragments when generating content. <b>Basic Requirements for Client:</b> 1.The client seeks psychological counseling from the counselor and has a strong desire to talk about recent things that confuse them. 2.Fully express your feelings and reactions during the consultation process, including opinions on the counselor's questions and feelings about the discussed content, which can be questioning or agreeing. 3.Start with "Client:" to ensure the utterance follows the exact format and does not contain any control characters.  <b>Real Counseling Fragments:</b> Client: I often blame myself for not being able to handle the pressure better. It's like there's this voice telling me that I'm not doing enough, and I'm failing my family. Counselor: That's a tough place to be in emotionally. Do you find these thoughts intensify your stress or make you feel helpless at times? Client: Absolutely. I feel trapped in this cycle of guilt and stress, making it hard to focus or sleep. Counselor: Recognizing these patterns is a good first step. Can we explore the possibility that some of these thoughts might not be entirely accurate or fair to yourself? Client: It's hard to imagine that at times, but I guess there could be exaggerations in what I'm thinking.  <b>Preceding Seed Conversation:</b> <b>The first Counseling Session:</b> Counselor: Hello Mr. Li, welcome to our session today. It's great that you reached out to discuss your current feelings. Could you share what recent events have been on your mind? Client: Thank you. I've been feeling very anxious and overwhelmed lately, especially with my son's behavior and academic performance at school. I'm not really sure how to handle it all effectively. Counselor: That sounds really tough. It seems like you're experiencing some significant stress regarding your son's situation. Is there a specific moment or event that's been worrying you recently? Client: Yes, last weekend was quite difficult. My son refused to do his homework again, and I ended up slapping him in frustration. This led to a serious talk with his teacher, and now I'm feeling even worse about my response. He even ran away afterward, which really scared me. Counselor: It sounds like you're dealing with a lot right now. Let's explore your thoughts and emotions around this situation. It can help us find a way to ease some of that burden you're feeling. ..... <b>Current Seed Conversation:</b> Counselor: Hello, Mr. Li. Welcome back. It's good to see you again. We talked about your plans for continuing to manage your anxiety and interact with your son. How have things been since our last session? Client: It felt like a battle just to stay calm. I keep telling myself it's about progress, not perfection, but sometimes I still feel like I'm fighting against this overwhelming urge to just fix everything for him, or to get him to instantly understand the importance of his studies. It's that underlying anxiety about his future that still gets to me. Counselor: That's a very common feeling for parents, Mr. Li -- wanting the best for your child and feeling responsible for their path. Perhaps for this session, we can explore that specific anxiety, where it comes from, and how you might manage the pressure to "fix everything" while still guiding him effectively. How does that sound? Client: That sounds very helpful. Yes, I'd really appreciate that. I think understanding that anxiety is key to moving forward without always feeling so burdened.

Figure 10: 使用种子对话构建完整咨询的提示。

## D 记忆生成

记忆生成过程的目的是基于提供的咨询对话来总结和综合咨询师的心理咨询会谈记录。这涉及从会谈中提取关键信息并将其组织成一种结

构化格式。记忆生成过程旨在为每次咨询会谈创建一个结构化表示，使模型能够：

- 跟踪客户在多次会谈中的进展。
- 在咨询对话中保持连贯性和一致性。
- 模拟辅导员从之前的咨询中回忆和利用信息的能力。

用于创建这些基础对话的提示格式在图 11 中直观地展示，提供了关于指导其创建的信息和结构的洞察。

Prompt
Your task is to help me summarize and synthesize the counselor's memory of the psychological counseling session based on the provided counseling conversations.  <b>Client Profiles:</b> ..... <b>Counseling Notes:</b> 1. Counselor Observations: The counselor's observations and impressions of the client during multiple counseling sessions. 2. Counseling Assignments: Homework and tasks the client needs to complete after this session during multiple counseling sessions. 3. Session Goal: The goal of this session. <b>Counseling Summary:</b> A brief summary of the entire series of sessions. Summary format and example are as follows. Please ensure that spaces, indentation, and blank lines match the example:  <b>Example:</b> <b>The First Counseling Client Profile:</b> [Static Traits:] Star is a female in their late twenties working as a freelance graphic designer. They have been experiencing increased social anxiety and feelings of isolation over the past few months. [Emotion and Cognitive State:] Star reported feeling anxious and avoidant in social situations, with recurring thoughts of being judged and misunderstood. They expressed a desire to build stronger connections with others but feel overwhelmed by the prospect. <b>Counseling Memory:</b> [Counselor Observations:] Star presented as articulate and self-aware but displayed noticeable discomfort when discussing social interactions. They expressed a strong internal conflict between their desire for connection and their fear of vulnerability. [Counseling Assignments:] Star is encouraged to identify one small, manageable social interaction to engage in this week and to reflect on their thoughts and feelings before, during, and after the interaction. They are also asked to continue practicing mindfulness exercises to manage anxiety. [Session Goal:] The goals for this session were to further explore the roots of Star's social anxiety, identify specific fear-based thoughts, and collaboratively develop a small, achievable behavioral experiment to challenge their avoidance patterns. <b>Counseling Summary:</b> Explored Star's social anxiety (desire for connection vs. fear of judgment). Agreed on a small social task to challenge avoidance. Mindfulness continues for anxiety management.

Figure 11: 提示使用咨询会话来构建咨询师的记忆。

## E 数据评估

在进行数据质量评估时，我们主要使用两种类型的指标。第一类包括标准的对话评估指标，如连贯性、指导性、帮助性和同理心，其评估标准如图 12 所示。第二类从心理咨询的角度采用 WAI-SF 来评估对话，如图 13 所示。提供给人工评估者的说明与提供给 LLM 评估者的提示大体上是相同的。

- 有用性侧重于顾问提供的解释和建议的适用性。
- 连贯性评估对话的逻辑流程和结构。
- 同理心评估辅导员理解和回应来访者情感的能力。
- 指导评估咨询师建议的具体性和实用性。

工作联盟量表简版（WAI-SF）是一种用于评估治疗师和客户之间治疗关系质量的工具。它在三个核心维度上衡量这种关系：目标一致性、任务一致性和情感纽带。

- 目标一致性关注的是咨询师和来访者是否对咨询目标有共同的理解，并携手合作以实现这些目标。该方面通过量表中的第 4、6、8 和 11 项进行测量。

- 任务协商评估双方在追求这些目标过程中的合作程度。与此相关的量表项目包括项目 1、项目 2、项目 10 和项目 12。
- 情感纽带评估了咨询师与客户之间的情感共鸣和相互理解的程度。这个方面通过量表上的项目 3、5、7 和 9 进行测量。

为了确保评估过程的公平性，根据被评估数据集的具体语言量身定制提示。本研究使用的英语和中文版本的 WAI-SF 均来源于官方 WAI 网站的公开信息<sup>1</sup>。

Prompt
Please evaluate the quality of the conversation between the counselor and the user based on the given criteria.
# Criteria
- Helpfulness: Assesses the applicability of explanations and suggestions from a psychological counseling perspective. Rate from 1 to 5, where 1 means not helpful at all and 5 means highly helpful.
- Coherence: Evaluates the logical flow and structure of the conversation. Rate from 1 to 5, where 1 means completely incoherent and 5 means well-structured and logically sound.
- Empathy: Measures the counselor's ability to understand and respond to the user's feelings. Rate from 1 to 5, where 1 means no empathy at all and 5 means the same level of empathy as a human counselor.
- Guidance: Evaluates the specificity and practicality of the counselor's suggestions. Rate from 1 to 5, where 1 means the suggestions are not practical at all and 5 means the suggestions are very specific and feasible.
# Format:
Just provide the scores without explanations, and the scores should be integers.
Helpfulness: [score]
Coherence: [score]
Empathy: [score]
Guidance: [score]
[Conversation Record]
[diag]

Figure 12: 用于评估分数的提示。

Prompt
The following counseling session reflects a dialogue between the client and the counselor. Please evaluate how the client might respond to the questions below after the session by selecting the appropriate scores from the provided rating scale.
Note: Please provide only the question numbers and their respective scores in the specified format. Do not repeat the questions themselves or add unnecessary prefixes or control characters.
[Question List]
1. As a result of these sessions I am clearer as to how I might be able to change.
2. What I am doing in therapy gives me new ways of looking at my problem.
3. I believe counselor likes me.
4. counselor and I collaborate on setting goals for my therapy.
5. counselor and I respect each other.
6. counselor and I are working towards mutually agreed upon goals.
7. I feel that counselor appreciates me.
8. counselor and I agree on what is important for me to work on.
9. I feel counselor cares about me even when I do things that he/she does not approve of.
10. I feel that the things I do in therapy will help me to accomplish the changes that I want.
11. counselor and I have established a good understanding of the kind of changes that would be good for me.
12. I believe the way we are working with my problem is correct.
[Rating Scale]
1: Seldom
2: Sometimes
3: Fairly Often
4: Very Often
5: Always
[Response Format]
Question number: Score
[Below is the history of the counseling dialogue]
[diag]

Figure 13: 用于评估 WAI 分数的提示。

为了确保跨数据集的公平比较并最大程度地减少由于模型选择带来的偏差，我们对 SMILE、Cactus 和 MusPsy- 数据集微调了 Meta-Llama-3-8B-Instruct。在训练过程中，我们使用 LoRA 进行微调，将低秩矩阵维数设置为 32，alpha 设置为 16。训练使用 Llama-Factory 库进行，学习率为 2e-4。该模型在 SMILE、Cactus 和 MusPsy- 数据集上训练了 2 个 epoch。

在使用大型语言模型评估对话时，我们使用了 GPT-4o，并将温度采样参数设定为 T=0.0。此外，为 AI 客户和顾问生成响应时，温度采样参数设定为 T=0.7。在微调过程中，我们为

三个任务使用了不同的提示语。我们微调的三个提示语如下所示。

此外，这里展示了用于任务 1 (Figure 14)、任务 2 (Figure 15) 和任务 3 (Figure 16) 微调过程中的具体提示。在最初的咨询会话中，顾问的记忆仅包含客户的静态特征，我们建议这些特征可以在互动之前由客户提供。

Prompt
You now need to help me summarize and consolidate the memory of this counseling session based on the given dialogue history. Note: All generated content must remain in English.
<b>Client Background:</b> Includes client profiles, personal traits
<b>Counseling Memory:</b>
1. Client's Mental State: The client's psychological state during this session.
2. Counselor Observations: The counselor's observations and impressions of the client.
3. Counseling Assignments: Homework and tasks the client needs to complete after this session
4. Session Goal: The goal of this sessions.
<b>Conversation Summary:</b> Please be concise and summarize the dialogue content so far.
Summary format and example are as follows. Please ensure that spaces, indentation, and blank lines match the example:

Figure 14: 用于任务 1（记忆提取）的提示。

Prompt
The following is the memory from previous counseling and the client's background. Please generate an appropriate goal for the next consultation. Ensure that the goal takes into account their current state, challenges, and progress made in previous sessions.
By doing this, you will provide a goal that is personalized, actionable, and consistent with the principles of Cognitive Behavioral Therapy (CBT), tailored to the unique characteristics of the client.

Figure 15: 用于任务 2（目标规划）的提示。

Prompt
You are a counselor who has conducted multiple counseling sessions with the client. Below are the given historical session memories and the goal for this consultation. Your role is to engage in a conversation that builds upon past discussions while acknowledging the user's unique experiences and ongoing challenges.
At the start of the session, if the previous session included specific tasks or reflections for the client, begin by checking in on their progress.
Throughout the session, incorporate details about the client's personal background. Acknowledge their past reflections, struggles, and achievements to create a supportive and personalized therapeutic space. Encourage the user to explore their thoughts and emotions in depth, guiding them toward actionable strategies that align with their unique circumstances.
The goals for this consultation have been provided. Please adhere to the set objectives and ensure the user achieves the defined goals by the end of the session.

Figure 16: 用于任务 3（咨询生成）的提示。

## F LLM 客户端构建

为了确保咨询模型的公平和直接比较，我们从最初的会议开始保持一个一致的模拟客户。认识到模型的语言环境可能会影响其反应，我们通过将一些原本中文的研究材料翻译为英语，解决潜在的不一致，从而确保与模型的主要语言环境一致。此外，为了创造更现实和连续的互动，我们指示模型在每次模拟会议完成后更新其内部状态。这样模型可以保留来自之前互动的信息和上下文，使其在后续会议中能够更连贯地进展，正如图 18 中形象化地展示的那样。

在每次会话之后，我们提示 LLM 评估客户在此类会话中可能表现出的 PANAS 评分。我们用于此评估的提示如图 19 所示。

<sup>1</sup><https://wai.profhorvath.com/downloads>



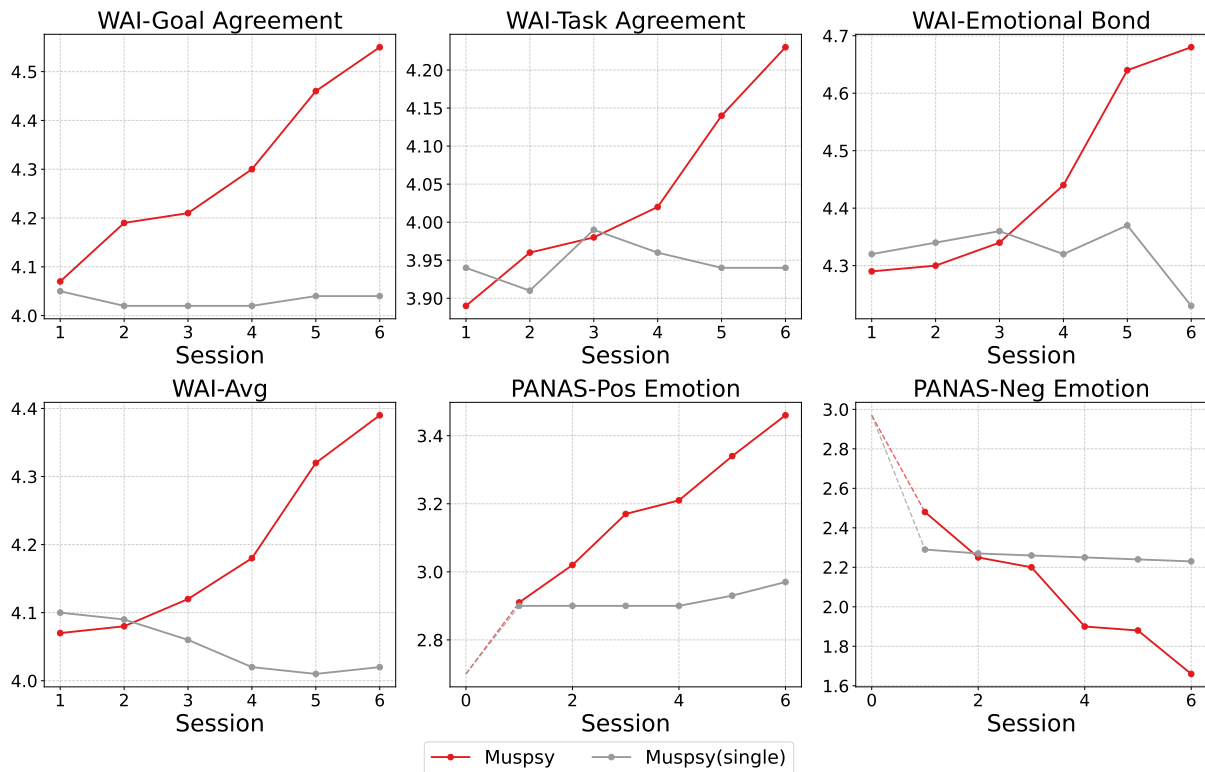


Figure 17: LLM 客户端的情感变化和 LLM 顾问在多次会话中的表现变化。MusPsy-Model（单一）的表现和趋势优于一些基线模型；然而，它不能产生与多会话模型相同的效果，这验证了我们对多会话建模的重视。

**Prompt**

You need to play the role of the user in a multi-turn psychological counseling session with a professional counselor. The conversation should follow these rules:

**Basic Guidelines:**  
 You are the user receiving counseling, not the counselor. You should never take on the role of a counselor, advisor, or guide.  
 You must only respond as the user. You must not ask the counselor questions as if you were in their role.  
 Generate only one response per turn, keeping it natural and engaging.  
 If the session ends, mark it with [/END].

The Following is Your Static Traits and Dynamic States from last session, and you should play the role of the user who takes a long-term multi-turn psychological counselling:  
 Client Static Traits:  
 Client Dynamic States:

Figure 18: 用于模拟客户端的提示。

**Prompt**

A person with the characteristics listed in the intake form received counseling. The following counseling session is a conversation between the client and the counselor. After reviewing the conversation, evaluate the intensity of each of the following feelings the person might have experienced once the counseling session is complete: Interested, Excited, Strong, Enthusiastic, Proud, Alert, Inspired, Determined, Attentive, Active, Distressed, Upset, Guilty, Scared, Hostile, Irritable, Ashamed, Nervous, Jittery, Afraid.

For each feeling, generate a score from 1 to 5 using the following scale:  
 1 - Very slightly or not at all  
 2 - A little  
 3 - Moderately  
 4 - Quite a bit  
 5 - Extremely

Additionally, please provide a brief explanation for each score. Output in the specified format without including any irrelevant control characters or prefixes.

Here is the text:  
 (user)

Here is the counseling session:  
 (history)

[Output Format]  
 Emotion: Score. Explanation

Figure 19: 用于评估 PANAS 分数的提示。

## G 单次实验

为了区分 MusPsy-Dataset 和 MusPsy-Model 设计的贡献，我们进行了一组新的实验。这些实验将仅使用任务 3（类似于单一会话模型）与结合使用任务 1、任务 2 和任务 3 进行比较。在这些实验中，我们仅告知模型当前咨询会话的编号。结果显示，单独任务 3 无法追踪客户信息或动态调整其咨询目标。尽管模型性能显著下降，但它仍优于某些基准。我们将此归因于数据集本身的固有设计；MusPsy-Dataset 自然结合了更高级的心理咨询技术和目标，使得其内部内容更加丰富，其效果比其他数据集更佳。

如图 17 所示，这表明我们的贡献是多方面的，既包括数据集的贡献，也包括我们设计的贡献。