个性化以人为中心的多模态推理：一项新任务、一个数据集和基线方法

Yaochen Zhu、Xiangqing Shen 和 Rui Xia中国南京理工大学计算机科学与工程学院
{yczhu, xiangqing.shen, rxia}@njust.edu.cn

arXiv:2304.02313v2 [cs.CL] 4 Mar 2024

摘要

人格特质、情感和信念会影响个体的行为选择和决策过程。然而，一方面，情感计算领域通常侧重于预测人格特质，却忽视了它们在行为预测中的应用。另一方面，多模态推理任务强调预测未来状态和行为，却往往忽略了个体人格特质的融入。本文提出一项名为个性化以人为中心的多模态推理（PHMR）（T^1）的新任务，旨在结合人格因素，利用个体过去经历中的多模态信息预测其未来行为。为此，我们基于六档电视节目构建了一个新数据集，涵盖225个角色和1.2万个样本。为该任务建立基准测试，我们提出了七种基线方法：三种改编自相关任务的方法、两种预训练模型，以及两种多模态大语言模型。实验结果表明，融入人格特质能够提升以人为中心的多模态推理性能。为进一步解决真实场景中缺乏人格标注的问题，我们还提出一项扩展任务——人格预测型以人为中心的多模态推理任务（T^2），并提供相应的数据集和方法。我们将在 GitHub 上公开数据集和代码。 [1]

1 引言

个体的人格特质、情感和信念是其个体差异的重要组成部分，并且会在不同情境下塑造其行为选择和决策过程。人格概括了个体的心理特征和行为倾向。个体的反应、性格特征——以多样人格特质为特征的人，面对相同情境时可能表现出显著差异。例如，当面对“人物*1*在公共场合为人物*0*唱歌”这一事件时，谨慎的和敏感的 人物*0* 可能感到尴尬，而自信的 和活泼的人物*0* 可能感到兴奋。

然而，一方面，在情感计算领域，研究通常侧重于预测和分析人格特质，却没有将其用于预测与个体相关的未来事件。另一方面，近年来出现了一类以人为中心的多模态推理任务，包括 Social-IQ (Zadehetal., 2019年)、VLEP (Lei etal.,2020年) 和 PMR (Donget al., 2022年)，旨在利用现有的多模态信息推断个体的心理状态和行为。现有研究主要关注与个体相关的客观信息，却忽视了在推理中纳入个体的人格特质。

在本研究中，我们引入了一项名为人格感知的人本多模态推理（PHMR）（T^1）的新任务。我们的任务目标是，在涉及多个个体和长期互动的复杂社会交往中，预测特定个体在未来场景中最有可能采取的行为。这是通过整合过去时刻的多模态信号并考虑其人格特质来实现的。

为支持这项任务，我们基于来自 TVQA（Lei 等人、2018年）中的六部电视节目构建了一个数据集，我们将其称为人格感知的人类中心多模态推理数据集（PHMRD）。我们通过 PDB[2] 网站获取主要角色的人格信息，并将其用作我们数据集中的人格标注。遵循多模态推理任务的主流设置，我们也将任务定义为一个多项选择问题。如图 1 所示，

[1]我们展示了 PHMRD 的 10 个示例，位于匿名 GitHub。

[2]https://www.personality-database.com

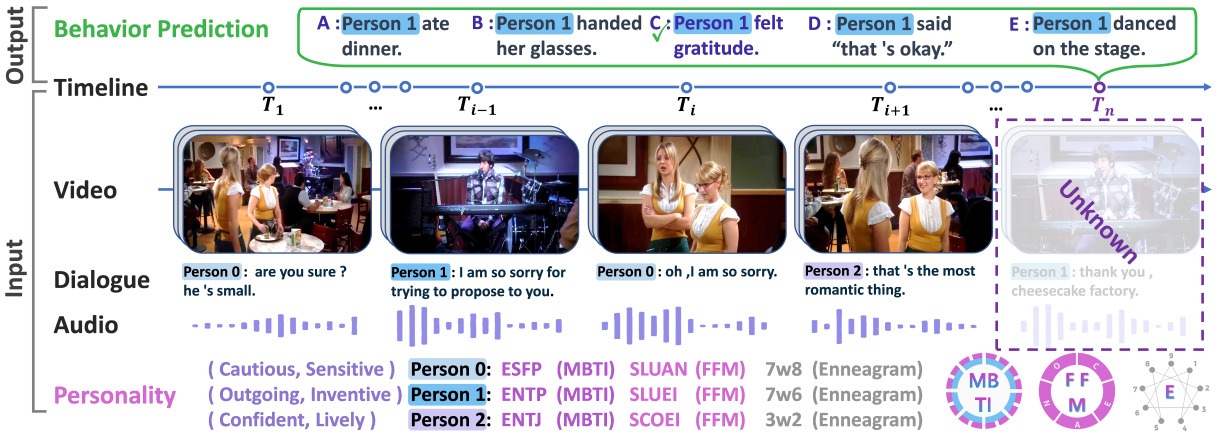

图 1：PHMR 和 PHMRD 的示意图。人物 *0*、*1*和 *2* 表示视频片段中出现的角色。我们的任务是根据视频、对话、音频和个性信息，预测 T_n 处最可能的行为描述。

人物*0*、*1*和*2*代表视频片段中的不同角色。给定以下信息：1）视频的图像帧序列；2）对话中的话语；3）视频的音频波形；4）相应角色的性格。我们的任务旨在选择一个最能描述人物*0*未来最可能行为的选项。例如，应推断出“人物*0*心怀感激”是最可能的选项。

我们通过设计七种基线方法对 PHMR 任务进行基准测试：其中三种改编自相关任务，两种为预训练多模态模型，另外两种为多模态大语言模型。我们采用预训练多模态模型提取多模态特征，并融入人格特质，以预测最可能的行为。人格以嵌入表示的形式呈现，并与其他多模态特征有效融合。

实验结果表明，融入人格特质能够提升单模态和多模态场景下的推理性能。我们提出的方法优于经过适配的基线方法，展现出更佳的性能。我们的多模态消融实验表明，在 PHMR 的单模态场景中，融入人格能够带来性能提升。

为进一步解决真实场景中缺乏人格标注的问题，我们提出了一项名为个性预测的人本多模态推理（T^2）的新任务，并构建了相应的数据集。考虑到可以根据一个人的行为和言语推断其人格，我们首先利用多模态信息预测其人格。随后，我们将这些预测作为人格标注的替代信息，以增强推理过程（T^1）。实验结果表明，我们的方法能够有效预测人格，并在无需依赖人格标注的情况下取得令人满意的多模态推理表现。

2 任务

我们将主要任务定义为个性化以人为中心的多模态推理（T^1）。我们关注涉及多个人物长期互动的场景，其中每个人的行为不仅受他人行动的影响，也受其自身人格特质的影响。

具体而言，我们为一个视频片段 $\mathcal{VC}$提供四类信息：

1. 视频（$\mathcal{V}$）：视频的图像帧序列。2. 音频（$\mathcal{A}$）：视频的音频波形。3. 对话（$\mathcal{D}$）：对话内容的文本话语、角色及相应的时间跨度 T^D。4. 人格（$\mathcal{P}$）：对出现在 $\mathcal{VC}$中的角色所标注的人格。

对于片段 $\mathcal{VC}$，它可以划分为 n 个时间片段 $T_1 \sim T_n$。我们移除时刻 T_n 的所有多模态信息。给定人格 $\mathcal{P}$ 以及位于 $T_1 \sim T_{n-1}$ 的多模态信息 $\mathcal{V}$、$\mathcal{A}$、$\mathcal{D}$，本任务的目的是预测特定个体最可能的行为 $\mathcal{B}$。

	训练集	Dev	Test	总计
# 示例	8768	1938	1910	12616
视频 (s)	74.67	75.84	75.32	74.95
# 对话	15.57	15.46	15.07	15.48
# 人格	3.72	3.81	3.74	3.73
平均选项	15.24	15.40	15.21	15.26
平均对话	6.79	6.60	6.58	6.73

表1：我们的PHMRD数据集统计。avg.表示平均词数。

在时间 T_n 时。这可以表示如下：

$$\mathcal{B} = \text{PHMR}(\mathcal{D}, \mathcal{V}, \mathcal{A}, \mathcal{P}), \quad (1)$$

其中，PHMR表示个性化以人为中心的多模态推理任务。

我们将PHMR建模为一项多项选择任务，遵循其他多模态推理任务的既定设置（Zadeh等人，2019年；Lei等人，2020年；Dong等人，2022年），其目标是从五个选项中选出最可能的答案，记为选项($\mathcal{C}$)。

3 数据集

3.1 数据集来源

我们选择 TVQA（Lei 等人，2018年）作为数据来源。TVQA 是一个基于电视节目的大规模视频问答数据集，涵盖与物体关系和人类活动相关的问题-答案对。我们的任务强调以人为中心的问答，主要关注个体的行为和心理学。

3.2 数据集构建流程

数据过滤 为了使数据更适合我们的任务设置，我们使用 ChatGPT（gpt-3.5- turbo）排除与个体行为和心理学无关的问答对。在给定实例中，我们认为查询“物体是什么颜色？”与个体无关，而询问“人 *0* 为什么喜欢这个物体？”则被视为与人类受试者相关。提示词和其他示例见附录A。上述示例可以由 ChatGPT 准确评估。为了评估 ChatGPT 的结果，我们手动标注了两百个样本。随后，我们对手动标注的样本和 ChatGPT 得出的结果进行卡帕检验。卡帕结果为 0。89，表明 ChatGPT 与手动标注样本之间具有高度一致性。

电视节目	训练集	Dev	Test	总计
# 《老友记》	2602	559	558	3719
# 《生活大爆炸》	2548	549	553	3650
# 《灵书妙探》	1525	329	333	2187
# Met	843	214	190	1247
# House	745	167	165	1077
# Grey	505	120	111	736
Sum	8768	1938	1910	12616

表2：关于六部电视节目样本数量的统计。BBT、Met、House和Grey分别代表《生活大爆炸》、HowI Met You Mother、House M.D.和Grey。

人格	**词汇**
FFM	[RLOAI, SCUEN $\cdots$, RCUEI, SLOAN]
MBTI	[INTJ, ESFP $\cdots$, ENFJ, ISTP]
九型人格	[1w2, 2w1, $\cdots$, 9w1, 1w9]

表3：不同人格的词汇统计。

行为描述改写 我们将多模态推理任务建模为多项选择问题，遵循多模态推理任务（Social-IQ、VLEP、PMR）的主流设置。PHMRD 中的正确选项由改写 TVQA 中的正确选项得出，错误选项也采用相同的方法生成。在初步实验中，我们观察到，当选项以问答对的形式呈现时，模型往往会过分强调问题的重要性。因此，我们手动改写问答对，将其转换为描述行为的陈述句。例如，给定 TVQA 中的问题：“Person*0* 在 Person*1* 在公共场合唱歌时有什么感受？”，以及答案：“Person*0* 感到尴尬。”，改写后的结果为：“Person*0* 在 Person*1* 在公共场合唱歌时感到尴尬”。图 1 展示了一个 PHMRD 示例。我们使用 ChatGPT 按照上述设置改写行为描述。我们选取 500 个样本，手动进行改写，以获得人工改写的行为描述。为了衡量标注者间的一致性，我们使用 MASI[3] (Passonneau，2006年) 指标，得分为 0.72。

[3]一种用于衡量集合间距离的度量，可量化集合数据标注之间的重叠程度。

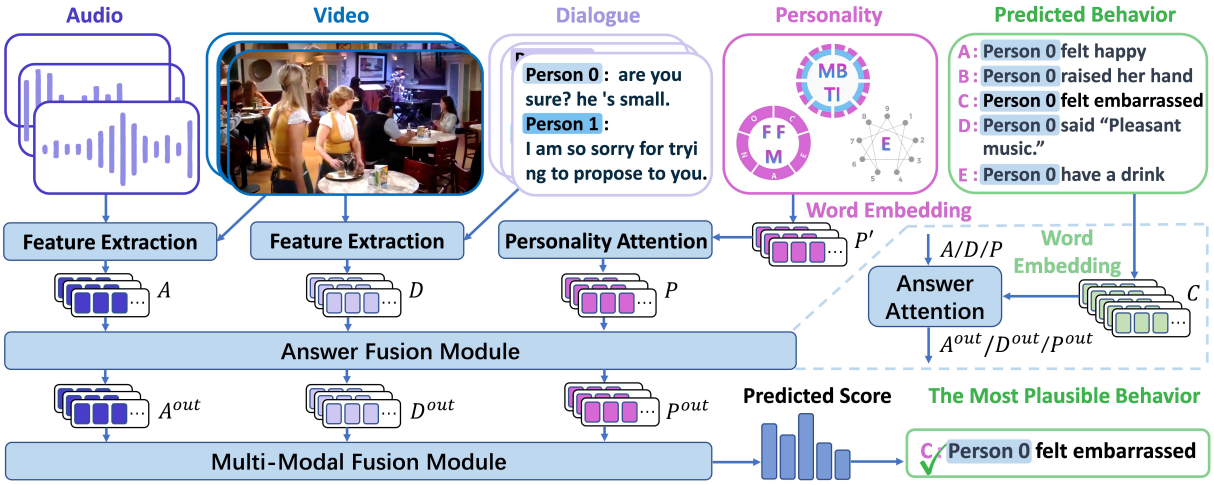

图 2：用于 PHMR 的 PRM^{pretrain} 模型。

人格标注 我们从 Personality Database（PDB）网站获取六部电视节目中角色的人格。每个角色的人格均由该网站上的众包投票确定。我们将得票比例最高的人格作为该角色的人格。

考虑到人格是通过多种评估标准来衡量的，我们标注了心理学中的主流方法，即迈尔斯-布里格斯类型指标（MBTI）、五因素模型（FFM）和九型人格，如表 3所示。例如，佩妮（电视剧《生活大爆炸》中的一个角色）具备 ESFP（MBTI）、SCUAN（FFM）和 7w8（九型人格）人格特质，是一位自信且活泼的个体。更多详情见附录 B。

3.3 数据集统计

该数据集包含来自六部电视剧的 225 个不同个体，每个个体都具有独特的人格特征。表 1概述了 PHMRD 数据集的相关统计信息。PHMRD 数据集中的每个片段平均时长为 74.95 秒，约包含 3.73 个角色。各电视剧中符合任务设置的样本数量见表 2。

4 方法

我们提出一个初步框架，人格感知推理模型（**PRM**），用于验证人格在我们的PHMR任务中的有效性。整体结构**PMR** 如图 2所示。

4.1 多模态信号表示

由于在大规模数据上预训练的多模态模型能够更好地表示来自不同模态的特征，我们利用预训练模型 Merlot Reserve（Zellers 等人，2022年），该模型是 TVQA 的 SOTA，用于特征提取。我们采用记号 $\mathcal{V}$、$\mathcal{A}$、$\mathcal{D}$、$\mathcal{P}$、$\mathcal{C}$ 来表示第 2 节中提到的视频、音频、对话、人格和选项。

对于视频模态 $\mathcal{V}$，我们采用 RESERVE-L，它使用 24 层的 ViT-L/16（Dosovitskiy等人，2021）对每一帧进行独立编码。它得到 $V' \in \mathbb{R}^{n_V \times d}$，其中 n_V表示图像帧序列的长度，而 d=1024表示特征维度。

对于音频模态 $\mathcal{A}$，我们采用音频频谱图 Transformer（Gong等人，2021）对每个子片段进行独立编码。这是通过将音频划分为 5 秒的时间间隔来实现的。其得到 $A' \in \mathbb{R}^{n_A \times d}$，其中 n_A 表示分割过程后的子片段数量。

对于对话 $\mathcal{D}$，$\mathcal{D}$ 的上下文表示为 $D' \in \mathbb{R}^{n_D \times l_D \times d_E}$。其中，$n_D$ 表示对话话语的数量，l_D 表示单个话语中的词数。具体而言，话语中每个词的表示均由 Merlot Reserve 的嵌入矩阵初始化。

为表示人格 $\mathcal{P}$，我们首先基于其离散分布构建与人格特质相对应的词汇表。人格嵌入矩阵采用随机初始化，并在训练中进行优化。我们可以选择使用单个人格特质，例如 [ESFP], , 或使用它们的集合，例如 [ESFP; SLUAN; 7w8]。最终，我们获得人格特征 $P' \in$

$\mathbb{R}^{n_P \times l_P \times d_E}$ 的相关角色在 $\mathcal{VC}$中。这里， n_P 表示 $\mathcal{VC}$中的相关角色数量， l_P 表示 P的长度。

多项选择 $\mathcal{C}$ 的表示形式为 $C \in \mathbb{R}^{n_C \times l_C \times d_E}$。其中， n_C 表示多项选择选项的数量，l_C 分别表示单个选项中的词数。具体而言，话语中每个词的表示均由 Merlot Reserve 的嵌入矩阵初始化。

4.2 个性化多模态推理

首先，将候选行为与多模态特征和人格特征相结合。具体而言，我们将 P 和 C 拼接后输入答案注意力。我们对 P_C进行平均池化，得到最终的 P^{out}。同样，对 D 和 A 进行处理，分别得到 D^{out} 和 A^{out}。

随后，将多模态特征与个性特征进行融合。我们首先分别将P^{out} 与 D^{out} 和 A^{out}拼接。然后，融合多模态信息以获得模态答案。最后，我们利用Softmax层推导最终结果如下：

$$Ans' = [D^{out}; P^{out}] + [A^{out}; P^{out}], \quad (2)$$

$$Ans = \text{Softmax}(Ans'). \quad (3)$$

除了上述专为该任务设计的方法外，我们还对相关任务（TVQA、PMR）中的若干方法进行了修改和调整，以适应我们的任务。详情见第6.2节。

5 扩展：基于预测人格的以人为中心的多模态推理

如前所述，在现实生活中获取人格标注颇具挑战。为克服这一挑战，我们提出了一项扩展任务，称为个性预测的人本多模态推理（T^2）。

如图3所示，鉴于多模态信息可用，我们首先预测最可能的人格特质 $\mathcal{P}'$（某一特定个体在 $\mathcal{VC}$中的）。我们通过以下过程来完成：

$$\mathcal{P}' = \text{Personality-Prediction}(\mathcal{D}, \mathcal{V}, \mathcal{A}). \quad (4)$$

其次，我们将 T^1 中的标注人格 $\mathcal{P}$替换为预测的人格 $\mathcal{P}'$，并重新评估其性能，方法是

$$\mathcal{B} = \text{PHMR}(\mathcal{D}, \mathcal{V}, \mathcal{A}, \mathcal{P}'). \quad (5)$$

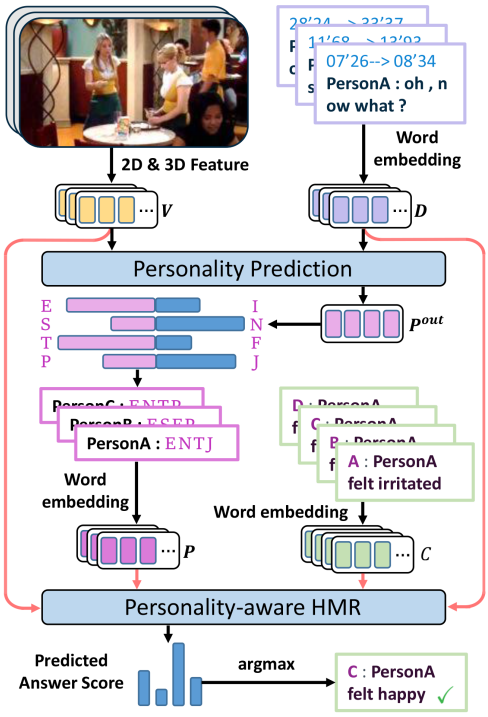

图 3：用于人格预测和人格感知 HMR 的 PRM $_{pretrain}$ 。

	训练集	Dev	Test	总计
# 样本	8152	1746	1747	11645
# 人格	3.29	3.25	3.27	3.28

表4：我们的 MPPD数据集的基本统计信息。

为了准确预测人格特质，并有效利用这些预测进行未来预测，必须确保 MPPD 与 PHMRD 完全分离。采取这一预防措施是为了避免数据泄漏，确保 MPPD 与 PHMRD 保持独立且互不重叠。为此，我们将 TVQA数据集中的所有片段都视为 MPPD 中的片段。随后，我们排除同时出现在 PHMRD 中的所有片段。

此外，我们从数据集中剔除信息不足的角色，例如 Person *0*，他们只会打招呼或一直吃东西。由此，我们得到了与每个片段相关的角色列表。根据这些角色获得相应的人格特质后，我们用姓名标签替换个体的姓名，刻意打破姓名与人格特质之间的关联。这一预防措施可以避免模型记住特定姓名与相应人格特质之间的联系。数据集的统计详

情见表 4.

6 实验

6.1 实验设置

PHMRD和 MPPD数据集按14:3:3 的比例划分为训练集、验证集和测试集。我们使用准确率作为主要评估指标。

我们在四块RTX 3090 GPU 上训练我们的 **PRM** 模型。对于所有模型，我们使用早停训练 20 个epoch。我们将特征融合中的线性层学习率设为$8e-5$，将人格特征提取中的人格注意力学习率设为 $4e-5$，其余部分的学习率设为 $1e-5$，并采用线性预热过程。我们报告四次运行的平均性能。

6.2 对比方法

我们建立了三种基于规则的方法、三种适配的多模态推理模型、一种模态分离变体和两种大型视频模型，作为与我们的**PRM** 模型进行比较的方法。

基于规则的方法。**随机**，**最长**和**最短** 分别选择随机、最长和最短选项。

自适应方法。**FVTA 自适应** (Liang 等人, 2018年) 专注于局部对齐视频与文本模态。**TVQA-adapt** (Lei 等人, 2018年) 整合人格和模态信息，以获取每个模态的特征。**Reserve-adapt** (Zellers 等人,2022年) 在预训练期间将视频与文本和音频融合，并将其与个性特征相结合，以生成最终预测结果。

多模态大语言模型。**Chat- UniVi** (Jin 等人, 2023年）采用视觉模型将图像转换为嵌入表示，并将对话和人格转换为提示。**Gem-ini**（Anil 等人，2023年）使用 16 张图像，并结合对话、人格和多项选择进行提示。

模态分离方法。**PRM**$_{vanilla}$ 是**PRM**的一种变体。它为三种模态中的每一种生成独立表示。这种方法使我们能够更轻松、更有效地进行模态消融实验。为应对在 Merlot Reserve 实验中观察到的将视频模态特征整合到音频和文本模态中所带来的挑战，我们还引入了一种原始方法。

模型	不含优先级经验回放	带优先级经验回放	Imp.
最短	16.70	-	-
随机	20.00	-	-
最长	**34.08**	-	-
FVTA 自适应	33.64	34.35	0.71
TVQA-adapt	**43.37**	43.92	0.55
PRM$_{vanilla}$	43.03	**44.14**	**1.11**
Reserve-adapt	52.43	53.07	0.64
PRM	**53.13**	**54.06**	**0.93**

表5：基于规则的方法、深度学习方法和视频大模型的准确率。优先级经验回放表示人格，而改进表示通过融入人格信息所实现的增强效果。

模型	无优先级经验回放	使用优先级经验回放	Imp.
Chat-UniVi	23.50	20.61	-2.89
Gemini	33.17	34.08	0.91

表6：零样本多模态大语言模型的准确率。

6.3 主要结果

表5汇报了两种设置下不同对比方法的结果。在有监督设置中，与 TVQA-adapt 相比，虽然**PRM**$_{vanilla}$ 模型在未纳入人格信息时表现不佳，但整合人格信息后取得了最佳表现。与 Reserve-adapt 相比，**PRM** 模型的表现显著提升。这凸显了人格信息对该任务的有效性，也表明我们的模型能够更有效地利用人格信息。

在零样本设置下，如表6所示，现有多模态大语言模型在 PHMR 任务中面临挑战，其表现逊于微调预训练模型。与 Chat-UniVi 相比，Gemini 通过纳入人格信息提升了表现。这种提升可能归因于 Gemini 的训练语料库纳入了人格与行为之间的对应关系，例如纳入了用于人格预测的面向任务的数据集。

6.4 不同人格类型的结果

表7 展示了不同人格类型实验得出的结果。实验结果表明，三种人格的组合表现最佳，其次是迈尔斯-布里格斯类型指标 和 FFM，它们的表现达到

人格	准确率	Imp.
基线	53.13	-
含迈尔斯-布里格斯类型指标	54.03	0.90
含FFM	54.04	0.91
含九型人格	53.89	0.76
含全部	**54.06**	**0.93**

表 7：不同人格下**PRM**的准确率。

模态	不含优先级经验回放	与优先级经验回放	Imp.
对话	51.02	51.27	0.25
音频	52.82	53.55	0.73
All	**53.13**	**54.06**	**0.93**

表 8：**PRM**不同模态的准确率。All 表示输入了所有模态信息。

具有可比性的结果。

统计显著性实验得到的 p 值低于 5%，表明纳入人格信息所带来的改进具有统计显著性。

6.5 不同模态的结果

如表8所示，我们基于**PRM**在各种模态组合下开展了实验。在单模态设置中，无论是基础性能还是人格带来的增强效果，音频的表现都优于对话。多模态结果高于单模态结果，且在人格增强方面的提升更为显著。

如表 9所示，**PRM**$_{vanilla}$ 呈现了不同模态的结果。音频模态的表现较低，因此削弱了原始方法在所有模态下的整体效果。然而，引入人格信息在多模态场景下带来了最显著的提升。

模态	不使用优先级经验回放	使用优先级经验回放	Imp.
对话	43.08	43.37	0.29
音频	36.59	36.70	0.11
视频	43.74	44.23	0.49
All	43.03	44.14	**1.11**

表 9：不同模态下**PRM**$_{vanilla}$的准确率。

测试领域	不含优先级经验回放	使用优先级经验回放	Imp.
《老友记》	52.92	53.40	0.48
BBT	55.57	56.23	0.66
《灵书妙探》	48.44	49.34	0.90
Met	56.14	59.12	**2.98**
House	48.68	49.49	0.81
Grey	57.65	58.85	1.20

表10：不同电视节目的 **PRM** 准确率。

模型	HL$_{\downarrow}$	RL$_{\downarrow}$	AP$_{\uparrow}$
PRM	25.52	25.71	77.68

表11：人格预测的结果。在此，HL 表示汉明损失，RL 表示排序损失，AP 表示平均精度。

6.6 不同电视节目上的结果

为了考虑六个电视节目中剧集类型和数量的多样性，我们将其划分为不同分区，以评估引入人格信息的影响。鉴于每个电视节目可用的剧集数量有限，分别单独训练模型被证明是一项挑战。因此，我们选择在完整数据集上训练模型，并在这六个电视节目的测试集上评估其性能。相应的结果见表 10。结果表明，在每种类型的电视节目中引入人格都会带来改进，并且对于片段数量相对较少的电视节目，改进幅度更大。

6.7 扩展任务上的实验 T^2

我们使用 **PRM** 模型在 MPPD数据集上训练并预测人格特质。具体而言，我们预测迈尔斯-布里格斯类型指标（MBTI），并将预测出的 MBTI 应用于 T^1。

人格预测的效果 在表中11，汉明损失和排序损失越低越好，而平均精度越高越好。结果表明，预测个性

模型	已标注	预测值	Dif.
PRM	54.03	53.82	**-0.21**

表 12：使用预测人格重新训练与使用标注值重新训练的结果差异。差异表示两者之间的差异。

可以作为真实值的合理近似。

人格预测 PHMR 的效果 为验证预测人格的有效性，我们用预测人格替代标注人格，并从头开始训练 **PRM** 以 T^1 。表 12展示了模型表现，结果仅下降了 0.21%。研究结果表明，可以利用多模态信息预测人格，从而缓解现实生活中人格标注不足的问题。

7 相关工作

7.1 多模态推理

多模态推理以单模态推理为基础，例如 VCR（Zellers 等人，2019年）、VCG（Park et al.，2020年）和 NExT-QA（Xiao et al.，2021），这些任务同时涵盖以物体为中心和以人为中心的推理。

以物体为中心的推理任务主要关注视频中的物体，较少涉及人物内容。AGQA（Grunde-McLaughlinetal.，2021）利用视频和关系图预测物体关系。Sherlock（Hessel et al.，2022年）根据图像中的现有线索推断缺失信息。TVQA（Lei 等人，2018年）根据视频、对话、问题和时间跨度选择答案，其中一些问题与人物有关。

以人为中心的推理任务中，人物往往在视频里只出现一次，因此在缺乏同一人大量样本的情况下很难获取人格信息。Social-IQ（Zadeh等人，2019年）聚焦于利用视频、对话以及行为与心理相关问题的选择题。WHYCAT（Ignat等人，2021）基于视频和描述评估动机。PMR（Dong等人，2022年）利用图像和问题选择最优回答。

上述任务提供有关问题的信息，而其他任务则在没有这些信息的情况下进行推理。这些任务通常不涉及个体的心理活动。VLEP（Lei 等人，2020年）根据视频和文本描述预测最可能的下一步行动。VAR（Liang 等人，2022年）基于视频和周围语境推断图像和描述，遵循 αNLI（Bhagavatula 等人，2020年），这是一项仅专注于文本模态的任务。

7.2 人格计算

人格计算（Yang 等人，2022年a）是一个研究领域，利用计算技术，从文本、多媒体和社交网络等各种来源研究人格。目前，研究中主要采用三种人格评估指标：五因素模型（FFM）（McCrae 和 Costa，1987年）、迈尔斯-布里格斯类型指标（MBTI）（Myers 等人，1985年）以及九型人格 (Palmer，1988年）。FFM的五个维度描述可预测的外显行为，而MBTI的四个维度主要解释行为，且与本能密切相关。九型人格 概述了个体的九种核心动机，每种动机都有其自身的思维、情感和行为模式。

与人格计算相关的工作主要有两个分支：预测型人格和应用型人格。人格预测任务的发展历程如下。早期研究（Flekova和Gurevych，2015）采用了更简单的二分类评价指标。后续研究（Ponce-López等人，2016年；Kampman等人，2018年；Aslan和Güdükbay，2019年）旨在建立对大五人格特质的预测。近年来，相关研究（Gjurkovic和Snajder，2018年；Stajner和Yenikent，2021；Sang等人，2022年）已将重点转向迈尔斯-布里格斯类型指标。也有一些研究（Zhang等人，2022年；Yang等人，2022b；Tayarani等人，2022年；Plepi等人，2022年）关注在实际应用中利用人格。

8 结论

在本研究中，我们提出了一项新的个性化以人为中心的多模态推理（PHMR）任务（T^1），并基于六档电视节目构建了一个新数据集 PHMRD。实验结果表明，融入人格信息能够提升以人为中心的多模态推理性能。此外，消融研究发现，三种不同的人格特质对性能提升的贡献程度各不相同。为进一步解决真实场景中缺乏人格标注的问题，我们提出了一项扩展任务，称为人格预测以人为中心的多模态推理（T^2）。实验结果表明，我们的方法能够准确预测人格，并且无需依赖人格标注即可取得令人满意的多模态推理性能。

局限性

个性化以人为中心的多模态推理是一项具有挑战性的任务。本研究是对该任务的初步探索；该任务并未被定义为生成任务。

受篇幅限制，本文重点介绍任务和数据集。所提出的基线相当简单，仍有很大的改进空间。

人格计算是一个相对宽泛的领域，我们的文献综述可能未能全面涵盖该领域的研究。

伦理声明

我们谨向Lei 等人（2018年）在 TVQA 方面的宝贵工作表示感谢。TVQA 根据麻省理工学院的一项许可协议授权，该协议允许出于任何目的对材料进行商业使用、修改、分发和私人使用。我们也将在之后公开发布 PHMRD。人格数据库（PDB）网站是一个开源网站，任何人无需注册即可获取信息。所有数据集和模型均为英语，这对英语使用者更有利。我们聘请了两名具有自然语言处理经验的研究生来核验 ChatGPT 的结果。我们向研究生支付每小时约 7 至 10 美元，远高于当地平均工资；如果他们对整个过程有顾虑，我们也会与他们进行建设性讨论。

参考文献

Rohan Anil, Sebastian Borgeaud, Yonghui Wu, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M. Dai, Anja Hauth, Katie Mil- lican, David Silver, Slav Petrov, Melvin Johnson, Ioannis Antonoglou, Julian Schrittwieser, Amelia Glaese, Jilin Chen, Emily Pitler, Timothy P.Lilli-crap, Angeliki Lazaridou, OrhanFirat, James Molloy, Michael Isard, Paul Ronald Barham, Tom Hennigan, Benjamin Lee, Fabio Viola, Malcolm Reynolds, Yuanzhong Xu, Ryan Doherty, Eli Collins,Clemens Meyer,Eliza Rutherford, Erica Moreira, Kareem Ayoub, MeghaGoel, George Tucker, Enrique Pi-queras, Maxim Krikun, Iain Barr, Nikolay Savinov, Ivo Danihelka, Becca Roelofs, Anaïs White, Anders Andreassen, Tamara von Glehn, Lakshman Yagati, Mehran Kazemi, Lucas Gonzalez,Misha Khalman, Jakub Sygnowski,以及其他作者。2023年Gemini：一系列能力极强的多模态模型。 *CoRR*, abs/2312.11805。

Süleyman Aslan 和 Ugur Güdükbay。2019年。使用长短期记忆网络和卷积神经网络进行基于视频的多模态外显人格识别。 *CoRR*，abs/1911.00381。

Chandra Bhagavatula、Ronan Le Bras、Chaitanya Malaviya、Keisuke Sakaguchi、Ari Holtzman、Hannah Rashkin、DougDowney、Wen-tau Yih 和 Yejin Choi。2020年。溯因常识推理。载于第八届国际学习表征会议，*ICLR 2020*，埃塞俄比亚亚的斯亚贝巴，*2020*年*4*月*26*日至*30*日。OpenReview.net。

Laura Ana Maria Bostan、Evgeny Kim 和 Roman Klinger。2019年。Goodnewseveryone：一个标注了情感、语义角色和读者感知的新闻标题语料库。收录于第*12*届语言资源与评估会议论文集，*LREC 2020*，法国马赛，*2020*年*5*月*11*日至*16*日，第1554–1566页。欧洲语言资源协会。

Jacob Devlin、Ming-Wei Chang、Kenton Lee 和 Kristina Toutanova。2019年。BERT：面向语言理解的深度双向Transformer预训练。收录于*2019*年北美计算语言学协会分会会议：人类语言技术论文集，*NAACL-HLT 2019*，美国明尼阿波利斯，*2019*年*6*月*2*日至*7*日，第*1*卷（长短篇论文），第4171–4186页。计算语言学协会。

Qingxiu Dong、Ziwei Qin、Heming Xia、Tian Feng、Shoujie Tong、Haoran Meng、Lin Xu、Zhongyu Wei、Weidong Zhan、Baobao Chang、Sujian Li、Tianyu Liu 和 Zhifang Sui。2022年。基于前提的多模态推理：基于联合文本与视觉线索的条件推理。载于第*60*届计算语言学协会年度会议论文集（第*1*卷：长篇论文），*ACL 2022*，爱尔兰都柏林，*2022*年*5*月*22*日至*27*日，第932–946页。计算语言学协会。

Alexey Dosovitskiy、Lucas Beyer、Alexander Kolesnikov、Dirk Weissenborn、Xiaohua Zhai、Thomas Unterthiner、Mostafa Dehghani、Matthias Minderer、Georg Heigold、Sylvain Gelly、Jakob Uszkoreit 和 Neil Houlsby。2021年。一张图像胜过16×16个词：用于大规模图像识别的Transformer。载于第*9*届国际学习表征会议，*ICLR 2021*，线上活动，奥地利，*2021*年*5*月*3*日至*7*日。OpenReview.net。

Christoph Feichtenhofer、Haoqi Fan、Jitendra Malik 和 Kaiming He。2019年。用于视频识别的SlowFast网络。载于*2019 IEEE/CVF* 国际计算机视觉会议，*ICCV 2019*，韩国首尔，*2019*年*10*月*27*日至*11*月*2*日，第6201–6210页。IEEE。

Lucie Flekova 和 Iryna Gurevych。2015。利用词汇资源之间的义项级链接对虚构人物进行性格画像。载于*2015* 年自然语言处理实证方法会议论文集，*EMNLP 2015*，葡萄牙里斯本，*2015*年*9*月*17*日至*21*日，第1805–1816页。计算语言学协会。

Matej Gjurkovic 和 Jan Snajder。2018年。Reddit：人格预测的金矿。发表于 .....论文集

第二届社交媒体中人们观点、人格与情感计算建模研讨会，*PEOPLES@NAACL-HTL 2018*，美国路易斯安那州新奥尔良，*2018*年*6*月*6*日，第87–97页。计算语言学协会。

Yuan Gong、Yu-An Chung 和 James R. Glass。2021。AST：音频频谱图Transformer。载于 *Interspeech2021*，国际语音通信协会第*22*届年会，布尔诺，捷克，*2021*年*8*月*30*日至*9*月*3*日，第571–575页。ISCA。

Madeleine Grunde-McLaughlin、Ranjay Krishna 和 Maneesh Agrawala。2021。AGQA：组合时空推理基准测试。载于 *IEEE* 计算机视觉与模式识别会议，*CVPR 2021*，线上，*2021*年*6*月*19*日至*25*日，第11287–11297页。计算机视觉基金会 / IEEE。

Kaiming He、张祥雨、任少卿和孙剑。2016年。用于图像识别的深度残差学习。载于 *2016*年*IEEE*计算机视觉与模式识别会议，*CVPR 2016*年，拉斯维加斯，内华达州，美国，*2016*年*6*月*27*日至*30*日，第770–778页。IEEE计算机学会。

Jack Hessel、Jena D. Hwang、Jae Sung Park、Rowan Zellers、Chandra Bhagavatula、Anna Rohrbach、Kate Saenko 和 Yejin Choi。2022年。夏洛克·福尔摩斯绑架案：用于视觉溯因推理的数据集。载于计算机视觉——*ECCV 2022*——第*17*届欧洲会议，*2022*年*10*月*23*日至*27*日，以色列特拉维夫，论文集，第*XXXVI*部分，第13696卷，《计算机科学讲义》，第558–575页。施普林格。

Oana Ignat、Santiago Castro、Hanwen Miao、Weiji Li 和 Rada Mihalcea。2021。Whyact：识别生活方式视频博客中的行动原因。载于*2021*年自然语言处理实证方法会议论文集，*EMNLP 2021*，线上活动／蓬塔卡纳，多米尼加共和国，*2021*年*11*月*7*日至*11*日，第4770–4785页。计算语言学协会。

Peng Jin, Ryuichi Takanobu, Caiwan Zhang, Cao Xiaochun和Li Yuan. 2023年. Chat-UniVi：统一视觉表征赋能大语言模型的图像与视频理解. *CoRR*,abs/2311.08046.

Onno Kampman、Elham J. Barezi、Dario Bertero 和 PascaleFung。2018年。 探究音频、视频和文本融合方法在端到端自动人格预测中的应用。载于 第*56*届计算语言学协会年会论文集，*ACL2018*，澳大利亚墨尔本，*2018*年*7*月*15*日至*20*日，第*2*卷：短篇论文，第606–611页。计算语言学协会。

RanjayKrishna、Kenji Hata、FredericRen、李飞飞和Juan Carlos Niebles。2017a。 视频密集事件描述。载于 *IEEE*国际计算机视觉会议，*ICCV 2017*，威尼斯，意大利，*2017*年*10*月*22*日至*29*日，第706–715页。IEEE计算机学会。

Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A. Shamma, Michael S. Bernstein, 和李飞飞. 2017b. 视觉基因组：利用众包密集图像标注连接语言与视觉. 国际计算机视觉期刊,123(1):32–73.

Ranjay Krishna、Yuke Zhu、Oliver Groth、Justin John- son、Kenji Hata、Joshua Kravitz、Stephanie Chen、Yannis Kalantidis、Li-Jia Li、David A. Shamma、Michael S. Bernstein 和 Li Fei-Fei。2017c。视觉基因组：使用众包的密集图像注释连接语言与视觉。国际计算机视觉杂志，123(1)：32–73。

Jie Lei、Licheng Yu、Mohit Bansal 和 Tamara L. Berg。2018年。TVQA：局部化、组合式视频问答。载于 *2018*年自然语言处理实证方法会议论文集，布鲁塞尔，比利时，*2018*年*10*月*31*日至*11*月*4*日，第1369–1379页。计算语言学协会。

Jie Lei、Licheng Yu、Tamara L. Berg 和 Mohit Bansal。2020年。接下来更有可能发生什么？视频和语言未来事件预测。载于 *2020*年自然语言处理实证方法会议论文集，*EMNLP 2020*年，在线，*2020*年*11*月*16*日至*20*日，第8769–8784页。计算语言学协会。

Linjie Li、Yen-Chun Chen、Yu Cheng、Zhe Gan、Licheng Yu 和 Jingjing Liu。2020年。HERO：视频语言全能表示的分层编-码器+预-训练。载于*2020*年自然语言处理经验方法会议论文集，*EMNLP 2020*，在线，*2020*年*11*月*16*日至*20*日，第2046–2065页。计算语言学协会。

Chen Liang、Wenguan Wang、TianfeiZhou 和 YiYang。2022年。视觉溯因推理。载于 *IEEE/CVF*计算机视觉与模式识别会议，*CVPR2022*，美国路易斯安那州新奥尔良，*2022*年*6*月*18*日至*24*日，第15544–15554页。IEEE。

Junwei Liang、Lu Jiang、Liangliang Cao、Li-Jia Li和Alexander G. Hauptmann。2018年。 面向视觉问答的焦点视觉-文本注意力。载于*2018*年*IEEE*计算机视觉与模式识别会议，*CVPR 2018*，盐湖城，犹他州，美国，*2018*年*6*月*18*日至*22*日，第6135–6143页。计算机视觉基金会 / IEEE计算机学会。

Robert R McCrae 和Paul TCosta。1987年。人格五因素模型跨测量工具和观察者的验证。《人格与社会心理学杂志》，52(1):81。

Isabel Briggs Myers、Mary H McCaulley 和 Robert Most。1985年。《手册：迈尔斯-布里格斯类型指标的开发与使用指南》。咨询心理学家出版社。

Helen Palmer。1988年。《九型人格：了解自己和生命中的他人》。Harper & Row。

Jae Sung Park、Chandra Bhagavatula、Roozbeh Mot- taghi、Ali Farhadi 和 YejinChoi。2020年。视觉彗星：对静态图像动态上下文的推理。载于 计算机视觉 - *ECCV 2020*- 第*16*届欧洲会议，英国格拉斯哥，*2020*年*8*月*23*日至*28*日，论文集，第*V*部分，收录于《计算机科学讲义》，第12350卷，第508–524页。施普林格。

Rebecca J. Passonneau。2006年。用于语义和语用标注的集合值项目一致性度量（MASI）。收录于第五届国际语言资源与评估会议论文集，*LREC 2006*，意大利热那亚，*2006*年*5*月*22*日至*28*日，第831–836页。欧洲语言资源协会（ELRA）。

Joan Plepi、Béla Neuendorf、Lucie Flek 和 CharlesWelch。2022年。统一数据视角主义与个性化：社会规范的应用。载于*2022*年自然语言处理实证方法会议论文集，*EMNLP 2022*，阿布扎比，阿拉伯联合酋长国，*2022*年*12*月*7*日至*11*日，第7391–7402页。计算语言学协会。

Víctor Ponce-López、Baiyu Chen、Marc Oliu、Ciprian A. Corneanu、Albert Clapés、Isabelle Guyon、Xavier Baró、Hugo JairEscalante 和 Sergio Escalera。2016年。ChaLearn LAP 2016：第一印象挑战赛第一轮——数据集和结果。载于计算机视觉—*ECCV 2016*研讨会—阿姆斯特丹，荷- 兰，*2016*年*10*月*8*日至*10*日及*15*日至*16*日，论文集，第三部分，第9915卷《《计算机科学讲义》，第400–418页。

Yisi Sang、Xiangyang Mou、Mo Yu、Dakuo Wang、Jing Li、Jeffrey Stanton。2022年使用电影剧本进行虚构角色的迈尔斯-布里格斯类型指标人格预测。*CoRR*，abs/2210.10994。

Rico Sennrich、Barry Haddow 和 AlexandraBirch。2016年。使用子词单元进行低频词的神经机器翻译。载于计算机语言学协会第*54*届年会论文集，*ACL 2016*，*2016*年*8*月*7*日至*12*日，柏林，德国，第*1*卷：长篇论文。计算机语言学协会。

Xindi Shang、Donglin Di、Junbin Xiao、Yu Cao、Xun Yang 和 Tat-Seng Chua。2019年。对用户生成视频中的对象与关系标注。载于 *2019*年国际多媒体检索会议论文集，*ICMR 2019*，渥太华，安大略省，加拿大，*2019*年*6*月*10*日至*13*日，第279–287页。ACM。

Sanja Stajner 和 SerenYenikent。2021。为什么从文本中进行迈尔斯-布里格斯类型指标人格检测是一项困难的任务？载于第*16*届欧洲计算语言学协会会议论文集：主会场，*EACL 2021*，在线，*2021*年*4*月*19*日至*23*日，第3580–3589页。计算语言学协会。

Mohammad Tayarani、Anna Esposito 和 Alessandro Vinciarelli。2022年。一个“呃”能透露你的哪些信息：利用量子进化特征选择算法将填充词映射为人格特质。*IEEE* 情感计算汇刊，13(1)：108–121。

YonghuiWu, Mike Schuster, Zhifeng Chen, Quoc V.Le, Mohammad Norouzi,Wolfgang Macherey,Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Lukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, GeorgeKurian, Nishant Patil, Wei Wang, CliffYoung, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, MacduffHughes, and Jeffrey Dean. 2016年。Google 的神经机器翻译系统：弥合人类翻译与机器翻译之间的差距。 *CoRR*, abs/1609.08144。

Junbin Xiao、Xindi Shang、Angela Yao 和 Tat-SengChua。2021年。NExT-QA：解释时间动作的问题回答的下一阶段。载于 *IEEE*计算机视觉与模式识别会议，*CVPR 2021*，线上，*2021*年*6*月*19*日至*25*日，第 9777–9786 页。计算机视觉基金会 / IEEE。

LiangYang、Shuqun Li、XiLuo、BoXu、Yuanling Geng、Zeyuan Zeng、Fan Zhang和Hongfei Lin。2022年a。计算人格：综述。 软计算，26(18):9587–9605。

Qi Yang、Sergey Nikolenko、Alfred Huang 和 Alek- sandr Farseev。2022年b。人格驱动的社交多媒体内容推荐。载于*MM '22*：第*30*届*ACM*国际多媒体会议，葡萄牙里斯本，*2022*年*10*月*10*日至*14*日，第7290–7299页。ACM。

Amir Zadeh、Michael Chan、Paul Pu Liang、Edmund Tong 和 Louis-Philippe Morency。2019年。Social-iq：人工社交智能的问答基准。载于 *IEEE*计算机视觉与模式识别会议，*CVPR 2019*，美国加利福尼亚州长滩，*2019*年*6*月*16*日至*20*日，第8807–8817页。计算机视觉基金会 / IEEE。

Rowan Zellers、Yonatan Bisk、Ali Farhadi和 Yejin Choi。2019年。 从识别到认知：视觉常识推理。载于*IEEE*计算机视觉与模式识别会议，*CVPR 2019*，美国加利福尼亚州长滩，*6*月*16-20*日，*2019*年，第6720–6731页。计算机视觉基金会 / IEEE。

Rowan Zellers、Jiasen Lu、Ximing Lu、Youngjae Yu、YanpengZhao、Mohammadreza Salehi、AdityaKusu- pati、JackHessel、Ali Farhadi 和 Yejin Choi。2022年 MERLOT RESERVE：通过视觉、语言和声音获得的神经脚本知识。载于*IEEE/CVF*计算机视觉与模式识别会议，*CVPR2022*，美国路易斯安那州新奥尔良，*2022*年*6*月*18*日至*24*日，第 16354–16366 页。IEEE。

Le Zhang、Songyou Peng 和 Stefan Winkler。2022年。Persemon：用于联合分析表观人格、情绪及其关系的深度网络。*IEEE* 情感计算汇刊，13(1)：298–305。

Min-LingZhang 和 Zhi-Hua Zhou。2014年。多标签学习算法综述。*IEEE Transactions on Knowledge and Data Engineering*，26(8):1819–1837。

ChatGPT 提示词与若干示例

A.1 数据过滤

提示词：“{什么颜色的物体？}”被视为与个体无关，而“{人*0*为什么喜欢这个物体？}”则被认为与人类受试者相关。“{QUESTION}”是否与人类行为或心理学有关?

同样，“人*0*腿上是什么物体？”被认为与个体无关，而“人*0* 在谈论什么？”则被视为与人类受试者相关。

B 人格类型学

B.1 迈尔斯-布里格斯类型指标

迈尔斯-布里格斯类型指标（MBTI）是一种内省式自我报告问卷，用于人格类型学，旨在识别人们看待环境和做出判断时的各种心理偏好。

为便于评估，迈尔斯-布里格斯类型指标将人格划分为四个维度（E/I、S/N、T/F 和 J/P）。这四个维度各有不同侧重点，具体如下:

- 关注注意力，将外向型（E）和内向型（I）作为区分，体现人们主要的能量来源。
- 关注认知，将判断（J）和感知（P）作为区分，说明人们如何收集信息。
- 关注判断，将直觉（N）和实感（S）作为区分，说明个体如何做出决策。
- 关注生活方式，将思考（T）和情感（F）作为区分，说明人们的行为方式。

每个维度包含两个相互对立的标签，这些标签可以组合成16种人格，如图所示。4（c）。

16种人格类型各有其自身的特征，如表 13。

B.2 FFM

五因素模型 (FFM)，也称为大五人格模型，是用于理解人格与学业行为之间关系的模型。该模型由几组相互独立的研究人员定义，他们对人类行为的语言描述词进行了因素分析。这些研究人员首先研究了大量与人格特质相关的语言描述词之间的关系。

在评估中，FFM 将人格划分为五个维度（I/N、O/U、S/R、A/E 和 L/C）。具体而言，这五个维度各有不同的侧重点，如下所示:

- 经验开放性（富有创造力/好奇（I） vs. 一致/谨慎（N））
- 尽责性（高效/有条理（O） vs. 奢侈/粗心（U））
- 外向性 (外向/精力充沛(S) 与 孤独/矜持(R))
- 宜人性（友善/富有同情心（A）vs. 批判/理性（E））
- 神经质（敏感/紧张（L）vs. 坚韧/自信（C））

如图所示的五种人格维度两极对立的详细名称 4（a）。

B.3 九型人格

九型人格沿着古代图腾的九个角展开，揭示了九种不同的内在动力，使每个人作为个体都具有与生俱来的独特性。九型人格理论所描述的九种人格类型并无好坏之分；不同人格类型的人应对世界的方式存在可辨识的根本差异。如今，人们普遍认识到人格属于我们自身，它们会过滤并诠释我们所见所闻，而九型人格理论的基本原则是，我们每个人都有九种可能的“过滤器”之一，它将为我们的生活勾勒蓝图，并在我们天性中维持一定的层次，用以形成我们与外部世界的沟通策略。

九种人格的名称，如图4(b)。

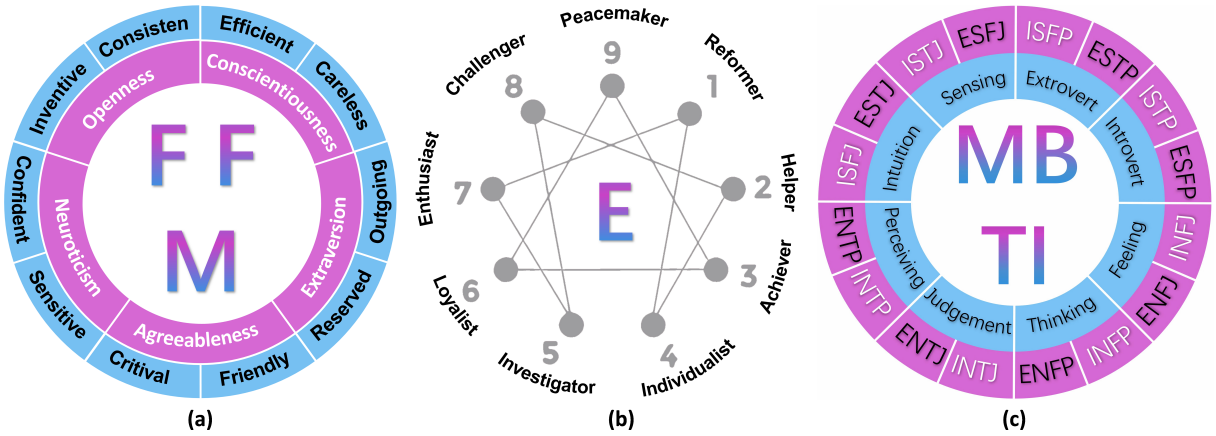

图4：三种人格。

MBTI	Name	描述
ISTJ	检查者	这些人往往内敛却固执，对生活持理性的看法。他们构成 他们的行动，并有条不紊、有目的地加以执行。
ISFJ	保护者	这些人往往以自己沉稳的方式显得温暖而不张扬。他们做事高效。 且富有责任感，在日常生活中仔细关注实际细节。
INFJ	顾问型	他们往往以深刻的思考和想象力来面对生活。他们内在的视觉， 个人价值观，以及一种沉静而有原则的人文主义，指引着他们处理一切事务。
INTJ	策划者	这些深思熟虑的战术家热衷于完善生活的种种细节，运用创造力和理性- 他们做任何事都带有个性。他们的内心世界通常是私密而复杂的。
ISTP	巧匠型	他们往往具有个人主义思维，追求目标而不太需要太多外部 联结。他们以求知欲和个人技能投入生活，采取的方式各不相同 根据需要.
ISFP	作曲家	他们往往思想开放，以脚踏实地的态度面对生活、新体验和他人 亲和力。他们专注于当下的能力帮助他们发掘令人兴奋的潜能。
INFP	治愈者	这些罕见的人格类型往往安静、思想开放且富有想象力，并且它们运用 以关怀且富有创造力的方式对待他们所做的一切。
INTP	建筑师	这些思维灵活的人喜欢以非传统的方式对待生活的诸多方面。 他们常常寻找不同寻常的道路，将尝试新事物的意愿与个人创造力结合起来。
ESTP	促动者	他们往往精力充沛、注重行动，能够灵活应对眼前的一切。 他们热爱发掘生活中的机遇，无论是与他人社交，还是进行更私人的交流。
ESFP	表演者	追求。 这些人热爱充满活力的体验，积极投入生活，并乐于 探索未知。他们往往十分善于社交，也经常鼓励他人一起参与
ENFP	倡导者	各项活动。 这些人往往会拥抱能够体现其希望感的宏大理念与行动， 对他人的善意。他们充沛的活力可以朝多个方向流动。
ENTP	发明家	他们往往大胆而富有创造力，善于在头脑中解构并重建各种想法 敏捷性。即使遇到任何阻力，他们也会积极追求自己的目标。
ESTJ	主管	他们拥有极大的坚韧，坚定地遵循自己明智的判断。他们 常常在人群中发挥稳定作用，能够在逆境中提供坚定的指引。
ESFJ	供给者	他们体贴周到、以人为本，并乐于参与社交群体。 他们的成就由坚定的价值观指引，并乐于为他人提供指导。
ENFJ	教师	这些热情、直率的类型喜欢帮助他人，他们往往有强烈的想法和 价值观。他们以创造性的能量支持自己的观点，以实现其目标。
ENTJ	统帅	他们是果断的人，热爱势头与成就。他们收集信息 以构建其创造性愿景，但在付诸行动之前很少长时间犹豫。

表13：迈尔斯-布里格斯类型指标的十六种类型及其描述。

九型人格 描述**完美型** 强调原则，不轻易妥协，明辨是非，对自己和他人都有很高的要求，追求完美。**助人型** 渴望与他人建立良好关系，以人为本，愿意迁就他人。**成就型** 具有竞争意识，以成就衡量自身价值，是工作狂。**自我型** 情绪化，害怕被他人拒绝，觉得他人不理解自己，坚持我行我素。**理想型** 喜欢思考和分析，求知欲强，但行动力不足，对物质生活要求较低。**怀疑型** 做事谨慎，不容易信任他人，疑虑很多，喜欢群体生活，并且努力工作。**活跃型** 乐观，喜欢新鲜事物，喜欢追随潮流，不喜欢压力。**领导型** 追求权力，强调实力，不依赖他人，并且富有正义感。**和平型** 做决定需要很长时间，害怕争执，祈求和谐共处。

表14：迈尔斯-布里格斯类型指标的16种类型及其描述。