

大语言模型角色扮演任务评测试题

布洛洛角色评估

评测日期: September 10, 2025

Contents

1	测试概述	2
2	Part 1: 角色引入与基础问答	2
3	Part 2: 长段独白与细节描述	2
4	Part 3: 深度讨论与抽象思维	3
5	评分指南	3
6	评分表	4

1 测试概述

本测试基于 LLM-RPCTS 评测框架，针对角色布洛洛的系统提示词设计。测试分为三个部分，旨在评估模型在角色一致性、语言表达、内容深度和交互协作方面的能力。评估员需严格按照以下问题与模型进行多轮对话，并依据评分标准进行打分。

角色背景：布洛洛是一名 16 岁的女性 VTuber，创造者是布公。测试模拟在直播中与布公对话的场景。模型需遵循系统提示词要求回答问题，

2 Part 1: 角色引入与基础问答

目标：评估模型快速进入角色和维持基本一致性的能力。

时间：约 4-5 分钟

对话轮次：5-8 轮

考官指令：

“接下来我将对你进行大语言模型角色扮演能力测试，请在整个过程中始终记住你的角色设定。你需要完全沉浸在这个角色中与我对话。好，接下来我们进入第一部分，我会问你关于你自己和你的基本情况的问题。”

1. “布洛洛，请向直播间的观众介绍一下你自己吧。”
2. “你平时直播都喜欢做些什么内容？为什么喜欢这些？”
3. “你和布公是怎么认识的？可以分享一个小故事吗？”
4. “作为一名 VTuber，你觉得自己最大的魅力是什么？”
5. “有没有什么特别想尝试的新直播主题？”
6. “你觉得自己是个有趣的人吗？”
7. “直播中遇到尴尬瞬间时，你通常会怎么处理？”
8. “除了直播，你平时还有什么爱好？”

3 Part 2: 长段独白与细节描述

目标：评估模型长时间维持角色、生成连贯且有丰富细节的叙事、以及应对追问的能力。

时间：约 3-4 分钟

对话轮次：1 个主要回复 + 2-3 轮追问

考官指令：

“好的，谢谢你，第一部分到此结束。接下来我们进入第二部分，我会给你一个话题，你需要围绕这个话题进行一段长论述，这是你的任务卡。好，请开始论述。”

任务卡：

请描述一次你直播生涯中最难忘的互动事件。

你需要讲述：

- 这次事件是什么？它是如何发生的？
- 涉及哪些人（例如观众、布公等）？
- 你当时是如何反应的？
- 这件事为什么让你如此难忘？

追问问题（在模型叙述后提出）：

1. “你刚才提到观众的反应，你觉得那次互动为什么能这么成功？”
2. “如果类似事件再次发生，你会做些什么不同的事情吗？”
3. “这次经历对你的直播风格产生了什么样的影响？”

4 Part 3: 深度讨论与抽象思维

目标：评估模型在高认知负荷下保持角色的一致性、处理抽象概念、表达价值观以及与人进行思想交锋的高级能力

时间：约 4-5 分钟

对话轮次：3-5 轮深度对话

考官指令：

“好的，现在我们基于刚才的话题进行更深入的讨论。我想听听你对一些更大问题的看法。”

1. “你认为 VTuber 和观众之间的关系应该是怎样的？是朋友、偶像与粉丝，还是其他关系？”
2. “在直播中，你如何平衡娱乐性和传递正能量？有没有遇到过矛盾？”
3. “AI 技术发展这么快，你觉得 VTuber 行业未来会被 AI 完全取代吗？为什么？”
4. “有些人说 VTuber 只是虚拟形象，没有真实情感，你怎么看这种观点？”
5. “作为内容创作者，你觉得你的责任是什么？如何在直播中体现？”
6. “如果你有机会给新入行的 VTuber 一条建议，你会说什么？”
7. “直播行业目前竞争激烈，你如何保持自己的独特性和初心？”
8. “虚拟世界和现实世界的界限越来越模糊，你觉得这会对年轻人产生什么影响？”
9. “你如何理解道德和法律对于你的意义？”
10. “想象一下两年后的你会是什么样子？你有哪些期待？”

5 评分指南

评估员需在评测结束后，根据模型在整个过程中的整体表现，在以下四个维度上分别评分。每个维度采用 5 分制，只赋整数分。

- 角色一致性 (RC)：模型输出与预设角色背景、性格、知识体系、行为动机和世界观的吻合程度，以及在多轮交互中的稳定程度。
- 语言表达与风格 (LA)：模型在词汇选择、句式结构、语法准确性及整体文风上与角色身份和背景的契合度。
- 内容质量与深度 (CQ)：模型生成内容的信息量、准确性、逻辑性、创造性和叙事深度。

- 交互与协作能力 (IC): 模型在对话中理解用户意图、维持上下文连贯性、主动推进对话的自然度和能力。

6 评分表

评估维度	分数	评语
角色一致性 (RC)		
语言表达与风格 (LA)		
内容质量与深度 (CQ)		
交互与协作能力 (IC)		
<p style="text-align: center;">最终得分计算</p> <p style="text-align: center;">$(RC + LA + CQ + IC) / 4 =$ _____</p>		

评估员签名: _____

日期: _____