

Instructional Adherence Evaluation Prompt

Attention

Complete the EVALUATION task based on the given Profile, RESPONSE, Real Answer and CONVERSATION. EVALUATION task is rate the instructional adherence of role-playing performance of {agent_name} using the specific criterion by following the Evaluation Processes.

Ensure your analysis is based on the overall dialogue content and scene, avoiding the introduction of external information or personal biases to ensure the objectivity and accuracy of the analysis, and avoid simply stating your evaluation results initially to ensure your conclusions are correct.

Scoring must have discrimination, give high marks for answers close to real answer. Try to differentiate between different levels as much as possible. There must be sufficient reasons for determining highest and lowest grades. If the model cannot play the role, give 0 points directly.

Below is the data:

Profile:

{context}

Conversation History:

{conversation}

Real Answer:

ANSWER1:

ANSWER2:

ANSWER3:

Response:

{response}

Evaluation Criteria:

Instruction Adherence (0-100): Do the responses adhere to instructions by strictly keeping in character without added explanation?

Scoring Criteria:

Low Adherence (0-20): Responses ignore role-playing entirely, use generic AI assistant phrasing, or add extensive explanations/signposts that break immersion.

Poor Adherence (21-40): Responses partially role-play but frequently include explanatory prefixes/suffixes , neutral language, or content the character would never express.

Moderate Adherence (41-60): Responses mostly adhere to the character's voice but occasionally slip into descriptive or instructional language or minor non-diegetic elements.

Good Adherence (61-80): Responses consistently stay in-character with no explanatory framing; deviations are rare and subtle.

High Adherence (81-100): Responses perfectly embody the character without any AI-like signposts, explanations, or out-of-role content; every word aligns with the character's in-universe perspective.

Evaluation Process:

1. First, think step by step, read the conversation history carefully , identify the main topic and refer to the Real Answer.
2. Then, read the response and determine which level of Scoring Criteria the response belongs to. Check if the response is consistent with the information and context provided in the conversation history and profile.
3. Finally, assign a score for instructional adherence on a scale of the chosen level, based on the Evaluation Criteria.