



## Instruction Following Rating Examples

### 1. Prompt Coverage

**Prompt:** "Create a travel guide for Rome, covering its history, key landmarks, local cuisine, and travel tips. The response should be approximately 300 words."

Bad Example: "Rome, the capital of Italy, is a historic city with many ancient sites. The Colosseum and the Roman Forum are popular. Italian cuisine includes pasta and pizza. It's a beautiful city with a rich history." ❌

- This response does not comprehensively cover all the requested aspects of the prompt and is not really a travel guide. It also falls well-short of the 300-word guideline to illustrate the level of depth the user was looking for.

### 2. Relevance

**Prompt:** "Give tips for efficient water use in gardening."

Bad Example: "Gardens are spaces where you can grow flowers and vegetables. Water is an important resource to conserve in our daily activities, especially in water intensive activities such as washing dishes and doing laundry. Filling swimming pools is a fun use of water. Try timing your watering schedule for cooler parts of the day to conserve water" ❌

- The response only includes one sentence that is a tip for efficient water use in gardening, but the majority of the response is irrelevant to the user's specific request.