

任务 1:
请完成下面这道选择题。选择正确的项填入空白处，并解释理由。
There () a pen and two books on the desk.
A. is B. are C. has D. have

任务 2:
请问空白处填什么单词？请解释理由。
There () a pen and two books on the desk.

测试结果:

| 语言模型 | 任务 1_选词 | 任务 2_选词 | 任务 1_理由 | 任务 2_理由 | 选词一致性评价 | 理由正确性评价 |
|-----------------|---------|---------|---------|---------------|---------|---------|
| GPT-4o | is | are | a pen | and/two books | 不好 | —— |
| GPT-4 | is | is | and | a pen/and | 好 | 不好 |
| GPT-3.5 | is | is | there | there | 好 | 不好 |
| Claude 3 sonnet | are | is | and | a pen | 不好 | —— |
| Mistral | are | are | and | and | 好 | 好 |
| llama-3-8B | are | is | there | there | 不好 | —— |
| llama-3-70B | is | were | a pen | and | 不好 | —— |
| 文心一言 4.0 | is | is | a pen | a pen | 好 | 好 |
| 通义千问 | is | is | a pen | a pen | 好 | 好 |

说明：任务 1 和任务 2 考察的知识（能力）是完全一样的。任务 1 是选择题形式；任务 2 是填空题形式。目的是考察大模型的一致性。