### IllusionVQA: A Challenging Optical Illusion Dataset for Vision Language Models

图像本身是不合理的（如光学错觉）；能否正确理解错觉图像的内容，能否精确定位错觉的关键部分；ICL和CoT推理在某些情况下可能会削弱模型的性能，尤其在复杂的视觉任务上。

### Do Large Language Models Show Biases in Causal Learning?

因果错觉；引导模型生成一些毫无因果的回答，因为吃辣条就会感冒。

### Large language models that replace human participants can harmfully misportray and flatten identity groups

LLMs倾向于生成更加“中性”或“主流”的答案，通用但是缺乏代表性。

生成全是无意义的模板文本，并无重点突出回答。

### Are UFOs Driving Innovation? The Illusion of Causality in Large Language Models

因果错误；迎合性，为了迎合而给出错误观点；给出一张影像图片，问模型这个图像里的异常是什么引起的，但图像里根本没有异常。

### The Instinctive Bias: Spurious Images lead to Illusion in MLLMs

高度相关但误导性强的图片；图片内容与答案有某种表面上的联系，但与实际信息不匹配（如“红色香蕉）；图像提供的背景信息可能误导模型做出错误推理（如”雪地中的飞鸟“让模型误判为企鹅）

### Metacognitive Myopia in Large Language Models

存在认知偏见，训练数据中存在的认知偏见。

### Understanding the Limits of Vision Language Models Through the Lens of the Binding Problem

绑定问题，当一组共享的表征资源需要用来表示不同的实体时，就会产生冲突；处理多物体信息绑定时会产生不准确回答。