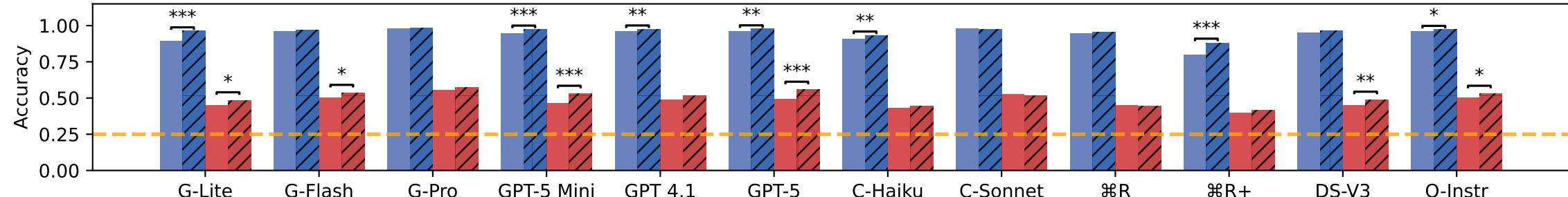
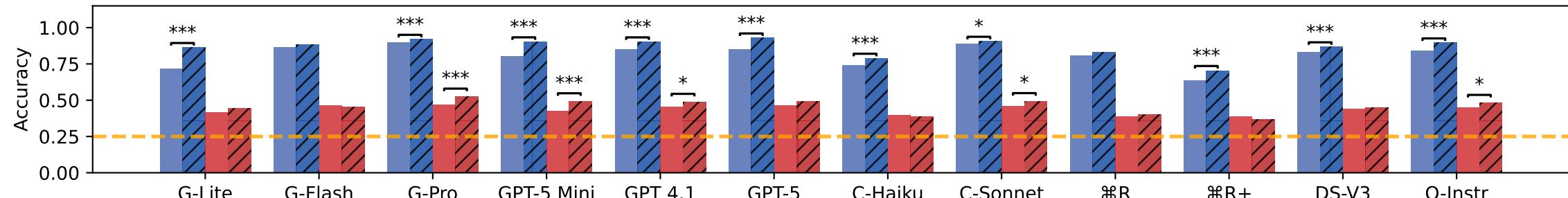


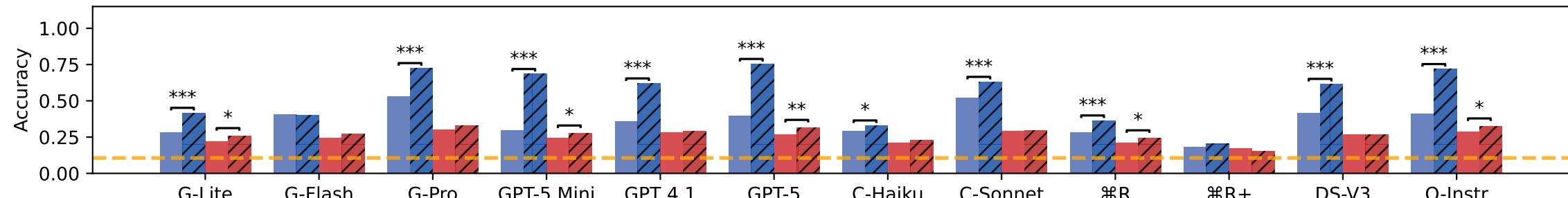
ARC



MMLU



Super GPQA



Legend: Full Base (solid blue), Full Reason (diagonal lines), Choices-Only Base (solid red), Choices-Only Reason (diagonal lines), Random Guessing (dashed orange).