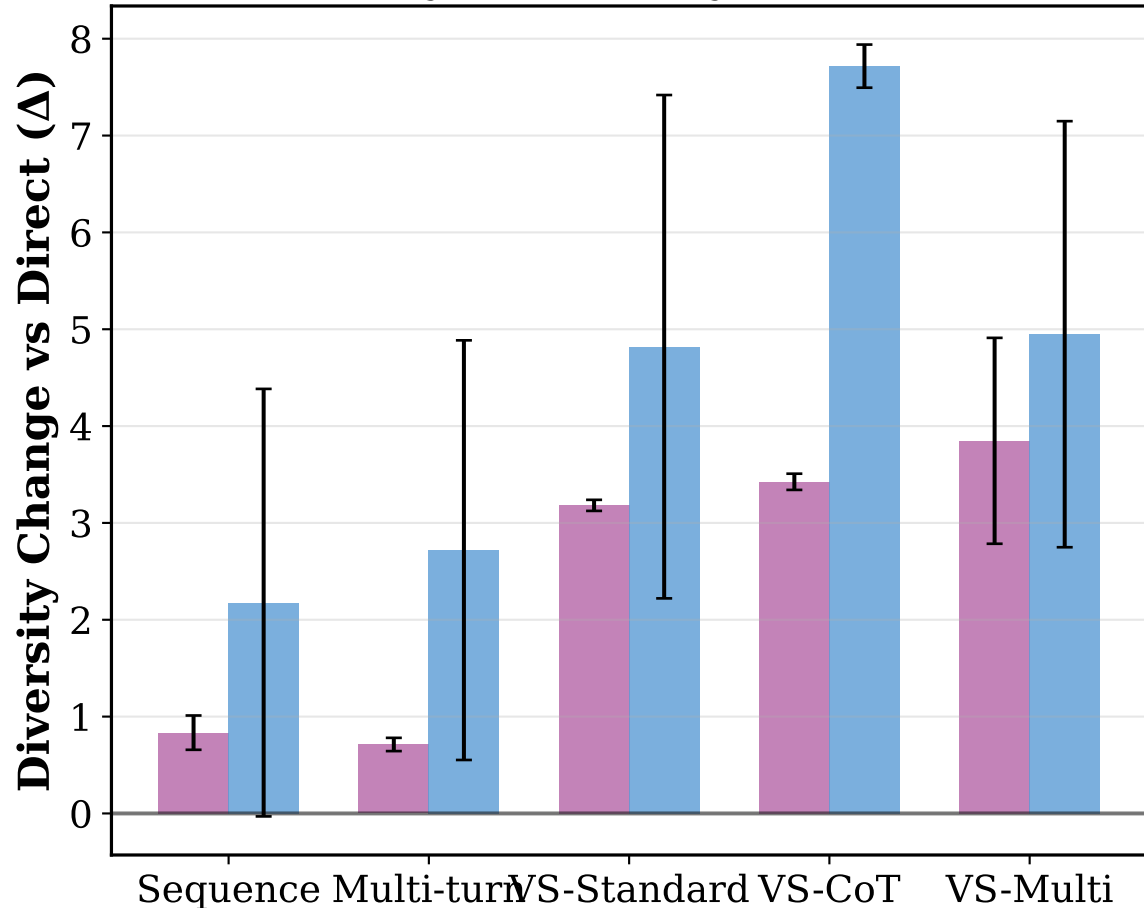


Small Models (GPT-4.1-Mini, Gemini-2.5-Flash)

Large Models (GPT-4.1, Gemini-2.5-Pro)

**(a) Diversity Impact by Model Size**



**(b) Quality Impact by Model Size**

