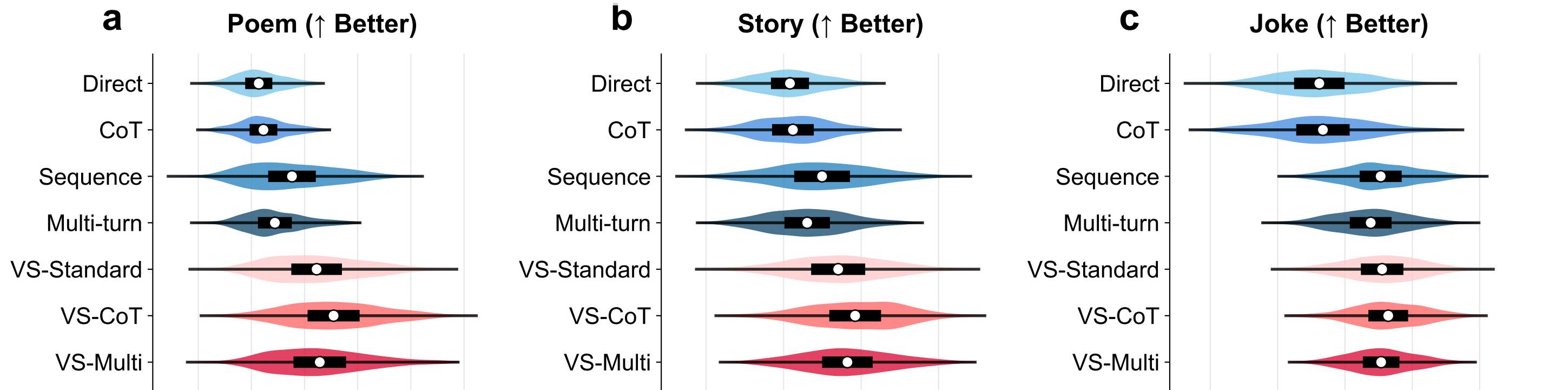
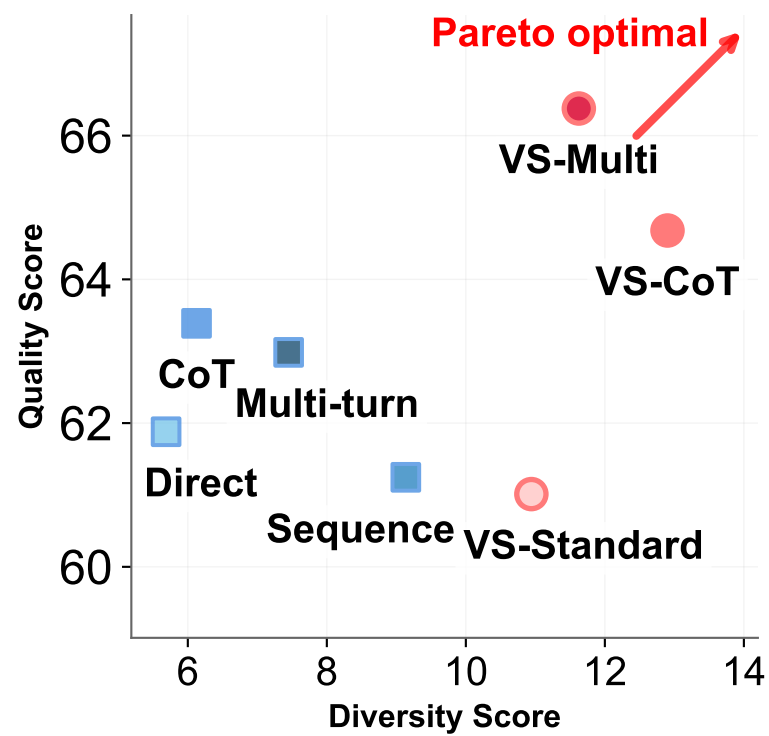


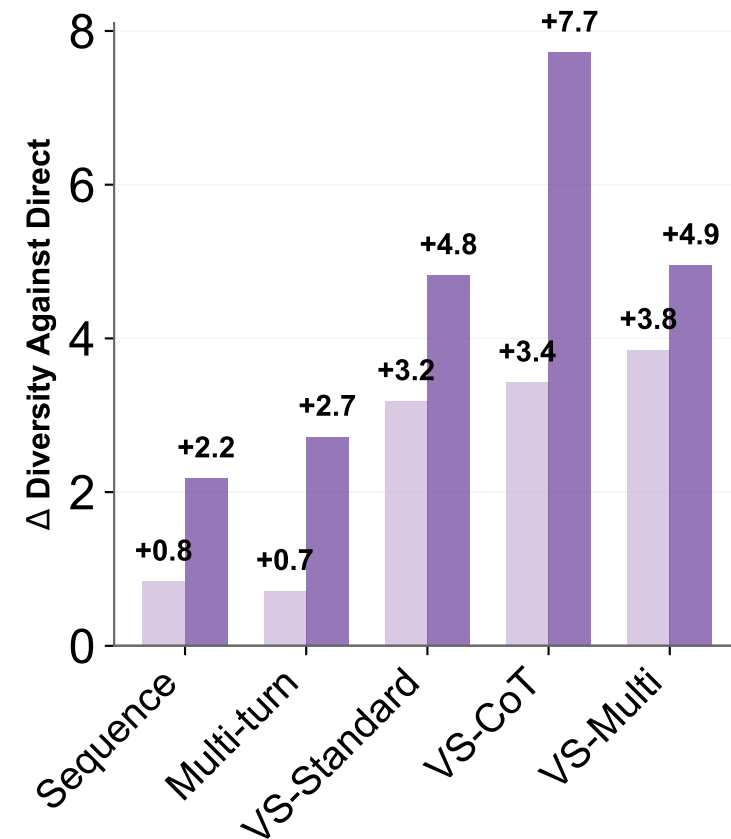
Direct CoT Sequence Multi-turn VS-Standard VS-CoT VS-Multi



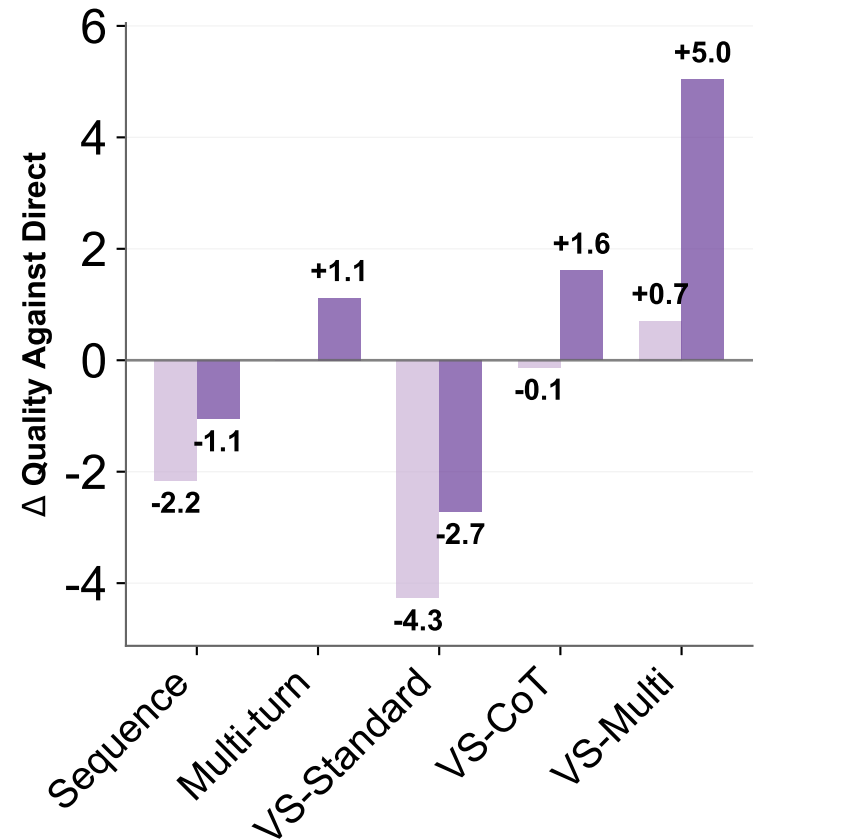
d Diversity vs. Quality (Poem)



e Emergent Trend: Δ in Diversity



f Cognitive Burden: Δ in Quality



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