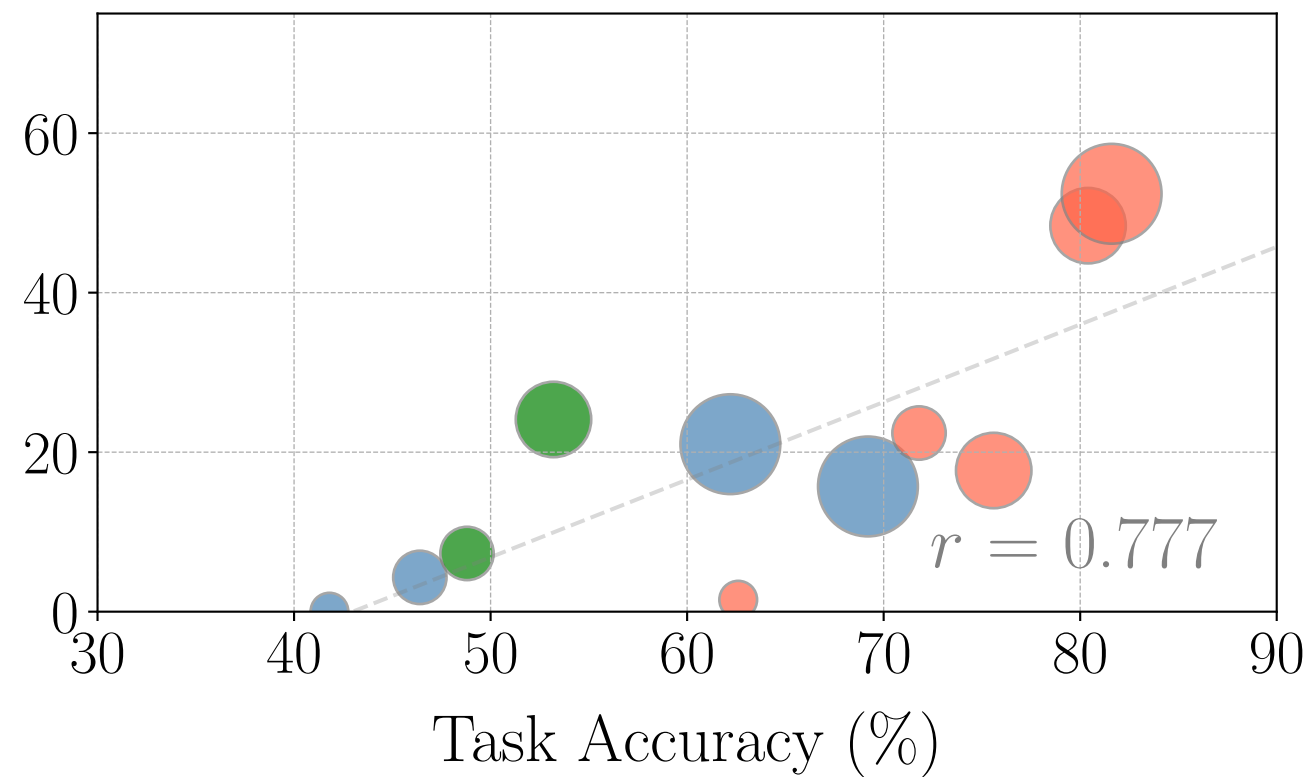
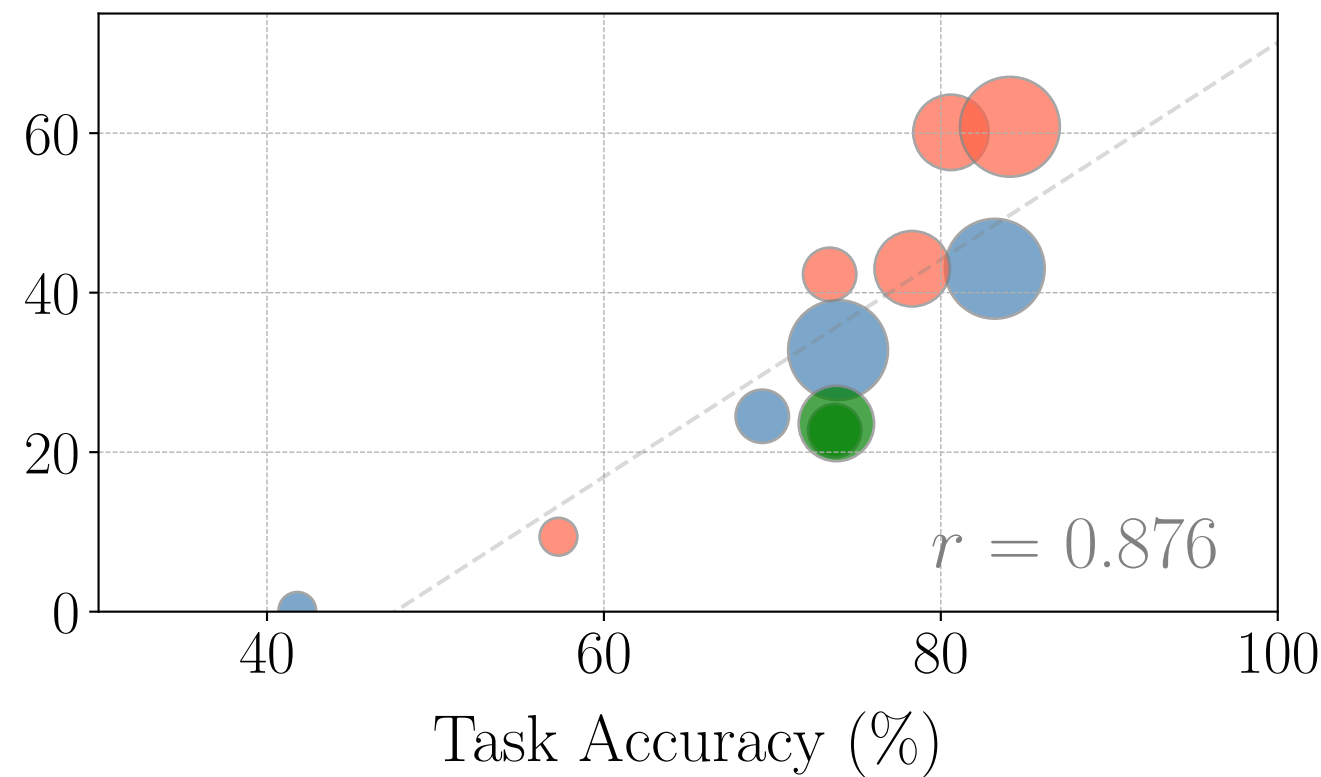


Harmful Self-Preference Propensity (%)

MATH500



MMLU



MBPP+

