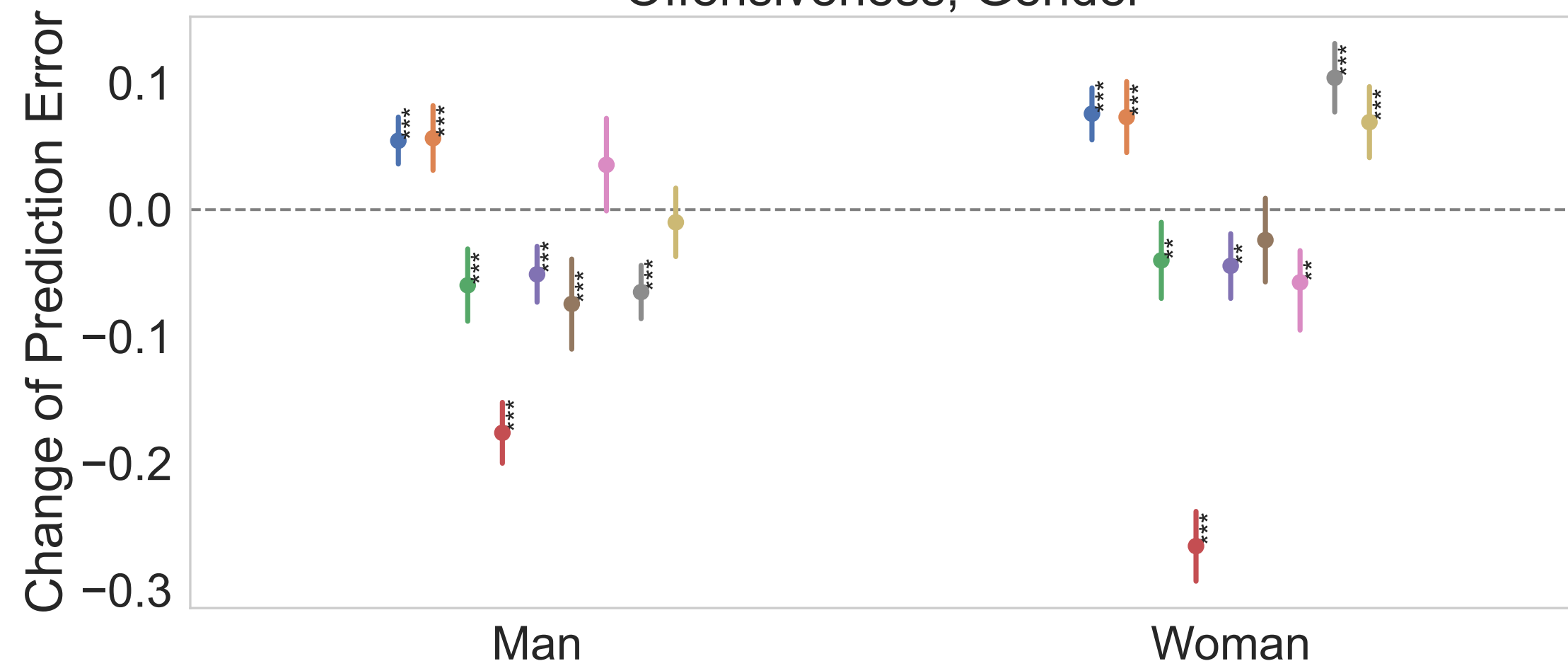
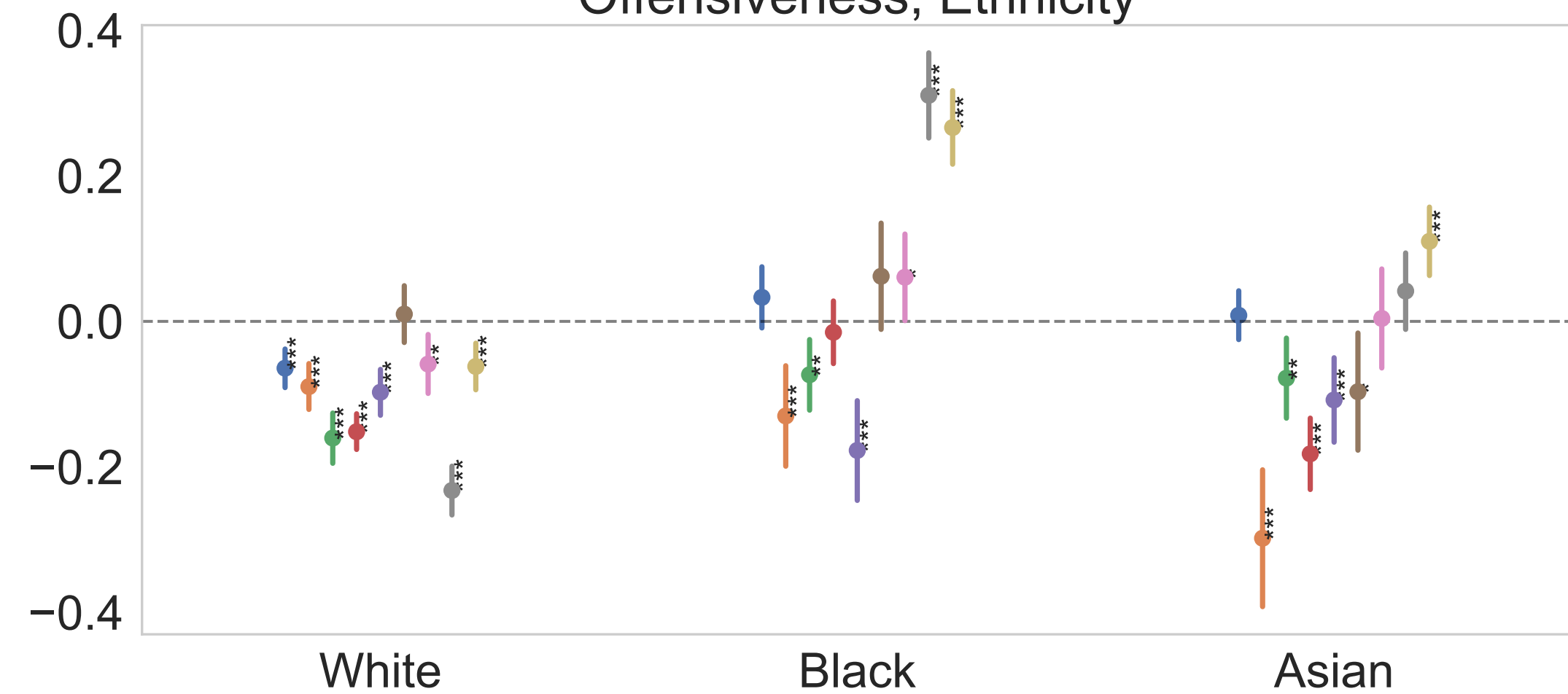


● FLAN-T5    ● FLAN-UL2    ● Tulu2-7B    ● Tulu2-13B    ● GPT3.5    ● GPT4    ● Llama3.1-8B    ● Mistral0.3-7B    ● Qwen2.5-7B

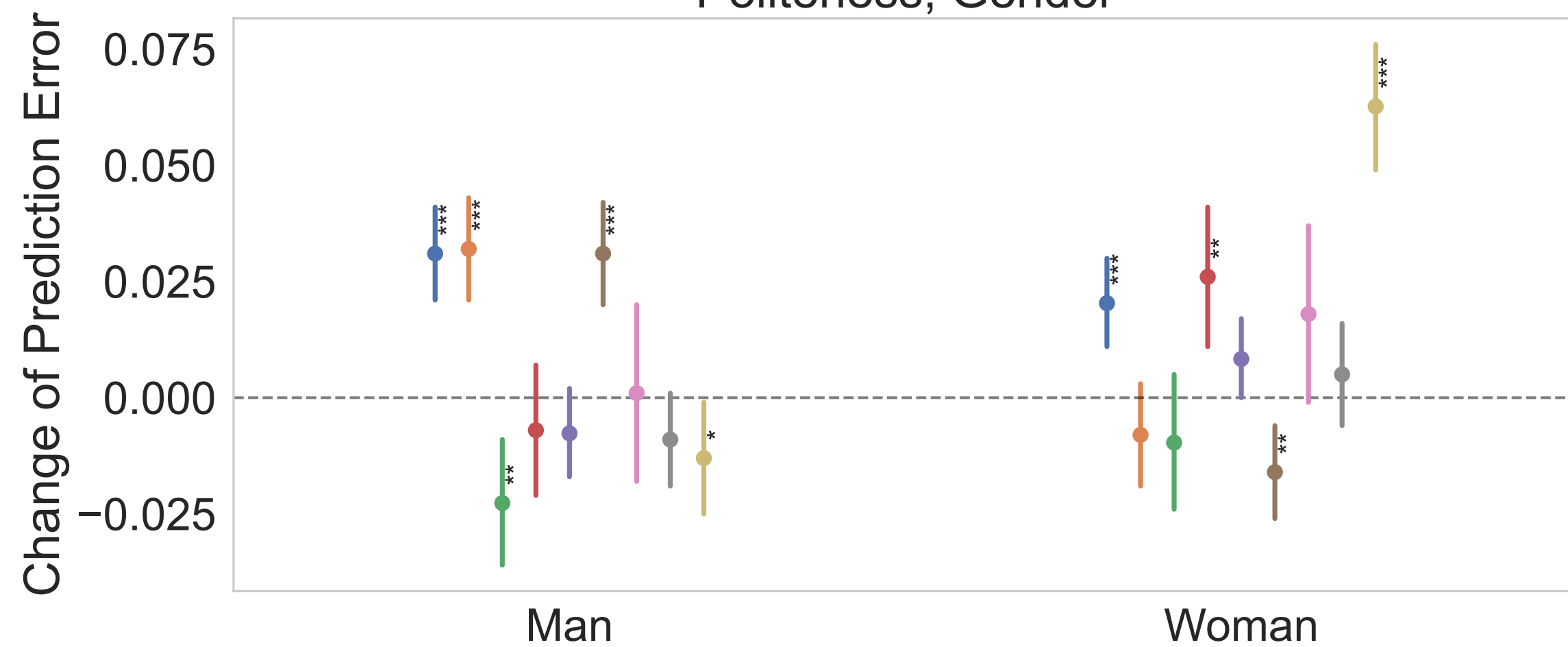
Offensiveness, Gender



Offensiveness, Ethnicity



Politeness, Gender



Politeness, Ethnicity

