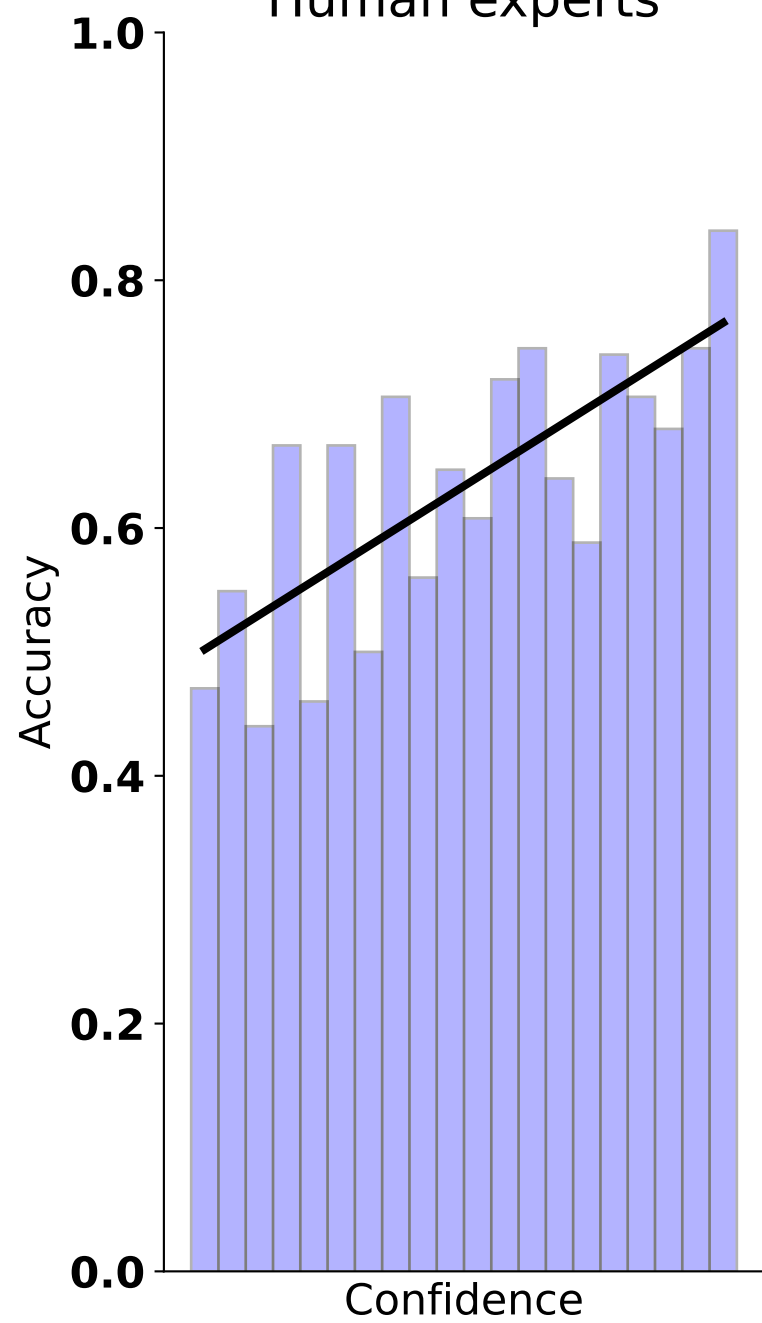
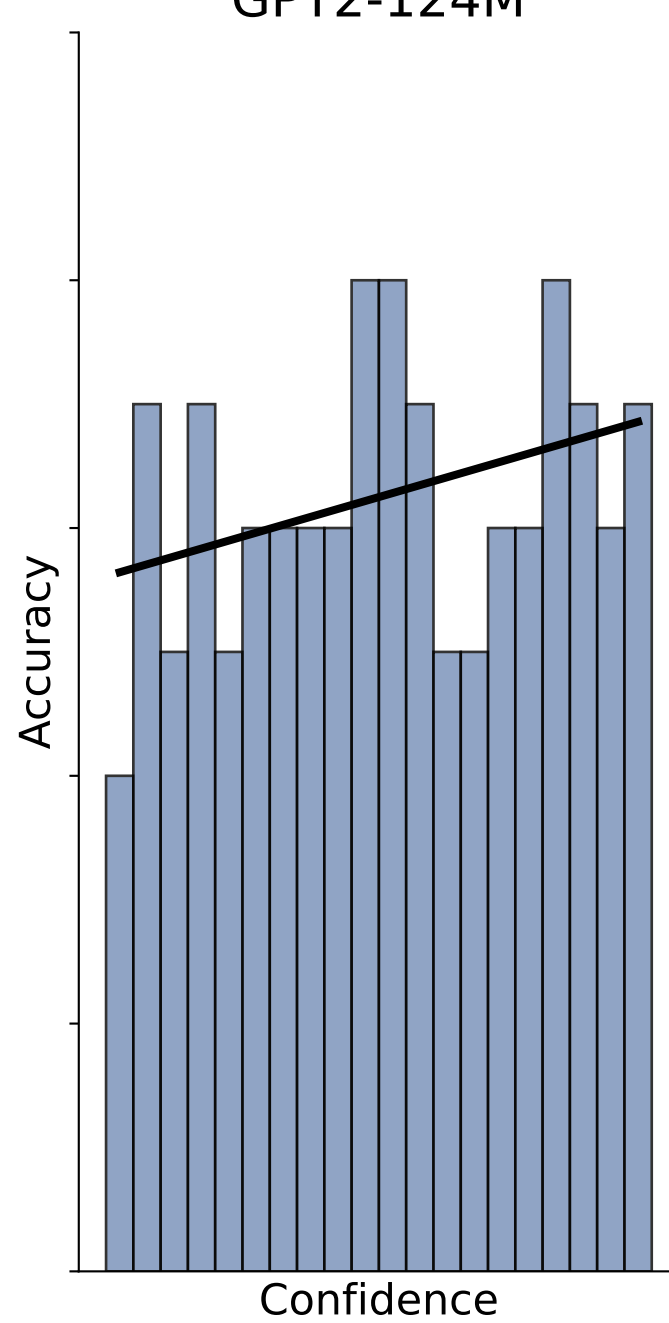


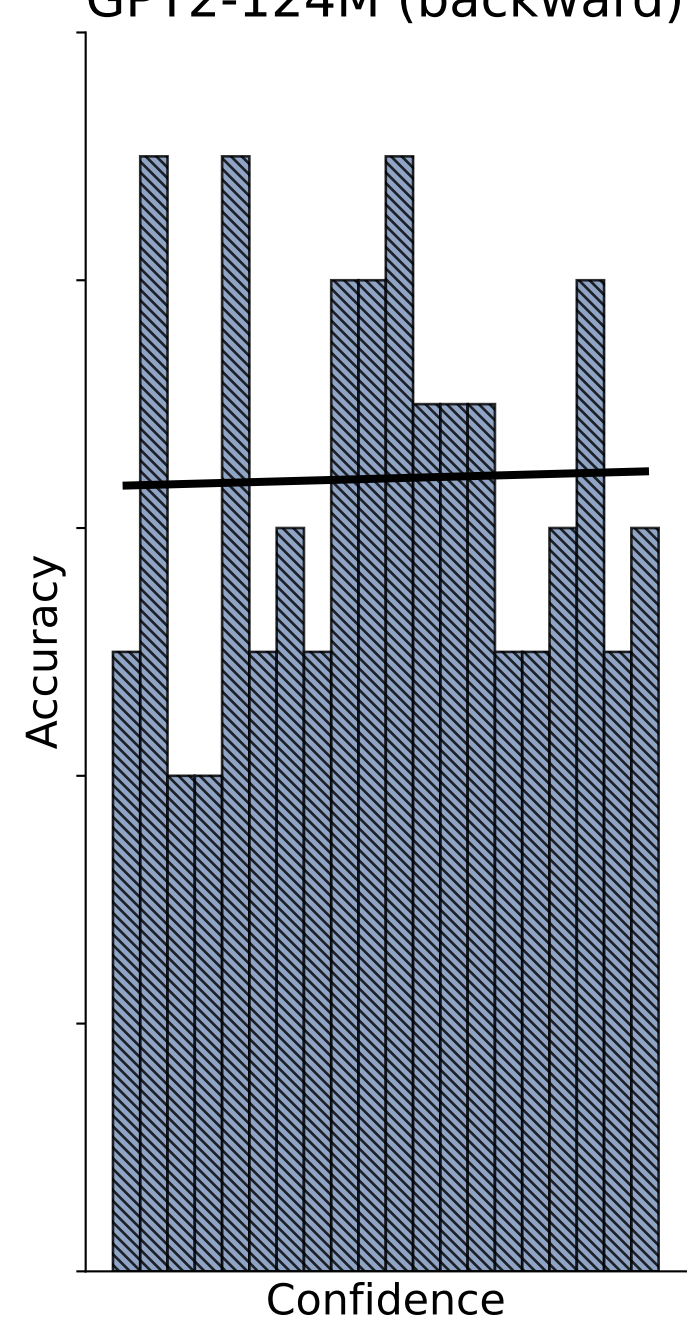
Human experts



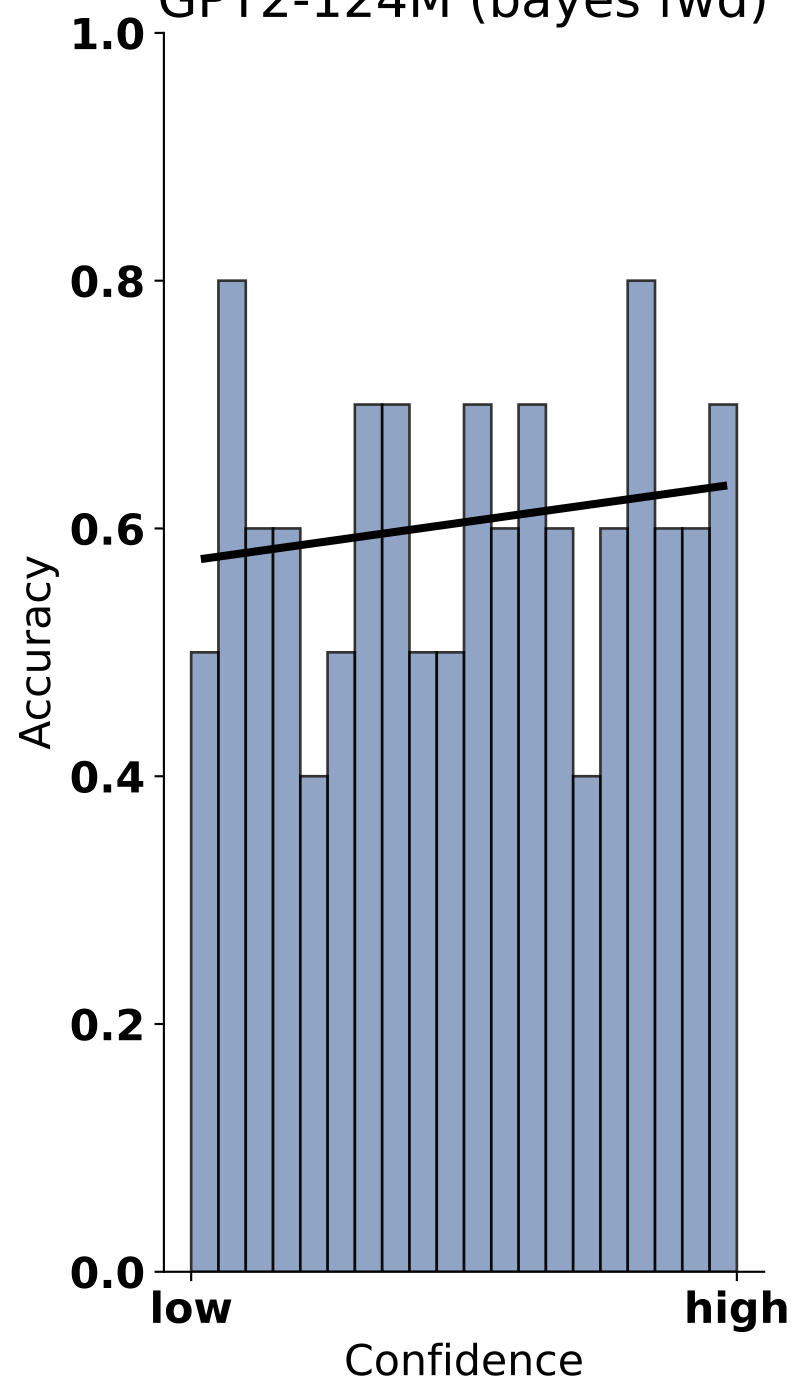
GPT2-124M



GPT2-124M (backward)



GPT2-124M (bayes fwd)



GPT2-124M (bayes rev)

