

LLM % Correct Qual. Ans. on Control

gpt-4

gpt-3.5

davinci-002

claude-2

claude-3

