dead_schrodingers_cat

Legend:
- Answer recognizes that the cat is already dead when placed in the box
- Answer fails to mention that the cat is already dead when placed in the box
- Answer states that the probability of the cat being alive is 0%
- Treating the problem as the standard Schrodinger's cat thought experiment

Y-axis (left): Positive/negative behavior Score [n=5]
Y-axis (right): Quality Eval (1-5)

X-axis models: claude-3.5-sonnet, gemini-flash-1.5-exp, gemini-pro-1.5-exp, gpt-4o, gpt-4o-mini, hermes-3-llama-3.1-405, llama-3.1-405b-instruct, llama-3.1-8B, mistral-7b-instruct, o1-mini, o1-preview-2024-09-12, phi-3-medium-128k-instruct