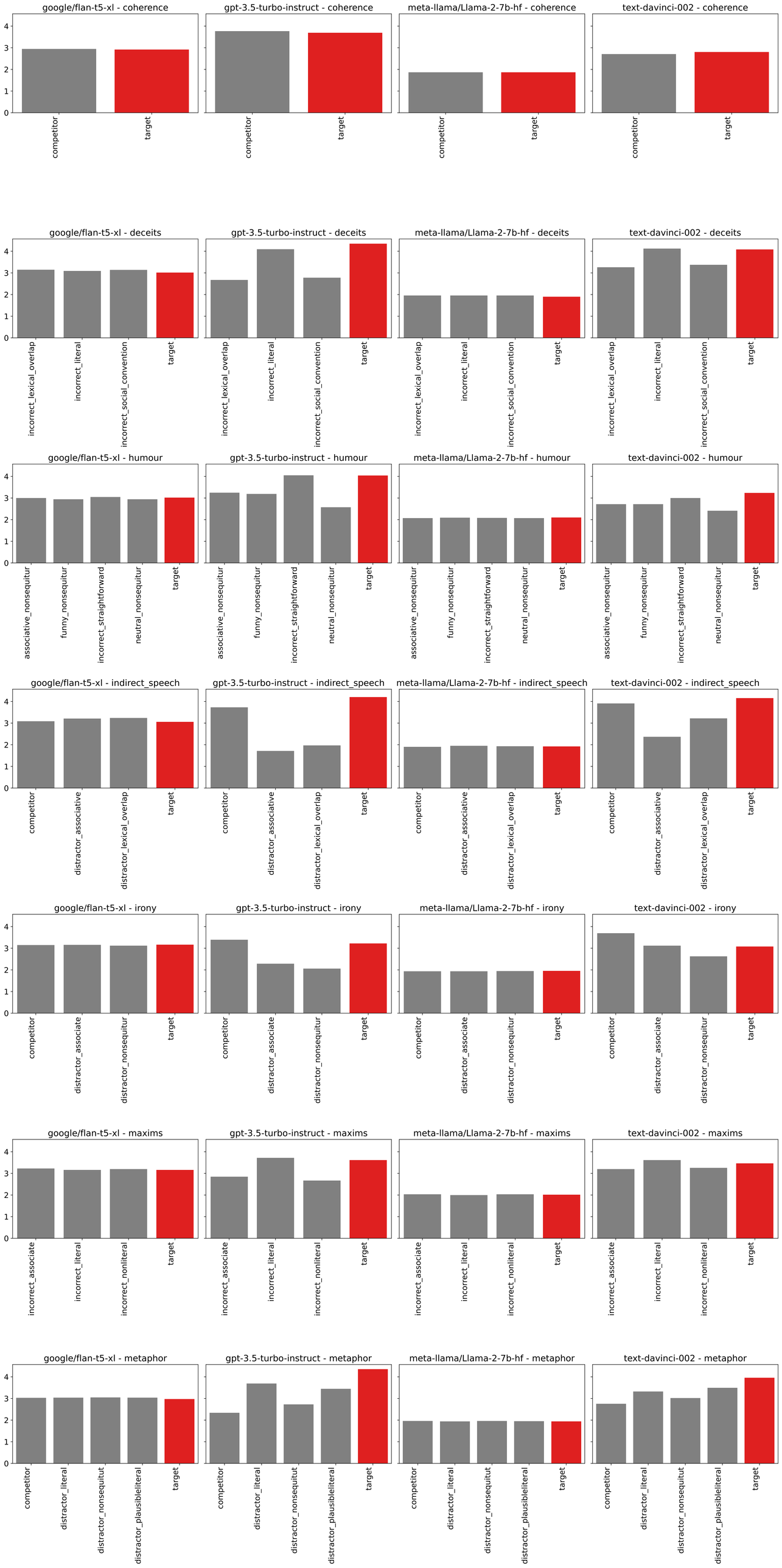
google/flan-t5-xl - coherence
gpt-3.5-turbo-instruct - coherence
meta-llama/Llama-2-7b-hf - coherence
text-davinci-002 - coherence
competitor
target
google/flan-t5-xl - deceits
gpt-3.5-turbo-instruct - deceits
meta-llama/Llama-2-7b-hf - deceits
text-davinci-002 - deceits
incorrect_lexical_overlap
incorrect_literal
incorrect_social_convention
target
google/flan-t5-xl - humour
gpt-3.5-turbo-instruct - humour
meta-llama/Llama-2-7b-hf - humour
text-davinci-002 - humour
associative_nonsequitur
funny_nonsequitur
incorrect_straightforward
neutral_nonsequitur
target
google/flan-t5-xl - indirect_speech
gpt-3.5-turbo-instruct - indirect_speech
meta-llama/Llama-2-7b-hf - indirect_speech
text-davinci-002 - indirect_speech
competitor
distractor_associative
distractor_lexical_overlap
target
google/flan-t5-xl - irony
gpt-3.5-turbo-instruct - irony
meta-llama/Llama-2-7b-hf - irony
text-davinci-002 - irony
competitor
distractor_associate
distractor_nonsequitur
target
google/flan-t5-xl - maxims
gpt-3.5-turbo-instruct - maxims
meta-llama/Llama-2-7b-hf - maxims
text-davinci-002 - maxims
incorrect_associate
incorrect_literal
incorrect_nonliteral
target
google/flan-t5-xl - metaphor
gpt-3.5-turbo-instruct - metaphor
meta-llama/Llama-2-7b-hf - metaphor
text-davinci-002 - metaphor
competitor
distractor_literal
distractor_nonsequitut
distractor_plausibleliteral
target