semantic top to pottom semantic random? semantic bottom to top thematic randoma thematic top to bottom thematic bottom to top deepseex solutions semantic randoma semantic random's thematic randomit semantic randomiz semantic random3 thematic random? thematic randoms thematic random? randoma

Positive value: Prompt A outperforms Prompt B Negative value: Prompt B outperforms Prompt A

Value represents proportion of disagreements * $p \le 0.05$, ** $p \le 0.01$, *** $p \le 0.001$ † Significant after Bonferroni correction

Prompt B

 $p \le 0.01 (**)$

 $p \le 0.05 (*)$

p > 0.05