

# Social Interaction QA

Cross Entropy

$-\log p_{\theta}^{\text{Vocab}}(\text{Correct Choice})$

$-\log \sum_{\text{Incorrect Choice}} p_{\theta}^{\text{Vocab}}(\text{Incorrect Choice})$

Model Family

- Cerebras (Param. and Data Scaling)
- Pythia 12B Param. (Data Scaling)
- Pythia 300B Tokens (Param. Scaling)

$10^{17}$

$10^{18}$

$10^{19}$

$10^{20}$

$10^{21}$

$10^{22}$

Compute (6ND)

$10^{17}$

$10^{18}$

$10^{19}$

$10^{20}$

$10^{21}$

$10^{22}$

Compute (6ND)