

Metric $\log p_{\theta}^{\mathsf{Vocab}}(\mathsf{Correct\ Choice})$ $p_{\theta}^{\mathsf{Vocab}}(\mathsf{Correct\ Choice})$ $p_{\theta}^{\mathsf{Choices}}(\mathsf{Correct\ Choice})$ Brier Score Accuracy Model Family Cerebras (Param. and Data Scaling) INCITE 7B Param. (Data Scaling)