

Benchmark: MMLU Moral Scenarios
Performance Metric: $p_{\theta}^{\text{Choices}}$ (Correct Choice)
Correlation Metric: Pearson

