

Evaluation Loss Comparison: CoT vs No CoT vs No Curriculum
(batch=128, n_layers=1, k_phases=8)

