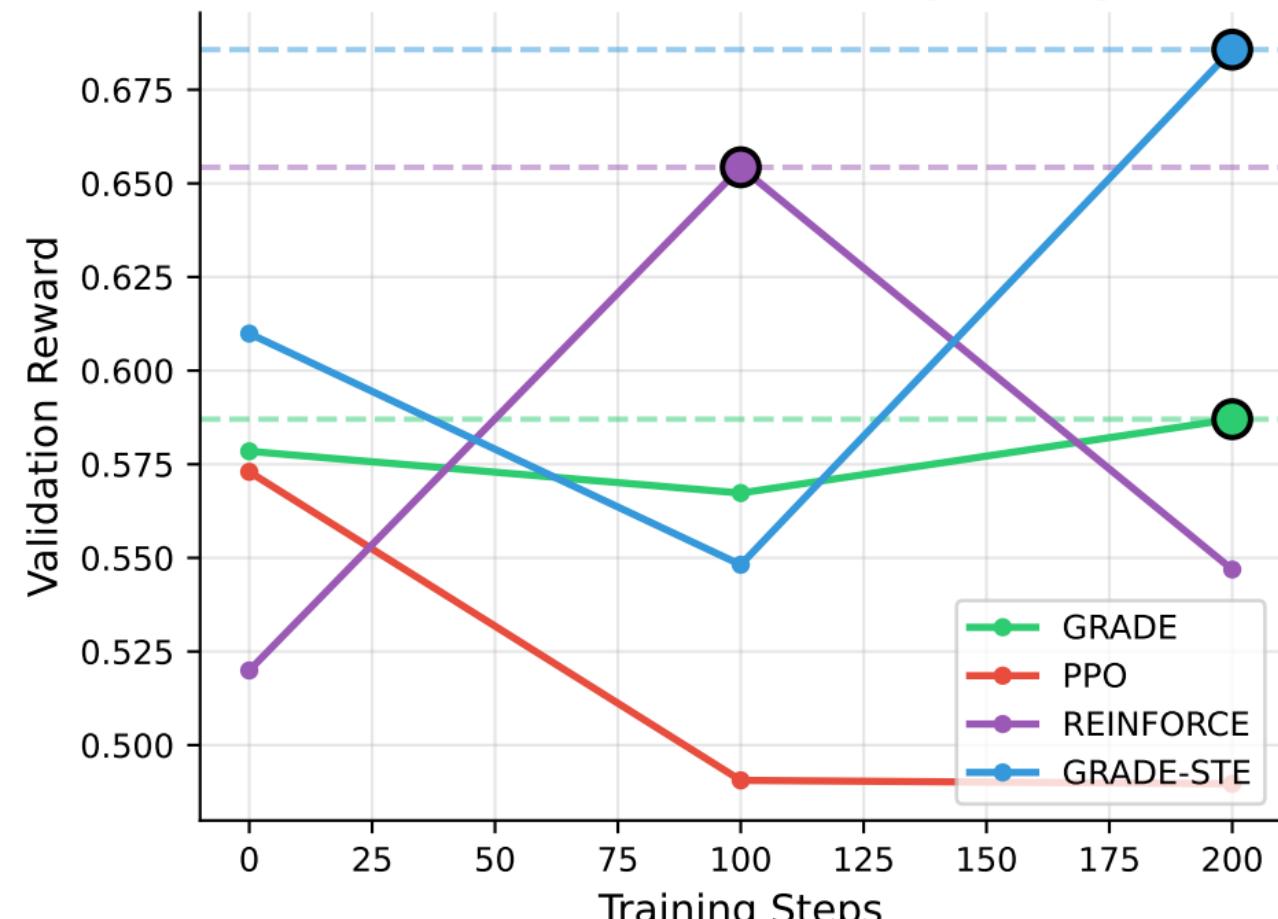


(a) Validation Reward During Training



(b) Generalization Gap (positive = overfitting)

