RL Performance of Learned Rewards Evaluated Under the GT Reward **Ground Truth XIRL** 60 XIRL Mixed Goal Classifier **Cumulative Ground Truth Reward** X-RLHF 50 **XTriplets** XIRL-Buckets 40 **XPrefs** XPrefs (g=0)30 20 10 0

150000

**SAC Training Steps** 

100000

200000

250000

50000