Env Score per Episode Env Score Std Deviation Baseline (SFT) 15 -Env Score Episode