

RL batched experiments q_a^π 's: near zero, batch size=1000

