

RL batched experiments q_a^π 's: between zero and one, batch size=1000

