

RL non-batched experiments - q_a^π 's: between zero and one, α : 0.1

