

RL batched experiments  $q_a^\pi$ 's: between zero and one, batch size=10

