

reward avg, blue = SARSA, dotted = testing

