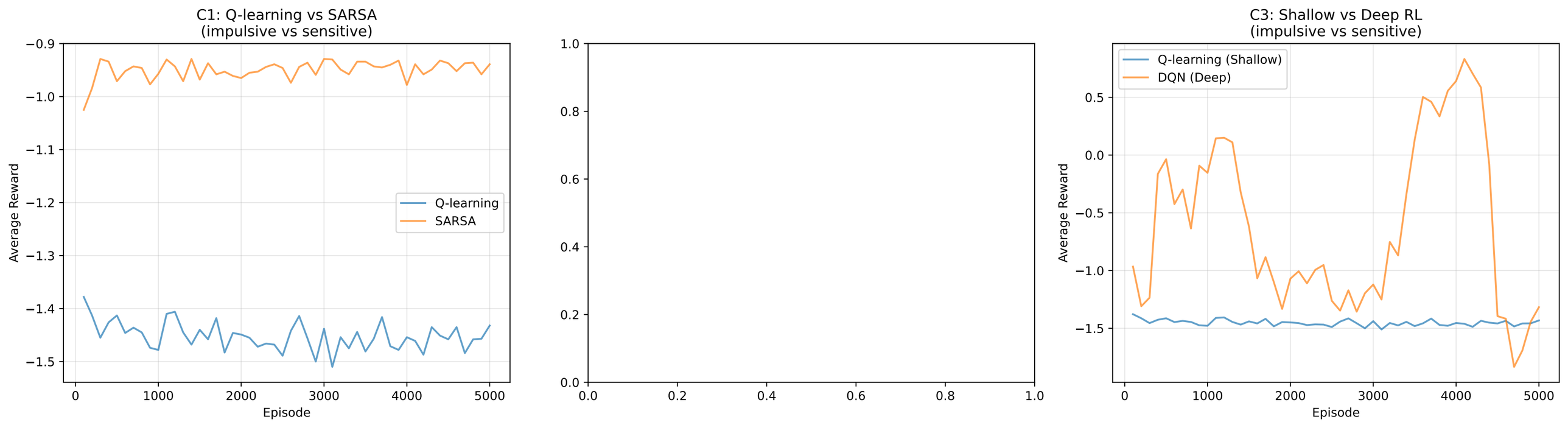
C1: Q-learning vs SARSA
(impulsive vs sensitive)
Average Reward
Episode
Q-learning
SARSA
C3: Shallow vs Deep RL
(impulsive vs sensitive)
Average Reward
Episode
Q-learning (Shallow)
DQN (Deep)