Training Learning Curves Federated Centralized Decentralized Double (Ranked) Grid-3x3 (Ranked) Grid-5x5 (Ranked) -140Mean Episode Reward -17.5-120-20.0-160-140-22.5-1800k100k 200k 0k100k 200k 0k100k 200k Time-Steps Time-Steps Time-Steps Double (Unranked) Grid-3x3 (Unranked) Grid-5x5 (Unranked) Mean Episode Reward -150-120-20 -175 -140-200 -25 0k100k 200k 0k100k 200k 0k100k 200k Time-Steps Time-Steps Time-Steps

(b)

(c)

(a)