Pure Coordination 5 Seed Episode Reward Mean Config Size 30 Small Medium Large 25 Policy Reward Mean 10 5 2 4 5 0 1e6 Total steps