Stag Hunt Evaluation Seed Episode Reward Mean Config Size Small Medium 80 Large Obstacle Policy Reward Mean 20 2 4 5 0 1e6 Total steps