Mean Reward of 20 trajectories ± standard deviation