MountainCar, ppo1, action: mean reward