

Episode Rewards

- 4,ant-6 (explo)
- 4,ant-6 (CAT+explo)
- 4,ant-6 (CAT)
- PPO

0.0

0.5

1.0

1.5

2.0

$\times 10^6$

Training Steps

