BipedalWalker, a2c_agent, observation: mean reward

| Random Steps | Policy Steps 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|
| 9 | -103.76 | -106.64 | -102.71 | -108.19 | -107.91 |
| 7 | -106.11 | -105.1 | -103.66 | -108.13 | -108.03 |
| 5 | -106.27 | -103.78 | -102.49 | -107.68 | -108.11 |
| 3 | -103.99 | -103.91 | -101.13 | -107.35 | -108.12 |
| 1 | -104.45 | -103.68 | -102.93 | -107.81 | -108.3 |