

DQN

Reward

50

40

30

20

10

0

0

1

2

3

4

5

6

7

8

9

10

Training steps

1e6

ER buffer 2 500 000 policies 25 000 policies 250 policies 1 policy

