

Policy loss

0.05
0.04
0.03
0.02
0.01
0.00

0

1

2

3

4

5

Steps

1e6

