

Mean reward

[-]

-29.8

-35.3

1

80

batch

