

Average action value (Q)

0

200

400

600

Epoch

0.8

1.0

1.2