

Cumulative reward per episode (average across three evaluation traces)

