Evaluation: cumulative rewards over time per episode