Training and Evaluation Traces for Fixed Exploration Training trace 1 100 Training trace 2 Training trace 3 50 Evaluation trace 0 cumulative reward -50 -100 -150 -200 · <del>-250</del> --300 · 100 200 300 400 500 training episode