Average Reward vs. Number of Training Iterations

- step_size=1
- step_size=0.1
- step_size=0.01
- step_size=0.001