Average Reward vs. Number of Training Iterations

- reward=constant
- reward=linear
- reward=squared
- reward=skewed