

**Training Stability Comparison: SAC vs DQN vs A2C (100k timesteps)**

