Action Preference: PPO vs Baselines