Trained x-agent vs. random o-agent