For world1, where the agent should ignore smaller reward to prioritize higher later reward, higher step count results in better prioritization of the later higher reward