PPO Loss - MDL - 7 hid. layers (Avg. across tasks)