Walker2d, sac_agent, observation: mean reward

	10	20	30	40	50
9	441.09	686.3	875.32	1347.75	1760.65
7	534.91	932.9	1313.02	1791.82	2298.6
5	612.85	1362.74	1658.44	1989.97	2956.98
3	1072.28	2275.43	2857.2	2815.55	3402.26
1	3125.68	3587.46	3657.09	3752.57	3777.97

Random Steps (y-axis), Policy Steps (x-axis)