Walker2d, tqc_agent, observation: mean reward

Random Steps \ Policy Steps	10	20	30	40	50
9	385.51	1054.48	1386.48	1606.3	1922.31
7	584.85	1487.44	1729.26	1962.86	2306.51
5	1003.03	1832.51	2334.09	2425.57	2734.15
3	1902.55	2456.05	2707.32	2712.48	3016.51
1	2787.1	3153.75	3422.76	3559.73	3725.29