

Table 11: Agent scores and human-normalized metrics on the 26 games of the Atari 100k benchmark. We show average scores over 5 seeds. Bold numbers indicate best performing method for each game.

Game	Random	Human	SimPLe	DreamerPro	TWM	IRIS	DreamerV3	STORM	Δ -IRIS	TWISTER (ours)
Alien	228	7128	617	635	675	420	959	984	391	970
Amidar	6	1720	74	85	122	143	139	205	64	184
Assault	222	742	527	374	683	1524	706	801	1123	721
Asterix	210	8503	1128	594	1116	854	932	1028	2492	1306
Bank Heist	14	753	34	217	467	53	649	641	1148	942
Battle Zone	2360	37188	4031	4200	5068	13074	12250	13540	11825	9920
Boxing	0	12	8	84	78	70	78	80	70	88
Breakout	2	30	16	14	20	84	31	16	302	35
Chopper Command	811	7388	979	1564	1697	1565	420	1888	1183	910
Crazy Climber	10780	35829	62584	53964	71820	59324	97190	66776	57854	81880
Demon Attack	152	1971	208	123	350	2034	303	165	533	289
Freeway	0	30	17	11	24	31	0	34	31	32
Frostbite	65	4335	237	723	1476	259	909	1316	279	305
Gopher	258	2412	597	1250	1675	2236	3730	8240	6445	22234
Hero	1027	30826	2657	4331	7254	7037	11161	11044	7049	8773
James Bond	29	303	100	142	362	463	445	509	309	573
Kangaroo	52	3035	51	2040	1240	838	4098	4208	2269	6016
Krull	1598	2666	2205	6989	6349	6616	7782	8413	5978	8839
Kung Fu Master	258	22736	14862	20650	24555	21760	21420	26182	21534	23442
Ms Pacman	307	6952	1480	1240	1588	999	1327	2673	1067	2206
Pong	-21	15	13	3	19	15	18	11	20	20
Private Eye	25	69571	35	40	87	100	882	7781	103	1608
Qbert	164	13455	1289	1094	3331	746	3405	4522	1444	3197
Road Runner	12	7845	5641	9122	9107	9615	15565	17564	10414	17832
Seaquest	68	42055	683	366	774	661	618	525	827	532
Up N Down	533	11693	3350	4021	15982	3546	7600	7985	4072	7068
# Superhuman	0	N/A	1	4	8	10	9	10	11	12
Normed Mean (%)	0	100	33	79	96	105	112	127	139	162
Normed Median (%)	0	100	13	28	51	29	49	58	53	77