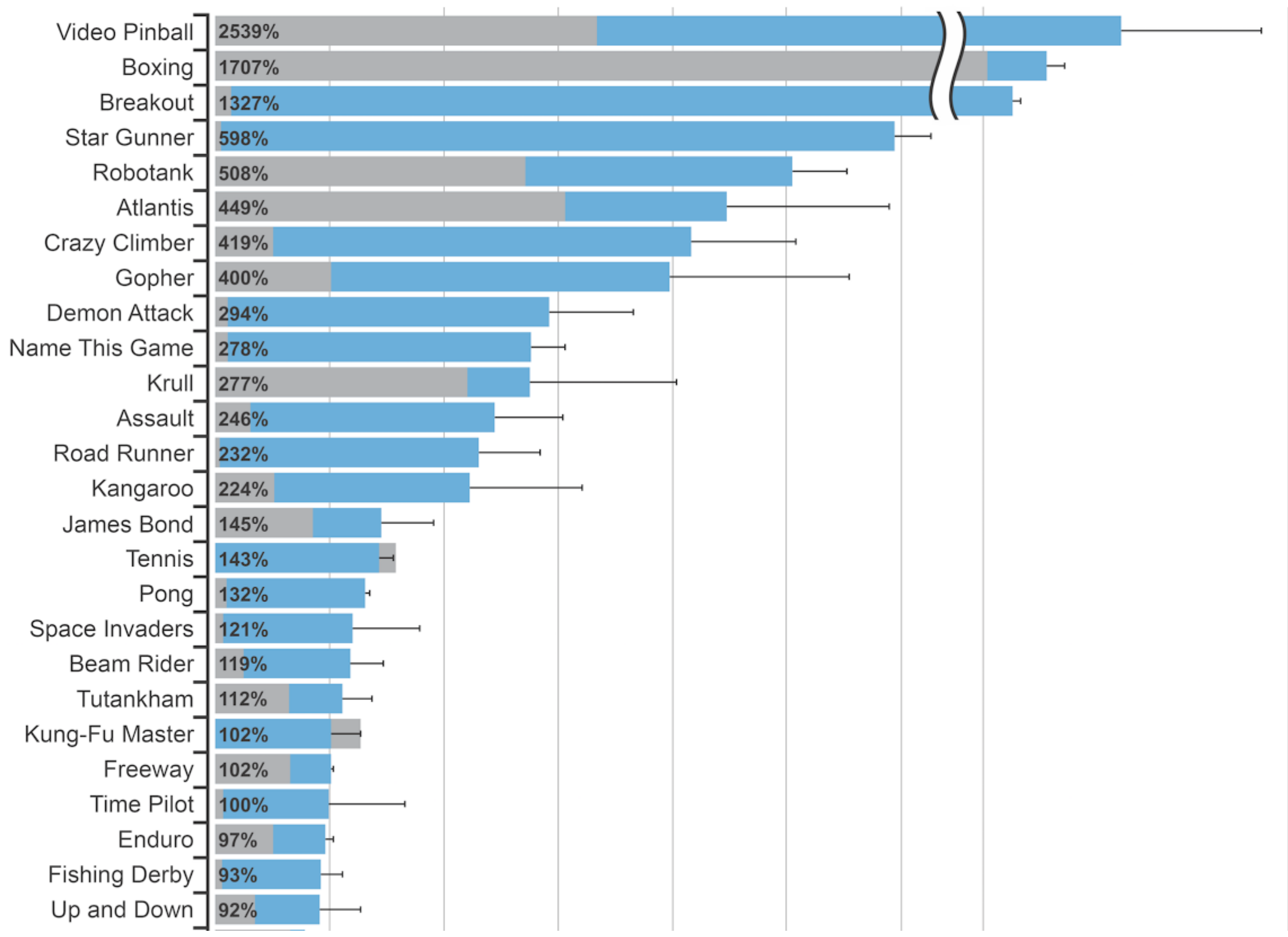
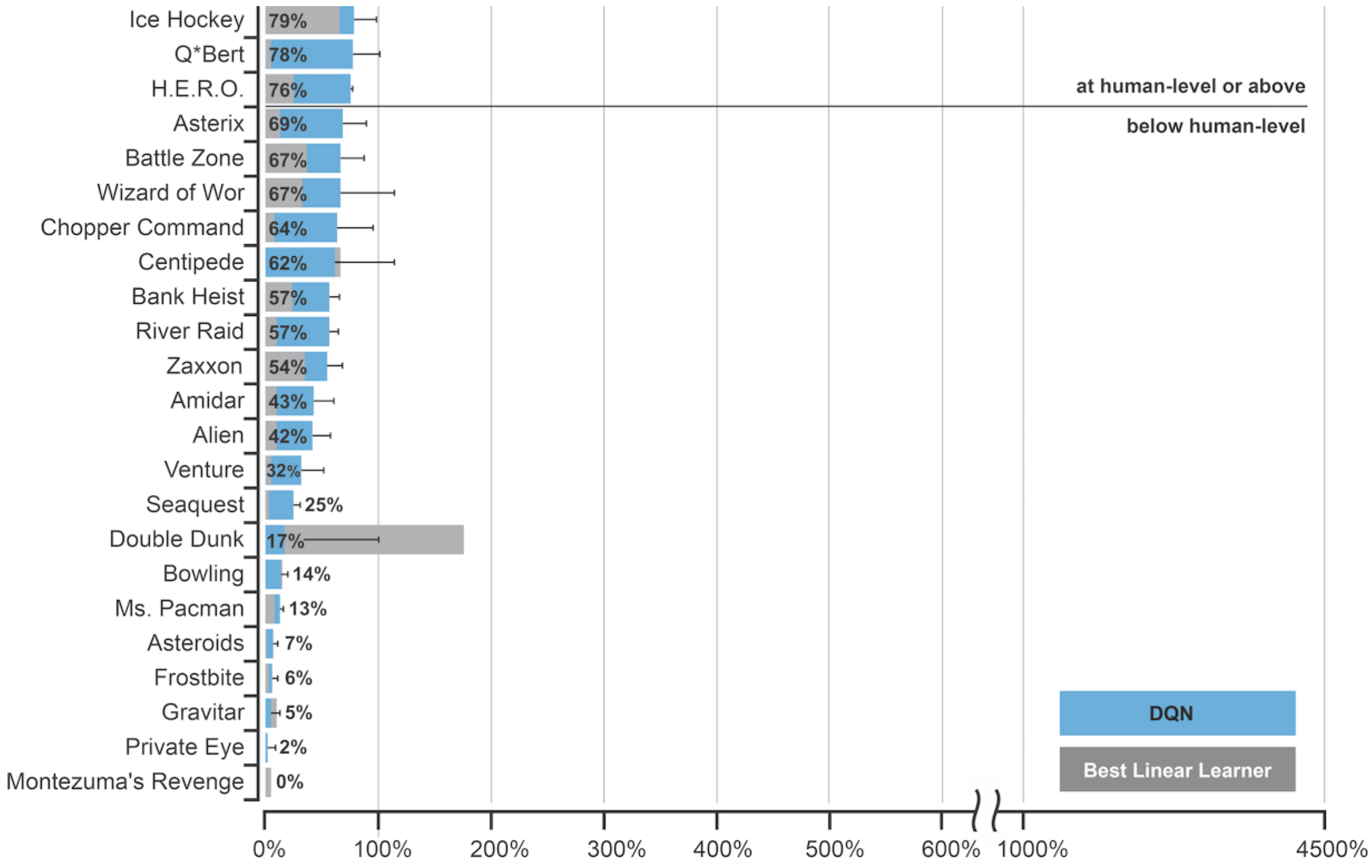
Deep Reinforcement Learning

Humans excel at solving a wide variety of challenging problems, from low-level motor control through to high-level cognitive tasks. Our goal at DeepMind is to create artificial agents that can achieve a similar level of performance and generality. Like a human, our agents learn for themselves to achieve successful strategies that lead to the greatest long-term rewards. This paradigm of learning by trial-and-error, solely from rewards or punishments, is known as reinforcement learning (RL). Also like a human, our agents construct and learn their own knowledge directly from raw inputs, such as vision, without any hand-engineered features or domain heuristics. This is achieved by deep learning of neural networks. At DeepMind we have pioneered the combination of these approaches - deep reinforcement learning - to create the first artificial agents to achieve human-level performance across many challenging domains.

Our agents must continually make value judgements so as to select good actions over bad. This knowledge is represented by a Q-network that estimates the total reward that an agent can expect to receive after taking a particular action. Two years ago we introduced the first widely successful algorithm for deep reinforcement learning. The key idea was to use deep neural networks to represent the Q-network, and to train this Q-network to predict total reward. Previous attempts to combine RL with neural networks had largely failed due to unstable learning. To address these instabilities, our Deep Q-Networks (DQN) algorithm stores all of the agent's experiences and then randomly samples and replays these experiences to provide diverse and decorrelated training data. We applied DQN to learn to play games on the Atari 2600 console. At each time-step the agent observes the raw pixels on the screen, a reward signal corresponding to the game score, and selects a joystick direction. In our Nature paper we trained separate DQN agents for 50 different Atari games, without any prior knowledge of the game rules.




Ice Hockey
Q*Bert
H.E.R.O.
Asterix
Battle Zone
Wizard of Wor
Chopper Command
Centipede
Bank Heist
River Raid
Zaxxon
Amidar
Alien
Venture
Seaquest
Double Dunk
Bowling
Ms. Pacman
Asteroids
Frostbite
Gravitar
Private Eye
Montezuma's Revenge
79%
78%
76%
69%
67%
67%
64%
62%
57%
57%
54%
43%
42%
32%
25%
17%
14%
13%
7%
6%
5%
2%
0%
at human-level or above
below human-level
DQN
Best Linear Learner
0%
100%
200%
300%
400%
500%
600%
1000%
4500%

Research

Applied

News & Blog

About Us

Careers

Press
Terms and Conditions
Privacy Policy – Updated
Alphabet Inc