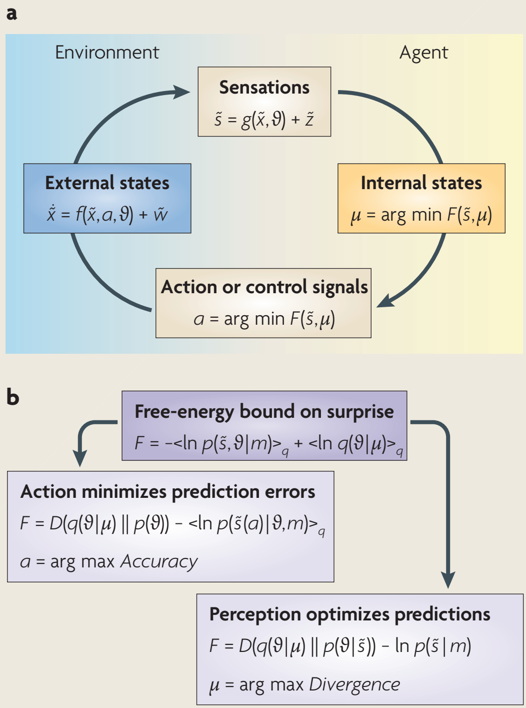
Box 1 | **The free-energy principle**

Part **a** of the figure shows the dependencies among the quantities that define free energy. These include the internal states of the brain $\mu(t)$ and quantities describing its exchange with the environment: sensory signals (and their motion) $\tilde{s}(t) = [s, s', s'' \ldots]^T$ plus action $a(t)$. The environment is described by equations of motion, which specify the trajectory of its hidden states. The causes $\vartheta \supset \{\tilde{x}, \theta, \gamma\}$ of sensory input comprise hidden states $\tilde{x}(t)$, parameters θ and precisions γ controlling the amplitude of the random fluctuations $\tilde{z}(t)$ and $\tilde{w}(t)$. Internal brain states and action minimize free energy $F(\tilde{s}, \mu)$, which is a function of sensory input and a probabilistic representation $q(\vartheta|\mu)$ of its causes. This representation is called the recognition density and is encoded by internal states μ.

The free energy depends on two probability densities: the recognition density $q(\vartheta|\mu)$ and one that generates sensory samples and their causes, $p(\tilde{s}, \vartheta|m)$. The latter represents a probabilistic generative model (denoted by m), the form of which is entailed by the agent or brain. Part **b** of the figure provides alternative expressions for the free energy to show what its minimization entails: action can reduce free energy only by increasing accuracy (that is, selectively sampling data that are predicted). Conversely, optimizing brain states makes the representation an approximate conditional density on the causes of sensory input. This enables action to avoid surprising sensory encounters. A more formal description is provided below.



Optimizing the sufficient statistics (representations)

Optimizing the recognition density makes it a posterior or conditional density on the causes of sensory data: this can be seen by expressing the free energy as surprise $-\ln p(\tilde{s}, | m)$ plus a Kullback-Leibler divergence between the recognition and conditional densities (encoded by the 'internal states' in the figure). Because this difference is always positive, minimizing free energy makes the recognition density an approximate posterior probability. This means the agent implicitly infers or represents the causes of its sensory samples in a Bayes-optimal fashion. At the same time, the free energy becomes a tight bound on surprise, which is minimized through action.

Optimizing action

Acting on the environment by minimizing free energy enforces a sampling of sensory data that is consistent with the current representation. This can be seen with a second rearrangement of the free energy as a mixture of accuracy and complexity. Crucially, action can only affect accuracy (encoded by the 'external states' in the figure). This means that the brain will reconfigure its sensory epithelia to sample inputs that are predicted by the recognition density — in other words, to minimize prediction error.