

How people distort explicit information about source reliability

Ota K, Ciston A, Haggard P, Gajdos Preuss T*, Charles L* (2026). How does explicit reliability guide choices? Effect of trustworthiness on choice and metacognition. *PLoS Comput Biol* 22(10): e1014818. <https://doi.org/10.1371/journal.pcbi.1014818>

* Equal contribution.

We studied how people use explicit information about the reliability of information sources. In five online experiments (140 participants), each trial showed six predictions, each from a source labelled with its probability of being correct (e.g. 50%, 55%, 65%); participants then chose between two options. We modelled choices with a hierarchical Bayesian model in which stated reliability is distorted on a log-odds scale and each prediction is weighted by its position in the sequence. Participants did not take reliability at face value: sources labelled as uninformative (50%) still biased choices, as if they were moderately reliable, and recent predictions weighed more than early ones. Participants were nevertheless partly aware of these influences: their introspective reports tracked how much each source actually influenced their choice.

Use of the DIPSC cluster. The cluster was used for the model recovery analysis. We simulated 640 synthetic datasets (24 participants, 300 trials each) from four candidate models and fitted all four models to each dataset, i.e. 2,560 hierarchical fits with Hamiltonian Monte Carlo in CmdStan (4 chains each). Leave-one-out cross-validation selected the true generating model in 98.4% of cases, which shows that the models can be reliably distinguished.

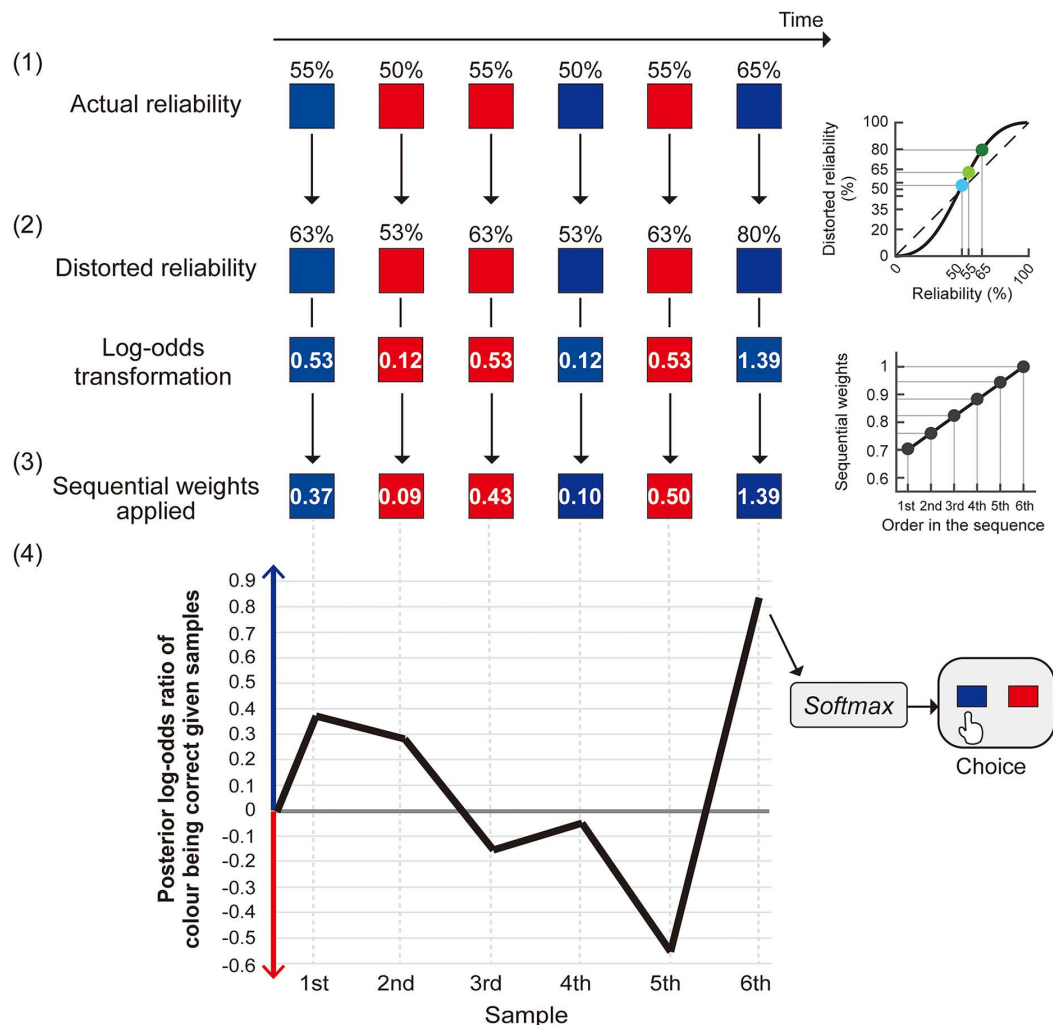


Figure. Computational model. (1) Each trial shows six sources, each labelled with its probability of being correct. (2) Stated reliability is distorted, then converted to log-odds. (3) Each source is weighted by its position in the sequence. (4) Weighted log-odds are summed, and a softmax rule maps the sum to a choice. Reproduced from Ota et al. (2026), Fig 3 (CC BY 4.0).