Heuristics as Bayesian inference under extreme priors

Paula Parpart[a,1], **Matt Jones**[b], **and Bradley C. Love**[a,c]

[a]Department of Experimental Psychology, University College London, London WC1H 0AP, United Kingdom; [b]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, C0, 80309; [c]The Alan Turing Institute, London NW1 2DB, United Kingdom

Simple heuristics are often regarded as tractable decision strategies because they ignore a great deal of information in the input data. One puzzle is why heuristics can outperform *full-information* models, such as linear regression, which make full use of the available information. We resolve this apparent paradox by demonstrating that heuristics are limiting cases of probabilistic inference. Inspired by regularization approaches from machine learning, we construct a Bayesian model that regulates sensitivity to cue-weight magnitudes. Parametric variation of the prior's strength generates a continuum of models, with linear regression at one extreme and the tallying heuristic, which equally weighs all cues, at the other extreme. A general difference between regression and heuristic approaches is that regression is fully sensitive to covariance among predictors, whereas heuristics completely ignore it. Based on this observation, we formalize a second Bayesian model with a prior on cue covariance and show that this model formally links both the tallying and the Take-The-Best heuristics with linear regression. These results, which cast heuristics as a limiting case with an infinitely strong prior, refute the strong version of "less-is-more" claims whereby entirely discarding some information sources (as heuristics do) is optimal. Indeed, intermediate models perform better across all our simulations, suggesting that down-weighting information is preferable to entirely ignoring it. These results have implications for work in psychology, machine learning and economics.

heuristics | Bayesian inference | decision making | ridge regression