# Information Theory Meets Expected Utility: The Entropic Roots of Probability Weighting Functions

Mikaela Akrenius, Indiana University Bloomington

FULBRIGHT FINLAND FOUNDATION

## Introduction

- The most widely applied **non-expected utility theories**[1,2,3] combine the classical core of expected utility theory[4] with a **concave utility function** and **non-linear probability weighting** to account for findings unexplained by EUT[5,6,7]

$$EV(X) = \sum_{i=1}^{n} \boldsymbol{\pi(p_i)} * \boldsymbol{v(x_i)}$$

- Despite of decades of research, the **psychological processes** underlying the estimated shape of these functions remain unclear
- Earlier work has suggested that **concave utility** could reflect **decision by sampling**[8], **efficient coding**[9, 10, 11], **heuristic processing**[12] or **attentional processes**[13, 14], whereas **probability weighting** has been suggested to result from **sensitivity to extreme probabilities**[1, 15], a **log-odds representation** of frequency and probability-related information[16], and/or **bounded rationality**[17, 18]
- Based on existing work in perceptual and cognitive psychology[19, 20, 21, 22] and neuroscience[23, 24], the present approach builds on the assumption that **a reduction in uncertainty carries utility**

## Valence-Weighted Distance (VWD) [25]

- The **perception of a probability** $p$ is influenced by the **amount of uncertainty reduction it carries**
  - relative to the **uniform** distribution (maximum entropy)
  - depending on **other probabilities** in the distribution (actual entropy)
- Formally:

$$VWD(p) = \frac{p^{\left(1-\frac{H_{dist}}{H_{max}}\right)}}{\sum_{i=1}^{n} p_i^{\left(1-\frac{H_{dist}}{H_{max}}\right)}}$$

where $Hdist = -\sum p_i \log_2(p_i)$[26], $Hmax = \log_2(n)$, and $n$ denotes the number of probabilities in the distribution that $p$ is embedded in

- This yields an $n$-dimensional function that has its fixed point at $1/n$ and curvature that depends on the entropy of the distribution

**Note:** Existing probability weighting functions cannot reflect changes in $n$ or distributional shape without changes in fitted parameters

**Figure 1.** Existing probability weighting functions[3, 27, 28, 29, 30]

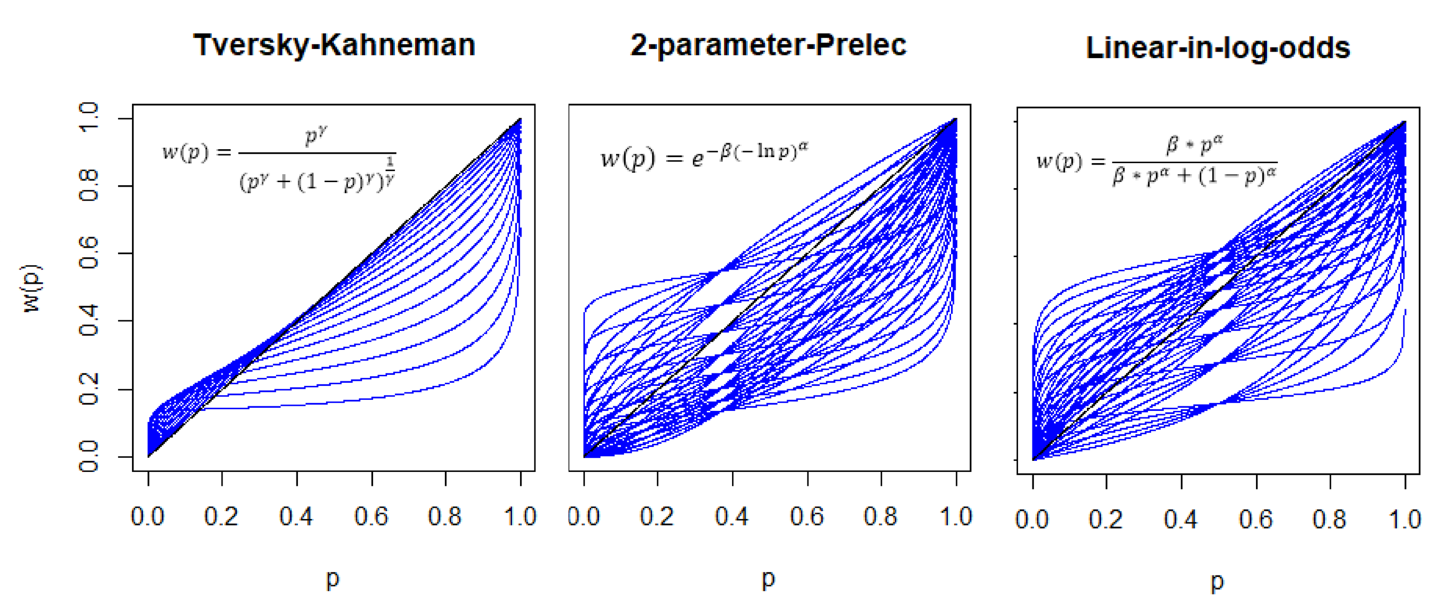



**Figure 2.** $VWD(p)$ when $n$ is varied from to 2 to 7 while retaining maximum entropy (left) and when entropy is varied while $n = 3$ (right)

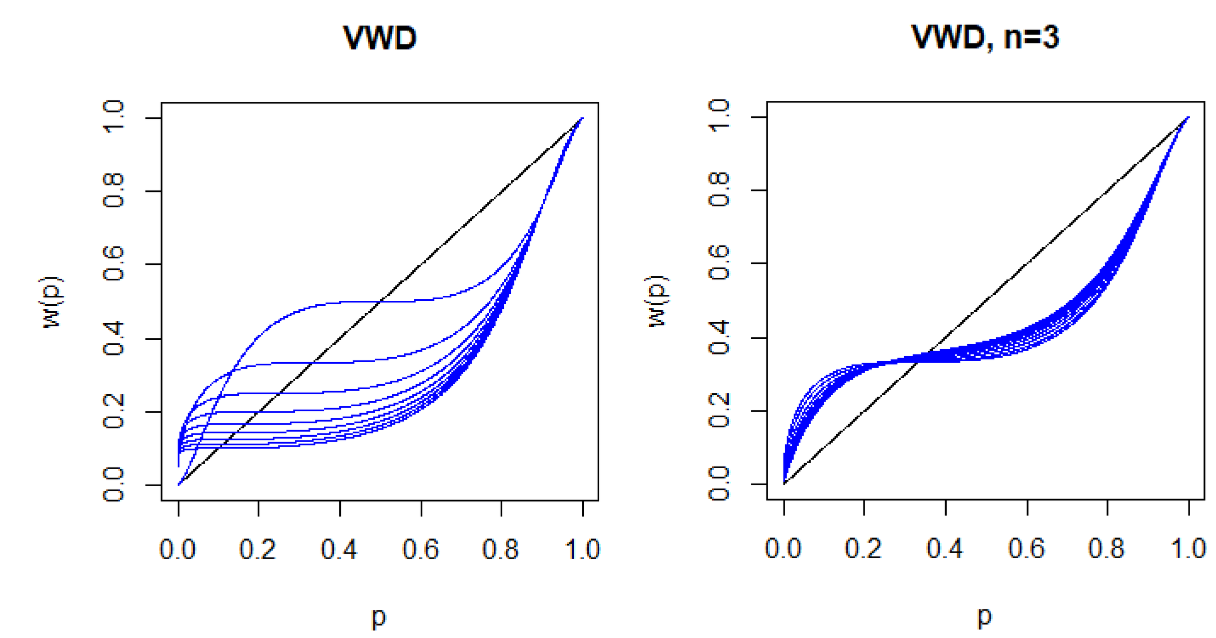
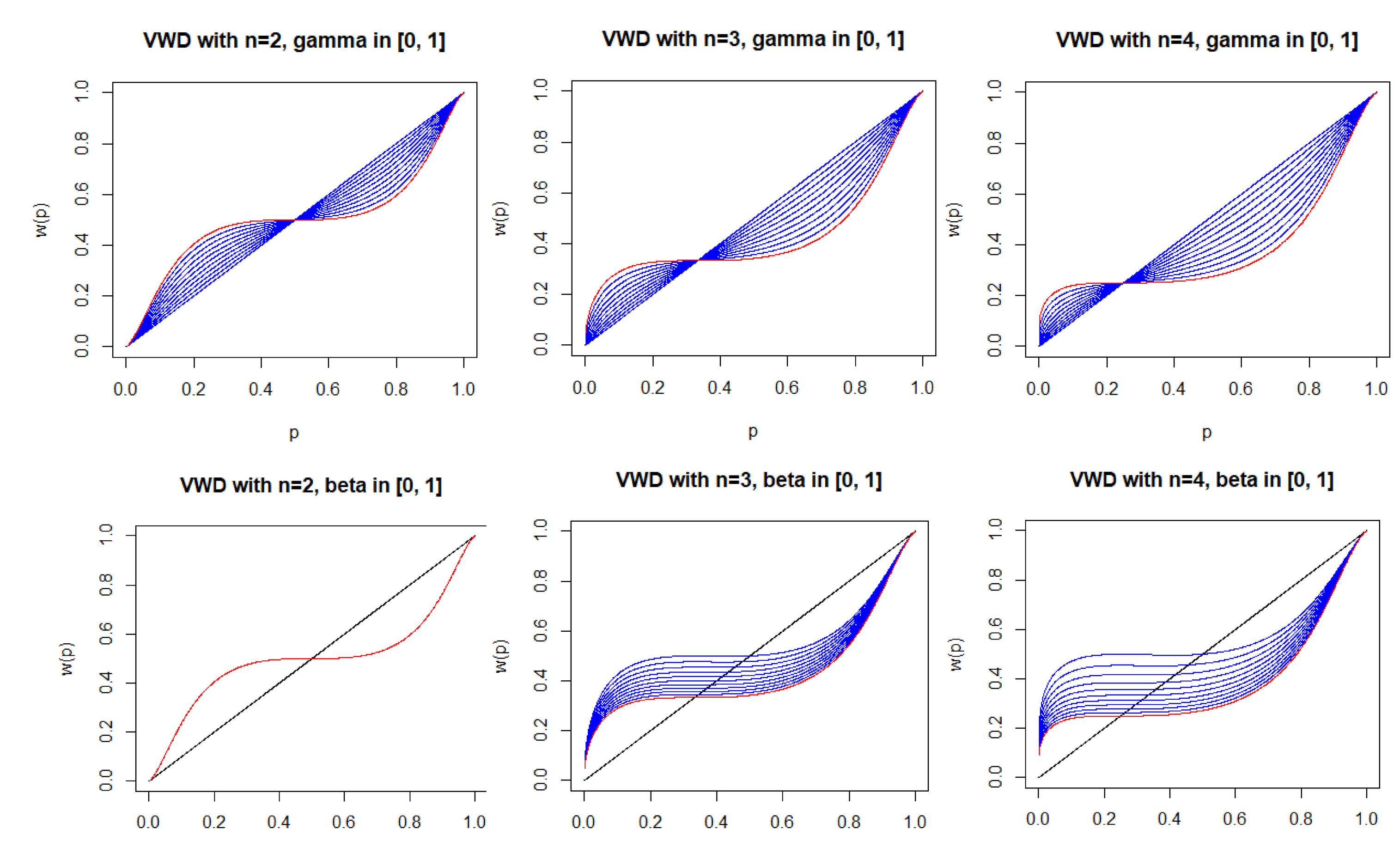
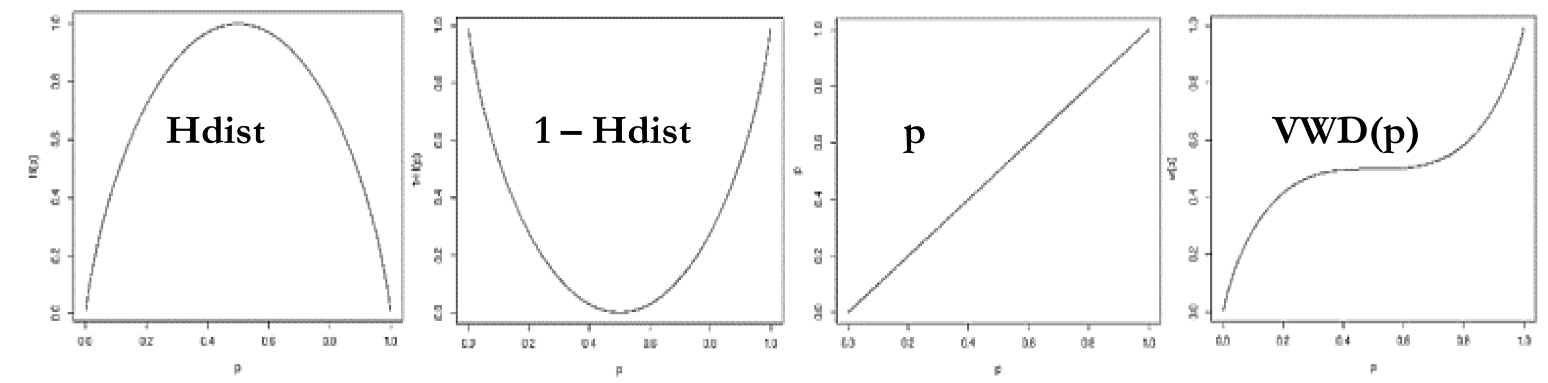



**Figure 3.** Parameterized $VWD(p)$ when γ is varied from 0 to 1 while fixing β to 1 (upper row) and when β is varied from to 0 to 1 while fixing γ to 1 (lower row) for $n = 2$ (left), $n = 3$ (middle), and $n = 4$ (right).




## Illustration with $n$ = 2

Utility of uncertainty reduction merged with utility of increased likelihood of gain




$$VWD(p) = \frac{p^{(1-Hdist)}}{p^{(1-Hdist)} + (1-p)^{(1-Hdist)}}$$

For $n = 2$, $Hmax = 1$

Compare to linear-in-log-odds: $w(p) = \frac{\beta * p^{\alpha}}{\beta * p^{\alpha} + (1-p)^{\alpha}}$

## Parameterized VWD

$$VWD(p_1) = \frac{p_1^{\left(1-\gamma*\frac{H_{dist}}{H_{max}}\right)}}{p_1^{\left(1-\gamma*\frac{H_{dist}}{H_{max}}\right)} + \beta * \sum_{i=2}^{n} p_i^{\left(1-\gamma*\frac{H_{dist}}{H_{max}}\right)} + (1-\beta)*(1-p_1)^{\left(1-\gamma*\frac{H_{dist}}{H_{max}}\right)}}$$

where β, γ ϵ [0, 1] reflect attention given to $n$ and entropy

## Contributions

**1. Probability weighting**

- Explains probability weighting with **a simple principle**
- Takes **context of** $\boldsymbol{p}$ ($n$, shape of distribution) into account
- Makes **novel, empirically testable predictions**
  - Location of fixed point determined by $n$
  - Curvature determined by shape of distribution

**2. Expected utility theory and information theory**

- Captures the psychological impact of **outcome-probability associations**[31] cf. [32, 33]

3. **Log-odds representation of frequency and probability**

- Provides a parsimonious explanation for the $1/n$ puzzle[16]

$$log\frac{VWD(p_i)}{1-VWD(p_i)} = log\frac{p_i^{\left(1-\frac{H_{dist}}{H_{max}}\right)}}{\sum_{j=1}^{n} p_j^{\left(1-\frac{H_{dist}}{H_{max}}\right)}} \quad \text{where } i \neq j$$

**Contact** Mikaela Akrenius, makreniu@indiana.edu, Cognitive Science Program, Indiana University, 1900 E 10th St, Bloomington, IN 47406–7512

SJDM 2020 Zoom

**References** 1 Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291. 2 Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization*, 3(4), 323–343. 3 Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. 4 von Neumann, J., & Morgenstern, O. (1944/1947). *Theory of games and economic behavior.* Princeton: Princeton University Press. 5 Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'École Américaine. *Econometrica*, 21, 503–546. 6 Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 75, 643– 669. 7 Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2, 61–104. 8 Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice. *Quarterly Journal of Experimental Psychology*, 62:6, 1041-1062. 9 Arkes, H. R. (1991). Costs and benefits of judgment errors: Implications for debiasing. *Psychological Bulletin*, 110(3), 486–498. 10 Summerfield, C. & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences,* 19(1), 27–34. 11 Bhui, R., & Gershman, S. J. (2018). Decision by sampling implements efficient coding of psychoeconomic functions. *Psychological Review*, 125:6, 985–1001. 12 Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review,* 113(2), 409-432. 13 Johnson, J. G., & Busemeyer, J. R. (2016). A computational model of the attention process in risky choice. *Decision, 3*(4), 254–280. 14 Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, 147, 147–169. 15 Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58, 697–720. 16 Zhang, H. & Maloney, L. (2012). Ubiquitous log odds: a common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience*, 6, 1–14. 17 Zhang, H., Ren, X., & Maloney, L. T. (2020). The bounded rationality of probability distortion. *Proceedings of the National Academy of Sciences*, 117(36), 22024–22034. 18 Juechems, K., Balaguer, J., Spitzer, B., & Summerfield, C. (2020). Optimal utility and probability functions for agents with finite computational precision. *PsyArXiv preprint*, February 2020. 19 Attneave, F. (1953). Psychological probability as a function of experienced frequency. *Journal of Experimental Psychology*, 46, 81–86. 20 Garner, W. R., & Clement, D. E. (1963). Goodness of Pattern and Pattern Uncertainty. *Journal of Verbal Learning and Verbal Behavior*, 2, 446–452. 21 Feldman, J. (2016). The simplicity principle in perception and cognition. *WIREs Cognitive Science*, 7, 330–340. 22 Crupi, V., Nelson, J. D., Meder, B., Cevolani, G., & Tentori, K. (2018). Generalized information theory meets human cognition: Introducing a unified framework to model uncertainty and information search. *Cognitive Science*, 42, 1410–1456. 23 Friston, K. (2010). The free energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11, 127–138. 24 Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36:3, 1–73. 25 Akrenius, M. (2020). Information theory meets expected utility: The entropic roots of probability weighting functions. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.) *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 3233–3239). Toronto, ON: Cognitive Science Society. 26 Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423 and 623–656. 27 Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527. 28 Tversky, A. & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102, 269–283. 29 Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, 94, 236–254. 30 Gonzalez, R. & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38(1), 129–166. 31 Luce, R. D. (2003). Whatever happened to information theory in psychology? *Review of General Psychology*, 7, 183–188. 32 Luce, R. D., Marley, A. J., & Ng, C. T. (2009). Entropy-related measures of the utility of gambling. In S. J. Brams et al. (Eds.), *The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn*, Studies in Choice and Welfare. 33 Yang, J. & Qui, W. (2014). Normalized expected utility-entropy measure of risk. *Entropy*, 16, 3590–3604.