

Optimal experimental design models of naïve human information acquisition
[For NIPS 2005 workshop, Dec 10, on value of information analysis/applications]

Jonathan D. Nelson

jnelson@salk.edu

<http://www.jonathandnelson.com/>

Gary Cottrell

gary@cs.ucsd.edu

<http://www.cs.ucsd.edu/users/gary/>

Flavia Filimon

ffilimon@cogsci.ucsd.edu

<http://www.cogsci.ucsd.edu/~ffilimon/>

Terry Sejnowski

terry@salk.edu

<http://www.sn1.salk.edu/~terry>

A number of statisticians have discussed general purpose utility functions for identifying useful experiments to conduct, as special cases of value of information (VOI) analysis. These utility functions are proposed for use in experiment selection problems where the goal is to obtain information about the true state of the world, in the absence of knowledge of the precise utility of each correct or incorrect decision. Theoretically, this probabilistic, decision-theoretic (Savage, 1954) approach to experiment selection contrasts with the logical falsificationist philosophy of science popularized by Karl Popper (1959). Examples of optimal experimental design (OED) utility functions include Bayesian diagnosticity (I. J. Good, 1950), mutual information (information gain; Shannon, 1948; Lindley, 1956) or Kullback-Liebler distance (Kullback & Liebler, 1951), impact (Wells & Lindsay, 1980; Klayman & Ha, 1987; Nickerson, 1996), and error minimization (probability gain; Baron, 1985).

Several investigators have studied whether OED theories of information acquisition may provide accurate descriptive models of human intuition and behavior. Do people's ideas about what medical tests are most useful correspond to the theoretical goal of improving the probability of correct diagnosis (Baron, Beattie and Hershey, 1988)? Do people's ideas about what features are most useful in classification tasks correspond to Bayesian diagnosticity (Skov & Sherman, 1986; Slowiaczek et al., 1992)? Do people's choices of queries on the 2-4-6 number concept task (Wason, 1960, 1966) maximize information gain (entropy reduction; Ginzburg & Sejnowski, 1996)? Several authors have also considered whether human eye movement can be modeled as an OED task (Legge, Klitz & Tjan, 1997; Renninger, Coughlan, Verghese, & Malik, 2005; Najemnik & Geisler, 2005; Nelson & Cottrell, in press). These models treat eye movements as experiments, where the visual information obtained by looking in a particular location is the experiment's outcome.

Within the OED framework for understanding and modeling human evidence acquisition, particular researchers' choices of utility functions have been rather arbitrary. Many articles have compared human ratings or choices with only a single utility function. It is therefore unclear whether a particular utility function approximates human intuitions in a wide range of tasks, and if so, which utility. We report progress along these lines (Nelson, 2005). One project was to re-implement the probability models in several articles, to ascertain whether the various utilities make contrary predictions. In most cases, all OED utility functions agree on which possible experiment is most useful, which experiment is next-most-useful, etcetera. Another project was to optimize the behavior of various utility functions, so as to maximize pairwise disagreement between them. Results of this optimization were used in an experiment to better differentiate which utility functions most closely approximate human intuitions.

Together, results to date suggest remarkable concordance between OED utility functions and human intuitions, on a variety of experimental tasks. In current work we seek to compare perceptual (eye movement) and cognitive (question or mouse click) means of information acquisition on a categorization task. Theoretical issues for future work include consideration of sequential scenarios, and of cases with interesting class-conditional feature dependencies, such as symmetries in natural objects. Better elucidation of naïve human intuitions about the usefulness of information should also facilitate identification of particular cases where intuitive strategies may fail.

References are given below. Nelson (2005) gives additional references.

- Baron, J. (1985). *Rationality and Intelligence*. Cambridge: Cambridge University Press.
- Baron, J., Beattie, J., & Hershey, J. C. (1998). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 42, 88-110.
- Ginzburg, I., & Sejnowski, T. J. (1996). Dynamics of rule induction by making queries: transition between strategies. *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, 121-125.
- Good, I. J. (1950). *Probability and the Weighing of Evidence*. New York: Charles Griffin.
- Klayman, J. & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information. *Psychological Review*, 94, 211-228.
- Kullback, S., & Liebler, R. A. (1951). Information and sufficiency. *Annals of Mathematical Statistics*, 22, 79-86.
- Legge, G. E., Klitz, T. S., and Tjan, B. S. (1997). Mr. Chips: an ideal-observer model of reading. *Psychological Review*, 104, 524-553.
- Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, 27, 986-1005.
- Najemnik, J. & Geisler, W.S. (2005) Optimal eye movement strategies in visual search. *Nature*, 434, 387-391.
- Nelson, JD (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact and information gain. *Psychological Review*, 112(4), xxx-xxx (appearing October, 2005).
- Nelson, JD; Cottrell, GW (in press). A probabilistic model of eye movements in concept formation. *Neurocomputing*.
- Nickerson, R. S. (1996). Hempel's paradox and Wason's selection task: Logical and psychological puzzles of confirmation. *Thinking and Reasoning*, 2, 1-32.
- Popper, K. R. (1959). *The Logic of Scientific Discovery*. London: Hutchinson & Co.
- Renninger, L.W., Coughlan, J.M., Verghese, P. & Malik, J. (2004). An information maximization model of eye movements. *NIPS 2004*.
- Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.
- Skov, R. B. & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, 22(2), 93-121.
- Slowiaczek, L. M., Klayman, J, Sherman, S. J. & Skov, R. B. (1992). Information selection and use in hypothesis testing: What is a good question, and what is a good answer? *Memory and Cognition*, 20(4), 392-405.
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New Horizons in Psychology* (pp. 135-151). Harmondsworth, England: Penguin.
- Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, 88, 776-784.