Naïve optimality: Subjects' heuristics can be
better motivated than experimenters' optimal models

Jonathan D Nelson

review of Oaksford & Chater / Bayesian Rationality

word counts:
abstract 45,
main text 1374,
references 564,
total 1983

address correspondence to:
Jonathan D Nelson

Computer Science and Engineering Dept., University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093-0404

Computational Neurobiology Laboratory, Salk Institute for Biological Studies
10010 N. Torrey Pines Rd., La Jolla, CA 92037

858-534-9028
jnelson@salk.edu
http://jonathandnelson.com/

Abstract

Is human cognition best described by optimal models, or by adaptive but suboptimal heuristic strategies? It is frequently hard to identify which theoretical model is normatively best justified. In the context of information search, naïve subjects' euristic strategies are better motivated than some 'optimal' models.

Oaksford and Chater (2007, henceforth "O&C") nicely synthesize the growing body of of research on Bayesian rationality. This approach offers a great deal of promise in explaining in a principled way not only human higher cognitive processes, but much of perception and animal cognition as well. One interesting question within the rational Bayesian approach is whether human cognition is best described by optimal models, or by adaptive but suboptimal heuristic strategies. O&C explain that in the context of reasoning about syllogisms, heuristic strategies approximate human behavior better than the full optimal Bayesian models. This comment seeks to illustrate, in the context of information search, that it is frequently possible to be mistaken about what model is best motivated; and further, that in practice, naïve subjects' heuristic strategies can be better motivated than some 'optimal' models!

Consider the task of deciding which of two medical tests to order, assuming cost constraints only allow one test, to best diagnose a patient's disease. We assume here that the patient either has Disease 1 or Disease 2, with equal (50 %) prior probability, and that Test 1 and Test 2 are each either positive or negative. How should a diagnostician decide which test to order?

A great deal of cognitive psychological and statistical thinking (since the Bayesian statistician I. J. Good's, 1950, 1975, work) claims that the optimal strategy is to conduct the test with highest *expected Bayesian diagnosticity* (also called *expected weight of evidence*) or *expected log diagnosticity*. [We use $T1{=}p$ to note that Test 1 is positive, $T1{=}n$ to note that Test 1 is negative, $D{=}d1$ to note that the patient has disease 1, etc.] A test's expected Bayesian diagnosticity (its expected utility, as measured with Bayesian diagnosticity) is:

$$\begin{aligned}
&eu_{BD}(T1)= \\
&P(T1=p) \;*\; max\left(\frac{P(T1=p \mid D=d1)}{P(T1=p \mid D=d2)},\; \frac{P(T1=p \mid D=d2)}{P(T1=p \mid D=d1)}\right) \\
&+ \\
&P(T1=n) \;*\; max\left(\frac{P(T1=n \mid D=d1)}{P(T1=n \mid D=d2)},\; \frac{P(T1=n \mid D=d2)}{P(T1=n \mid D=d1)}\right)
\end{aligned}$$

and its expected log diagnosticity is:

$$eu_{LogD}(T1) = P(T1=p) * \log\ max\left(\frac{P(T1=p \mid D=d1)}{P(T1=p \mid D=d2)}, \frac{P(T1=p \mid D=d2)}{P(T1=p \mid D=d1)}\right) + P(T1=n) * \log\ max\left(\frac{P(T1=n \mid D=d1)}{P(T1=n \mid D=d2)}, \frac{P(T1=n \mid D=d2)}{P(T1=n \mid D=d1)}\right)$$

However, many subjects follow the *feature difference strategy* (Skov & Sherman, 1987; Slowiaczek, Klayman, Sherman & Skov, 1992; Nelson, 2005). This strategy involves calculating the absolute difference in feature likelihoods for each test, and ordering the test with the highest fDiff:

$$fDiff(T1) = \left| P(T1=p \mid D=d1) - P(T1=p \mid D=d2) \right|$$ and

$$fDiff(T2) = \left| P(T2=p \mid D=d1) - P(T2=p \mid D=d2) \right|$$

Thus, Skov and Sherman, and Slowiaczek et al., concluded that many subjects use a suboptimal heuristic strategy that is highly correlated with the optimal strategy. Remarkably, however, both the claims (1) that Bayesian diagnosticity (and/or log diagnosticity) are theoretically optimal, and (2) that the feature difference strategy only imperfectly approximates optimal behavior, are in disrepute.

Both expected Bayesian diagnosticity and expected log diagnosticity are poorly behaved as optimal models. For instance, suppose that Test 1 were positive in 99% of people with Disease 1, and in 100% of the people with Disease 2. Suppose further that Test 2 were positive in 1% of people with Disease 1, and 99% of people with Disease 2. Test 1 leads, on average, to 50.5% probability of identifying the correct disease; Test 2 leads, on average, to 99% probability of correctly identifying the true disease. Clearly, Test 2 would be more helpful than Test 1 to differentiate between Disease 1 and Disease 2. Yet diagnosticity and log diagnosticity maintain that Test 1 is infinitely more useful than Test 2! Both diagnosticity measures hold that any test that offers greater-than-zero probability of obtaining 100% certainty of the true disease is infinitely useful. This bizarre claim is not a desirable property of an "optimal" model. (Nelson, 2005, 2008, discusses these and other theoretical flaws with the diagnosticity measures, and how redefining a single point cannot fix them.)

Better-motivated theoretical models of the value of information, such as information gain-KL distance (Lindley, 1956; Oaksford & Chater, 1994), probability gain (error reduction, Baron, 1981, 1985), and impact (Wells & Lindsay, 1980; Klayman and Ha, 1987, pp. 219–220; Nickerson, 1996; Nelson, 2008) behave reasonably in this medical diagnosis scenario, and do not suffer from the diagnosticity measures' above-mentioned theoretical flaws.

Does the feature difference strategy also approximate these better-motivated theoretical models? In fact, it exactly corresponds to impact! The feature with the highest fDiff also always has the highest impact, irrespective of the prior probabilities of the diseases and the specific feature probabilities (Nelson, 2005, footnote 2).

In other words, closer analysis of the supposedly-optimal theoretical models used by some experimenters, and the supposedly-suboptimal heuristics used by some subjects, showed that the subjects' heuristic strategy corresponds to a normative model (impact) that is theoretically superior to the normative model that the experimenters had in mind! Put in the context of Marr's (1983) levels of analysis, consideration of subjects' behavior at the algorithmic level can inform thinking about the kinds of computational-level models (normative theories) that are most appropriate.

There may be cases where we know precisely what nature is trying to optimize, as well as the precise statistics of current or evolutionary environments. But in the case of perception (e.g. Hoffman, in press), memory, information search, category formation, and other topics of interest in Bayesian and rational analyses, the correct function to optimize is seldom clear. Cohen (1981) suggested that normative theories ultimately derive their normativity from reflective intuition. In any case, the relationship between the (algorithmic) feature difference strategy and the (computational-level) impact model illustrates how algorithmic and computational level analyses can mutually inform each other, as Chater, Oaksford, Nakisa and Redington (2003) discussed.

Baron (2004) noted that utilities are formed on the basis of reflection, and are constantly being modified. As a pragmatic matter, cognitive science would be wise to treat candidate normative models in similar fashion (also see McKenzie, 2003). When there are clear and robust discrepancies between human behavior and a particular theoretical model, the normative status of the theoretical model should be questioned, and not only the rationality or adaptiveness of the human behavior.

Do all subjects use the feature difference strategy? No. As O&C discuss, the means with which information is presented is important. Different people use a variety of strategies, especially when environmental probabilities are presented in the *standard probability format*, with explicit prior probabilities and likelihoods. Gigerenzer & Hoffrage (1995) and Cosmides & Tooby (1996) show that the standard probability format is not the most meaningful, and that frequency formats better facilitate Bayesian reasoning. Personal experience of environmental probabilities may be even more effective. When environmental probabilities are learned through personal experience, the vast majority of subjects maximize probability of correct guess (*probability gain*), rather than impact or information gain (Nelson, McKenzie, Cottrell, & Sejnowski, submitted).

One way to conceptualize standard probability format subjects' use of the feature difference strategy is that that strategy exactly corresponds to impact, one normative model. But in a deeper sense, if we suppose that subjects actually want to maximize probability gain, subjects' use of the feature difference strategy may reflect that subjects have good intuitions about which algorithms best approximate their true goals. Impact (which the feature difference strategy implements) more reliably approximates probability gain than do Bayesian diagnosticity or log diagnosticity (Nelson, 2005), and impact is easily calculated when the standard probability format is used.

Is human cognition is optimal in the general sense? Organisms' adaptation to their environs can be impressive. For instance, insects' flight length distributions appear well-calibrated to natural environmental probabilities (Viswanathan, Buldyrev, Havlin, da Luz, Raposo, & Stanley, 1999). But the modern world is so different from evolutionary history that broad claims of optimality might be implausible. Perhaps people's love of sugar, salt and fat was presumably adaptive in evolutionary history. But today, people might be better served by taste that would more quickly satiate when adequate sugar, salt and fat have been obtained. Similarly, a disproportionate fear of plane crashes, as compared to car crashes, may reflect a perceptual difficulty that would not have arisen if word-of-mouth and personal experience, as opposed to modern media, were used to convey environmental probabilities. It appears that the rate of human genetic evolution has been increasing (Hawks, Wang, Cochran, Harpending, & Moyzis, 2007). This may facilitate some adaptation to modern environments, including those of our own making, over phylogenetic time. For the moment, however, it is most helpful to convey probabilities in a way that takes advantage of our evolutionary experience in learning environmental probabilities via experience, with feedback. When people understand the environmental probabilities, they will be in the best position to optimize their information-acquisition goals as well.

Sources Cited

Baron, J. (1985). *Rationality and intelligence*. Cambridge, England: Cambridge University Press.

Baron, J. (2004). Normative models of judgment and decision making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making*, pp. 19-36.

Chater, N., Oaksford, M., Nakisa, R. & Redington, M. (2003). Fast, frugal, and rational: How rational norms explain behavior. *Organizational Behavior and Human Decision Processes*, 90, 63–86.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317–370.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58, 1-73.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats, *Psychological Review*, l02(4), 684-704.

Good, I. J. (1950). *Probability and the weighing of evidence*. New York: Griffin.

Good, I. J. (1975). Explicativity, corroboration, and the relative odds of hypotheses. *Synthese*, 30, 39–73.

Hawks, J., Wang, E. T., Cochran, G. M., Harpending, H. C., & Moyzis, R. K. (2007). Recent acceleration of human adaptive evolution. *Proceedings of the National Academy of Sciences of the USA*, 104(52), 20753–20758.

Hoffman, D. (in press). The interface theory of perception: Natural selection drives true perception To swift extinction. In S. Dickinson, M. Tarr, A. Leonardis & B. Schiele (Eds.), *Object Categorization: Computer and Human Vision Perspectives*. Cambridge, UK: Cambridge Univ. Press

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information. *Psychological Review*, 94, 211–228.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, 27, 986–1005.

Marr, D. (1983). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman

McKenzie, C. R. M. (2003). Rational models as theories – not standards – of behavior. *Trends in Cognitive Sciences*, 7(9), 403-406

Nelson, J. D. (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact and information gain. *Psychological Review*, 112(4), 979-999.

Nelson, J. D. (2008). Towards a rational theory of human information acquisition. In Oaksford, M & Chater, N (Eds.), *The probabilistic mind: Prospects for rational models of cognition* (pp. 143-163). Oxford: Oxford University Press.

Nelson, J. D.; McKenzie, C. R. M.; Cottrell, G. W.; Sejnowski, T. J. (submitted). Experience-based, but not summary statistics-based, human information acquisition optimizes probability gain.

Nickerson, R. S. (1996). Hempel's paradox and Wason's selection task: Logical and psychological puzzles of confirmation. *Thinking and Reasoning*, 2, 1–32.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608–631.

Oaksford, M., & Chater, N. (2007). *Bayesian Rationality: the Probabilistic Approach to Human Reasoning*. Oxford, UK: Oxford Univ. Press

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, 22, 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 & Cognition*, 20, 392–405.

Viswanathan, G. M., Buldyrev, S. V., Havlin, S., da Luz, M. G., Raposo, E. P., & Stanley, H. E. (1999, October 28). Optimizing the success of random searches. *Nature*, 401, 911–914.

Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, 88, 776–784.