models of cognition. Bayesian statistical methods are important, useful, and should play a central role in analyzing all models of cognition, including Bayesian ones. The target article views this as a side issue, but I think it is a fundamental element of the path to enlightenment.

Cognitive systems optimize energy rather than information

doi:10.1017/S0140525X11000355

Arthur B. Markman and A. Ross Otto
Department of Psychology, University of Texas, Austin, TX 78712.
markman@psy.utexas.edu rotto@mail.utexas.edu
http://homepage.psy.utexas.edu/homepage/Faculty/Markman/PSY394/kreps11.html

Abstract: Cognitive models focus on information and the computational manipulation of information. Rational models optimize the function that relates the input of a process to the output. In contrast, efficient algorithms minimize the computational cost of processing in terms of time. Minimizing time is a better criterion for normative models, because it reflects the energy costs of a physical system.

Two parallel developments in the 1940s set the stage both for the cognitive revolution of the 1950s and for the discussion presented in the target article. The development of information theory explored ways to characterize the information content of a message and ways to consider how to best pass messages (Shannon 1949). At the same time, the architecture for digital computing led to advances in discrete mathematics that facilitated the analysis of the efficiency of algorithms (Turing 1950).

One consequence of the cognitive revolution was that that it became common to characterize the mind as a computational device. Thus, researchers began to formulate theories of mental processes in computational terms. As Marr (1982) points out, a process can be defined at either a computational level or an algorithmic level of description. At the computational level, the process is defined by a mapping between information available at the start and end of the process. For example, Anderson (1990) advocates a Bayesian, "rational-level" analysis of the information relationship between inputs and outputs of a system. At the algorithmic level, a process is specified in terms of a set of steps that implements this computational-level description. Any given algorithm can be analyzed for its efficiency in time. The efficiency of a cognitive process can be established at either the computational level of description or at the algorithmic level. The Bayesian approaches described in the target article are focused on defining the optimality of a cognitive process at the computational level (Anderson 1990; Tenenbaum & Griffiths 2001). Anderson (1990) does point out that computational costs can also play a role in determining a rational model, but, in practice, these considerations did not have a significant influence on the structure of his rational models.

The danger in casting optimality purely at the computational level is that human cognition is implemented by a physical system. Indeed, it has been proposed that any characterization of the optimality of actions or beliefs should take into account the resource-limited nature of the human cognitive apparatus (Cherniak 1986; Stanovich & West 1998). As the target article points out, the brain consumes a significant amount of energy. Thus, energy minimization is likely to be an important constraint on cognitive processing.

The idea that energy-minimization is an important constraint on cognitive processing is implicit in the focus on efficient computational procedures. We do not suppose that the metabolic cost of cognition is completely invariant of the type of thinking that people are engaged in, but marginal changes in metabolic rates attributed to different types of cognition pale in comparison to the metabolic cost of simply keeping the brain running. Thus, the time taken by a process is a good proxy for energy conservation. On this view, for example, habits minimize energy, because they allow a complex behavior to be carried out quickly (e.g., Logan 1988; Schneider & Shiffrin 1977).

Of course, effort-minimization is not the only constraint on cognitive processing. It is crucial that a process be carried out to a degree sufficient to solve the problem faced by the individual. This view was central to Simon's (1957b) concept of *satisficing*. This view suggested that cognitive processes aim to expend the minimal amount of effort required to solve a problem. On this view, the costs of additional effort outweigh the gains in decision accuracy. This idea was elaborated in the effort accuracy framework developed by Payne et al. (1993). Their work examined the variety of strategies that people utilize in order to balance decision accuracy with effort – the cognitive costs of gathering and integrating information about choice attributes – in decision-making. Payne et al. point out that these strategies differ both in the effort required to carry them out as well as in their likelihood of returning an accurate response. People negotiate the trade-off between effort and accuracy by selecting decision strategies that minimize the effort required to yield an acceptable outcome from a choice.

A key shortcoming, then, of the Bayesian Fundamentalist approach is that it optimizes the wrong thing. The ideal observer or actor defined purely in terms of information is quite useful, but primarily as a point of comparison against human cognitive or sensory abilities rather than as a statement of what is optimal as a cognitive process (e.g., Geisler 1989). A definition of optimal behavior needs to take energy minimization into account. Thus, the key limitation of Bayesian Fundamentalism is that it focuses selectively on optimality of information processing rather than on the combination of information and time.

Enlightenment grows from fundamentals

doi:10.1017/S0140525X11000367

Daniel Joseph Navarro and Amy Francesca Perfors
School of Psychology, University of Adelaide, Adelaide, SA 5005, Australia.
daniel.navarro@adelaide.edu.au amy.perfors@adelaide.edu.au
http://www.psychology.adelaide.edu.au/personalpages/staff/danielnavarro/
http://www.psychology.adelaide.edu.au/personalpages/staff/amyperfors/

Abstract: Jones & Love (J&L) contend that the Bayesian approach should integrate process constraints with abstract computational analysis. We agree, but argue that the fundamentalist/enlightened dichotomy is a false one: Enlightened research is deeply intertwined with – and to a large extent is impossible without – the basic, fundamental work upon which it is based.

Should Bayesian researchers focus on "enlightened" modelling that seriously considers the interplay between rational and mechanistic accounts of cognition, rather than a "fundamentalist" approach that restricts itself to rational accounts only? Like many scientists, we see great promise in the "enlightened" research program. We argue, however, that enlightened Bayesianism is deeply reliant on research into Bayesian fundamentals, and the fundamentals cannot be abandoned without greatly affecting more enlightened work. Without solid fundamental work to extend, enlightened research will be far more difficult.

To illustrate this, consider the paper by Sanborn et al. (2010a), which Jones & Love (J&L) consider to be "enlightened" as it seeks to adapt an ideal Bayesian model to incorporate insights about psychological process. To achieve this, however, it relies heavily upon work that itself would not have counted as "enlightened." The comparison between Gibbs sampling and particle filtering as rival process models grew from "unenlightened" research that used these algorithms purely as methodological tools. As such, without this "fundamentalist" work the enlightened paper simply would not have been written.

Enlightened research can depend on fundamentals in other ways. Rather than adapt an existing Bayesian model to incorporate process constraints, Navarro and Perfors (2011) used both Bayesian fundamentals (an abstract hypothesis space) and process fundamentals (capacity limitations on working memory) as the foundations of an analysis of human hypothesis testing. Identifying a conditionally optimal learning strategy, given the process constraint, turned out to reproduce the "positive test strategy" that people typically employ (Wason 1960), but only under certain assumptions about what kinds of hypotheses are allowed to form the abstract hypothesis space. This analysis, which extended existing work (Klayman & Ha 1987; Oaksford & Chater 1994) and led us to new insights about what kinds of hypotheses human learners "should" entertain, could not have been done without "fundamentalist" research into *both* the statistical and the mechanistic basis of human learning.

Not only do "enlightened" papers *depend* on fundamental ones, we suggest that they are a natural *outgrowth* of those papers. Consider the early work on Bayesian concept learning, which contained a tension between the "weak sampling" assumption of Shepard (1987) and the "strong sampling" assumption of Tenenbaum and Griffiths (2001). When strong sampling was introduced, it would presumably have counted as "fundamentalism," since the 2001 paper contains very little by way of empirical data or consideration of the sampling structure of natural environments. Nevertheless, it served as a foundation for later papers that discussed exactly those issues. For instance, Xu and Tenenbaum (2007a) looked at how human learning is shaped by explicit changes to the sampling model. This in turn led Navarro et al. (in press) to propose a more general class of sampling models, and to pit them all against one another in an empirical test. (It turned out that there are quite strong individual differences in what people use as their "default" sampling assumption.) The change over time is instructive: What we observe is a gradual shift from simpler "fundamentalist" papers that develop the theory in a reduced form, towards a richer framework that begins to capture the subtleties of the psychology in play.

Even J&L's own chosen examples show the same pattern. Consider the Kemp et al. (2007) article, which J&L cite as a prime example of "fundamentalist" Bayesianism, since it introduces no new data and covers similar ground to previous connectionist models (Colunga & Smith 2005). Viewing the paper in isolation, we might agree that the value added is minor. But the framework it introduced has been a valuable tool for subsequent research. An extension of the model has been used to investigate how adults learn to perform abstract "second order" generalizations (Perfors & Tenenbaum 2009) and to address long-debated issues in verb learning (Perfors et al. 2010). A related model has even been used to investigate process-level constraints; Perfors (in press) uses it to investigate whether or not memory limitations can produce a "less is more" effect in language acquisition. It is from the basic, fundamental research performed by Kemp et al. (2007) that these richer, more enlightened projects have grown.

Viewed more broadly, the principle of "enlightenment growing from fundamentals" is applicable beyond Bayesian modelling; our last example is therefore an inversion. We suggest that J&L understate the importance of computational considerations in good process modelling. For instance, one of their key examples comes from Sakamoto et al. (2008), who consider mechanistic models of category learning. That paper might be characterized as a "fundamentalist" work in process modelling, insofar as it gives no consideration to the computational level issues that pertain to their choice of learning problem. As consequence of this "process fundamentalism," the "rational" model that paper employs is not actually a rational model. It is highly mis-specified for the problem of learning time-inhomogeneous categories. In recent work (Navarro & Perfors 2009), we discuss this concern and introduce extensions to the experimental framework aimed at highlighting the computational considerations involved; at present, we are working on model development to build on this. However, the goal in our work is *not* to deny the importance of process, but to learn which aspects of human behaviour are attributable to computational level issues and which aspects reflect process limitations. In this case, that goal is met by building on fundamental work on the process level (i.e., Sakamoto et al.'s 2008 paper) and adding computational considerations. In general, attaining the goal of "enlightened" research is possible only if fundamentals on both levels are taken seriously – if researchers deny neither psychological mechanism *nor* ideal computation.

Like J&L, we believe that it is the *interaction* between the twin considerations of computation and process that leads us to learn about the mind. However, this should not lead us to abandon work that focuses on only one of these two components. Enlightened research is constructed from the building blocks that fundamental work provides.

The illusion of mechanism: Mechanistic fundamentalism or enlightenment?

doi:10.1017/S0140525X11000379

Dennis Norris
MRC Cognition and Brain Sciences Unit, Cambridge CB2 7EF, United Kingdom.
dennis.norris@mrc-cbu.cam.ac.uk
http://www.mrc-cbu.cam.ac.uk/people/dennis.norris

Abstract: Rather than worrying about Bayesian Fundamentalists, I suggest that our real concern should be with Mechanistic Fundamentalists; that is, those who believe that concrete, but frequently untestable mechanisms, should be at the heart of all cognitive theories.

Jones & Love (J&L) suggest that we should reject Bayesian Fundamentalism in favour of Bayesian Enlightenment, thus combining Bayesian analysis with mechanistic-level models. This raises two questions: Who are these Bayesian Fundamentalists and what is a mechanistic-level model?

First, let us go in search of Bayesian Fundamentalists. As I read the target article, I began to wonder how it could be that I'd never encountered a Bayesian Fundamentalist. If these ideas are so pervasive, then surely J&L could quote at least one author who has made a clear statement of the Bayesian Fundamentalist programme? From the first line of the abstract it appears that the main proponent of Bayesian Fundamentalism must be Anderson (1990) with his Rational Analysis framework, and his suggestion that behaviour can often be explained by assuming that it is optimally adapted to its purpose and the environment. In criticising rational analysis, J&L argue that "Rather than the globally optimal design winning out, often a locally optimal solution . . . prevails. . . . Such non-behavioral factors are enormously important to the optimization process, but are not reflected in rational analyses, as these factors are tied to a notion of mechanism, which is absent in rational analyses" (sect. 5.3, paras. 3 and 5).

A similar concern about the limitations of rational analysis can be found in the following quotation: "My guess is that short-term memory limitations do not have a rational explanation. . . . [T]hey reflect the human trapped on some local optimum of evolution" (Anderson 1990, pp. 91–92). These cautionary words on the dangers of relying entirely on rational explanations were written by the arch-Fundamentalist himself. Is there really a difference, then, between these two positions?