that combine these sources of Bayesian information to produce simulations, neural net architectures that reactivate previously experienced states have much potential for doing so.

Because simulators and situated conceptualizations occur in nonhumans, they offer a natural account of conceptual processing across species (Barsalou 2005). If so, the kind of Bayesian analysis just described applies comparatively, perhaps via somewhat common forms of optimality arising continuously across evolution. Where humans are likely to differ is in the linguistic control of this architecture, with words activating simulators, and larger linguistic structures specifying situated conceptualizations compositionally and productively (Barsalou 1999; 2008b).

Bayesian analysis can also be applied to linguistic forms, similarly to how it can be applied to simulators and situated conceptualizations. On activating a word, the probability that other words become active reflects a distribution of priors over these words, constrained by likelihoods, given other words in the context. As research shows increasingly, the statistical structure of linguistic forms mirrors, to some extent, the structure of conceptual knowledge grounded in the modalities (e.g., Andrews et al. 2009; Barsalou et al. 2008; Louwerse & Connell 2011). Of interest is whether similar versus different factors optimize the retrieval of linguistic forms and conceptual knowledge, and what sorts of factors optimize their interaction.

Finally, the grounded perspective assumes that cognition relies inherently on the body, the physical environment, and the social environment, not just on classic cognitive mechanisms (Barsalou 2008a). Because cognition does not occur independently of these other systems, characterizing their structure is essential, analogous to the importance of characterizing the physical environment in Bayesian analysis.

For all these reasons, grounded cognition offers a natural approach for practicing and achieving Bayesian Enlightenment. As cognition emerges from bodily and neural mechanisms through interactions with physical and social environments, numerous forms of optimization undoubtedly occur at many levels. Fully understanding these optimizations seems difficult – not to mention unsatisfying – unless all relevant levels of analysis are taken into account. Indeed, this is the epitome of cognitive science.

Mechanistic curiosity will not kill the Bayesian cat

doi:10.1017/S0140525X11000215

Denny Borsboom,[a] Eric-Jan Wagenmakers,[a] and Jan-Willem Romeijn[b]

[a]*Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands;* [b]*Department of Philosophy, University of Groningen, 9712 GL Groningen, The Netherlands.*

dennyborsboom@gmail.com ej.wagenmakers@gmail.com j.w.romeijn@rug.nl http://sites.google.com/site/borsboomdenny/dennyborsboom http://www.ejwagenmakers.com/ http://www.philos.rug.nl/~romeyn

Abstract: Jones & Love (J&L) suggest that Bayesian approaches to the explanation of human behavior should be constrained by mechanistic theories. We argue that their proposal misconstrues the relation between process models, such as the Bayesian model, and mechanisms. While mechanistic theories can answer specific issues that arise from the study of processes, one cannot expect them to provide constraints in general.

Jones & Love (J&L) argue that Bayesian approaches to human behavior should attend more closely to cognitive and neural mechanisms. Because mechanisms play such an important role in their target article, it is important to get a clear idea of what mechanisms are and what they are good for. J&L unfortunately do not clarify the term. They get closest when, in section 5.1, they mention the "notion of *mechanism* (i.e., process or representation)" (para. 3, emphasis J&L's). This treatment is, in our view, less accurate than would be needed to support the strong claims the target article makes with regard to the status of Bayesian approaches to cognition. When the concepts of mechanism and process are fleshed out, these claims might well turn out to be untenable.

Roughly, processes and mechanisms relate as follows. A *process* concerns the change of a system over time. The easiest way to think about this is as a path through a set of possible states the system can be in. A *process model* is a description of this path, detailing how each new state (or its probability) depends on its previous state(s). In the behavioral sciences, such a model can often be represented by a flowchart. A *mechanism*, by contrast, is not a process but a system. It typically has parts that work together to implement an input-output relation. For instance, smoking (input) robustly produces lung cancer (output), through a causal mechanism (smoke brings tar into the lungs which leads to mutations). A *mechanistic model* is a representation of the way the parts of the system influence one another. Typically, this is represented as a directed graph or a circuit diagram. Mechanisms are closely tied to the notion of *function*, because they are often studied and discovered by pursuing questions of the "how does this work?" variety (e.g., "how does smoke cause cancer?").

Now, a Bayesian model is a process model, not a mechanistic model. This is not, as J&L believe, because "the Bayesian metaphor is tied to a mathematical ideal and thus eschews mechanism altogether" (sect. 2.2, para. 3), but simply because it describes how a rational agent moves through an abstract state-space of beliefs (probabilities of hypotheses) when confronted with evidence (data): all the model says is how a rational agent is to move to new belief state at $t+1$, given the prior belief state and evidence available at time t. This has nothing to do with the fact that the model is mathematically formalized. Mechanistic and causal models have mathematical formalizations just as well (e.g., see Pearl 2000). The Bayesian model is simply not a mechanistic model because it is a process model. To argue that the Bayesian model fails to capture mechanisms is much like arguing against relativity theory because it provides no mechanistic detail on how clocks slow down when moved.

Clearly there have to be mechanisms that allow the belief-updating process to run, and these mechanisms are likely to reside in our brain. One may profitably study these mechanisms and even provide support for Bayesian models with that. A good question, for instance, that may receive a mechanistic answer is, "How do people implement belief updating?" (Ma et al. 2006). Note that, by a suitable choice of variables and probabilistic relations, any sequence of belief states can be viewed as resulting from a Bayesian update (cf. Albert 2001). But say that we have independently motivated our starting points and found a convincing fit with the behavioral data of the belief dynamics (e.g., Brown et al. 2009). J&L then seem to suggest how this model might be given a mechanistic underpinning when they say that "belief updating of Bayes' Rule [amounts] to nothing more than vote counting" (sect. 7, para. 3). To us, the vote-counting idea seems just about right, since vote counting is about all that neurons can do if they are supposed to be ultimately implementing the process. We would add that mechanisms might also support the Bayesian account by providing independent motivations for choosing the variables and relations that make up the model.

Another good question is, "Why do people deviate from optimality in circumstance X?" The Bayesian model cannot explain such deviations directly, since it presupposes optimality. However, without a clear definition of optimality, as given by the Bayesian model, it would be impossible to detect or define such deviations in the first place: Without the presence of rationality, the concept of bounded rationality cannot exist. What's more, suboptimal behavior can be elucidated by giving it the semblance of optimality within a Bayesian model. Those models then suggest what potentially irrational assumptions real agents make; the Bayesian models of reasoning behavior (Oaksford & Chater 2007) are a case in point.

J&L are not satisfied by this type of mechanistic support for Bayesian models; they argue that mechanistic theories should *constrain* the Bayesian model. However, it is unclear why exactly we should believe this. Surely, it does not matter for the empirical adequacy of the Bayesian process models whether peoples' beliefs are physically realized as activation networks in their frontal lobe, as global properties of their brain states, or as bursts of currents running in their big left toe. What matters is that the behavioral data are fitted within an independently motivated and predictively accurate model. In fact, if it turned out that dualism were correct after all, and belief revision actually went on in Cartesian mental stuff, that would not hurt the Bayesian analysis one bit – as long as the mental stuff updated its beliefs properly. Thus, the relation between Bayesian explanation and mechanistic accounts is asymmetric: While the finding that there is a mechanistic realization of Bayesian belief revision supports the Bayesian view, not finding such a mechanistic realization does not refute the theory. The only facts that can refute the Bayesian explanation are empirical facts about human behavior.

More varieties of Bayesian theories, but no enlightenment

doi:10.1017/S0140525X11000227

Jeffrey S. Bowers[a] and Colin J. Davis[b]

[a]*School of Psychology, University of Bristol, Clifton, Bristol BS8 1TU, United Kingdom;* [b]*Department of Psychology, Royal Holloway University of London, Egham Hill TW20 0EX, and School of Psychology, University of Bristol, Clifton, Bristol BS8 1TU, United Kingdom.*

j.bowers@bris.ac.uk **c.davis@rhul.ac.uk**

http://psychology.psy.bris.ac.uk/people/jeffbowers.htm

http://www.pc.rhul.ac.uk/staff/c.davis/

Abstract: We argue that Bayesian models are best categorized as *methodological* or *theoretical*. That is, models are used as tools to constrain theories, with no commitment to the processes that mediate cognition, or models are intended to approximate the underlying algorithmic solutions. We argue that both approaches are flawed, and that the Enlightened Bayesian approach is unlikely to help.

We agree with many points raised by Jones & Love (J&L) in the target article, but do not think that their taxonomy captures the most important division between different Bayesian approaches; and we question their optimism regarding the promise of the Enlightened Bayesian approach.

In our view, the critical distinction between Bayesian models is whether they are being used as a tool or a theory, what we have called the Methodological and Theoretical Bayesian approaches, respectively (Bowers & Davis, submitted). According to the Methodological approach, Bayesian models are thought to provide a measure of optimal performance that serves as a benchmark against which to compare actual performance. Researchers adopting this perspective highlight how often human performance is near optimal, and such findings are held to be useful for constraining a theory. (Whatever algorithm the mind uses, it needs to support behaviour that approximates optimal performance.) But there is no commitment to the claim that the algorithms that support perception, cognition, and behaviour approximate Bayesian computations.

By contrast, according to the Theoretical approach, the mind is claimed to carry out (or closely approximate) Bayesian analyses at the algorithmic level; this perspective can be contrasted with the view that the mind is a rag-bag of heuristics. For example, when describing the near-optimal performance of participants in making predictions about uncertain events, Griffiths and Tenenbaum (2006) write: "These results are inconsistent with claims that cognitive judgments are based on non-Bayesian heuristics" (p. 770).

Unfortunately, it is not always clear whether theorists are adopting the Methodological or the Theoretical approach, and at times, the same theorists endorse the different approaches in different contexts. Nevertheless, this is the key distinction that needs to be appreciated in order to understand what claims are being advanced, as well as to evaluate theories. That is, if Bayesian models are used as a tool to constrain theories, then the key question is whether this tool provides constraints above and beyond previous methods. By contrast, if the claim is that performance is supported by Bayesian-like algorithms, then it is necessary to show that Bayesian theories are more successful than non-Bayesian theories.

In our view there are two main problems with the Methodological Bayesian approach. First, measures of optimality are often compromised by the fact Bayesian models are frequently constrained by performance. For instance, Weiss et al. (2002) developed a Bayesian model of motion perception that accounts for an illusion of speed: Objects appear to move more slowly under low-contrast conditions. In order to accommodate these findings, Weiss et al. assumed that objects tend to move slowly in the world, and this prior plays a more important role under poor viewing conditions. One problem with this account, however, is that there are other conditions under which objects appear to move more quickly than they really are (Thompson et al. 2006). Stocker and Simoncelli's (2006) response to this problem is to note that their Bayesian theory of speed perception could account for the latter phenomenon as well:

> [I]f our data were to show increases in perceived speed for low-contrast high-speed stimuli, the Bayesian model described here would be able to fit these behaviors with a prior that increases at high speeds. (Stocker & Simoncelli 2006, p. 583)

The modification of Bayesian models in response to the data is widespread, and this renders the models more as descriptions of behaviour than as tools with which to measure optimality.

Second, even if a Bayesian model provides a good measure of optimal performance, it is not clear how the tool contributes to constraining theories. Under these conditions, a model can be supported or rejected because it does or does not match optimal performance, or more simply, a model can be supported or rejected because it does or does not capture human performance. The match or mismatch to data is sufficient to evaluate the model – the extra step of comparing to optimal performance is superfluous.

With regard to the Theoretical Bayesian approach, the key question is whether a Bayesian model does a better job in accounting for behaviour compared to non-Bayesian alternatives. However, this is rarely considered. Instead, proponents of this approach take the successful predictions of a Bayesian model as support for their approach, and often ignore the fact that non-Bayesian theories might account for the data just as well. We are not aware of any psychological data that better fit a Bayesian as compared to a non-Bayesian alternative.

What about the promise of the Bayesian Enlightenment approach? On our reading, this perspective encompasses both the theories that we would call Methodological (e.g., the adaptive heuristic approach of Gigerenzer), and the theories that we would call Theoretical (e.g., demonstrations that Bayesian computations can be implemented in neural wetware are considered Enlightened). Thus, the above criticisms apply to the Bayesian Enlightenment approach as well.

With regard to Enlightened theories that take the form of heuristics, it is not clear that Bayesian models are providing any constraints. For example, we are not aware of any instance