Mechanistic curiosity will not kill the Bayesian cat

doi:10.1017/S0140525X11000215

Denny Borsboom, Eric-Jan Wagenmakers, and Jan-Willem Romeijn

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.

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 that their 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 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 believe 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 answer some of the questions that concern the change of a system over time. The easiest 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 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?”).

Another good question is, “Why do people deviate from optimality in circumstances 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
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 (Thorupson 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:

[If] 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 heuristics 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

More varieties of Bayesian theories, but no enlightenment

doi:10.1017/S0140525X11000227

Jeffrey S. Bowersa and Colin J. Davisb

aSchool of Psychology, University of Bristol, Clifton, Bristol BS8 1TU, United Kingdom; bDepartment 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@bristol.ac.uk c.davis@rhul.ac.uk

http://psychology.psyrhul.ac.uk/people/jeffbowers.htm

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

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 (Thorupson 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:

[If] 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 heuristics 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