computation, it also has some advantages which we discuss. Therefore, it is disingenuous to claim that this type of research in causal induction arises from or differentially supports a Bayesian perspective.

The second way to disambiguate the causal models, temporal precedence, has been the object of much associative analysis (e.g., Pavlov 1927). The causal arrows in the models can be replaced with "leads to" and, if the observer can discriminate order, they could easily discriminate the common cause, A leads to B and C, from the chain, A leads to B that leads to C models. Bayesian fundamentalists researching causal induction (e.g., Lagnado & Sloman 2006) have, indeed, shown that with temporal information people can discriminate the models. However, the timing information does not arise directly from Bayesian computations. Again it is the *supervisor* that disambiguates the data by using timing. However, learning orders emerges easily from a connectionist perspective. Associative nets, like neuronal processes, have activations that decay with time. In ($A \rightarrow B \rightarrow C$), when C finally comes along, A activation will have decayed more than B activation, so a stronger link between B and C will form. Baetu and Baker (2009) have reported a series of experiments in which they have studied how people form causal chains from experience with the individual links (generative and preventative). People are good at assembling these independently learned links into chains. But, most important, a modification of a simple associative model (the autoassociator; McClelland & Rumelhart 1988) generates representations of causal structure, and predicts participants' behavior.

Finally, the associative structures have a certain face validity for psychological processes that the Bayesian frames do not. In associative terms, causal chains are represented as associative strengths and not likelihood ratios. Order and timing can flow naturally from them. They represent causes and effects of different magnitudes, and not just binary (absent/present) events. Causes and activations may be weak or strong, and not just present or absent.

What does this say about J&L's thesis? First, we agree that Bayesianism must progress beyond fundamentalism. Indeed, much of the research concerning the *supervisor* can lead the unwary believer to the unwarranted conclusion that this work discriminates Bayesian computations from others. Second, J&L argue that the Bayesian analysis can prosper at all levels. In our rather simple case, it certainly does not account for the *supervisor*. It is not clear how it could in a principled way. Third, they argue that Bayesian analyses should become closely linked to psychological mechanism, and we agree; but we argue that associative structure may already be there. For instance, we are now closer to understanding how a prediction-error algorithm (e.g., Rescorla & Wagner 1972) might be implemented in the brain (Kim et al. 1998; Waelti et al. 2001).

In conclusion, we agree with J&L that Agnostic Bayesian nets offer a powerful method for artificial intelligence and that, if elaborated, can learn about or represent any finite data set – but so could an associative net. However, the question for psychological process is one of parsimony and mechanistic plausibility, and we are not convinced that J&L have demonstrated this contribution. We would be more convinced if they had described a single instance where a Bayesian analysis produced a realistic psychological mechanism or empirical result.

Integrating Bayesian analysis and mechanistic theories in grounded cognition

doi:10.1017/S0140525X11000197

Lawrence W. Barsalou

Department of Psychology, Emory University, Atlanta, GA 30322.

barsalou@emory.edu

http://www.psychology.emory.edu/cognition/barsalou/index.html

Abstract: Grounded cognition offers a natural approach for integrating Bayesian accounts of optimality with mechanistic accounts of cognition, the brain, the body, the physical environment, and the social environment. The constructs of *simulator* and *situated conceptualization* illustrate how Bayesian priors and likelihoods arise naturally in grounded mechanisms to predict and control situated action.

In the spirit of Bayesian Enlightenment, as suggested by Jones & Love (J&L), grounded cognition offers architectural mechanisms that naturally afford Bayesian analysis. In particular, the constructs of *simulator* and *situated conceptualization* illustrate the potential for integrating explanations across Marr's (1982) computational, algorithmic, and implementation levels. Many other grounded constructs also undoubtedly offer similar potential.

In *perceptual symbol systems*, a *simulator* is a dynamical system distributed across the modalities that process a category's properties, aggregating information about diverse instances, comparable to a concept in traditional theories (Barsalou 1999; 2003a). The *beer* simulator, for example, aggregates information about how beer looks, smells, and tastes, how it is consumed, how we feel afterwards, and so on. If someone experiences a wide variety of beers (e.g., American, Belgian, English, German, Czech, Indian, Thai, etc.), the beer simulator captures diverse multi-modal states about the category (modeled naturally with neural net architectures; e.g., Pezzulo et al. 2011). On a given occasion, the beer simulator dynamically produces one of many specific beer simulations, from an infinite set possible. A natural way of thinking about the space of possible simulations within a simulator is as a space of Bayesian priors, with some simulations being more likely than others. Furthermore, the strength of a given prior can be assessed empirically (rather than simply assumed), reflecting well-established and readily measured factors, including how frequently and recently category instances have been experienced, how ideal or preferred they are, their similarity to other instances, and so forth (Barsalou 1985; 1987).

In grounded approaches to the conceptual system, a *situated conceptualization* is a situation-relevant simulation from a simulator embedded in the representation of a likely background situation (Barsalou 2003b; 2008c; Yeh & Barsalou 2006). One situated conceptualizations of *chair*, for example, represents a chair on a *jet*, embedded in a jet setting, accompanied by relevant actions and mental states. A natural way of thinking about a situated conceptualization is as representational structure that captures and produces Bayesian likelihoods. Entering a jet setting, for example, may activate the situated conceptualization for jet chairs, producing the expectancy that this specific type of chair will be experienced shortly. Similarly, seeing a jet chair activates expectancies about how to interact with it, how it will feel to operate, and so on. Again, the statistical structure of situated conceptualizations can be assessed empirically, through objective assessments of the environment, subjective estimates of co-occurrence, et cetera. In general, much evidence demonstrates that situated conceptualizations produce part-whole inferences to support diverse forms of reasoning (e.g., Barsalou et al. 2003; 2005; Wilson-Mendenhall et al. 2011).

Together, simulators and situated conceptualizations naturally produce Bayesian inferences. Before entering a building for the first time, priors associated with the *chair* simulator produce situation-independent inferences about chairs likely to be encountered (e.g., kitchen chairs, easy chairs, office chairs), whereas likelihoods emerging from relevant situated conceptualizations produce inferences about likely chairs to be found in this particular context (e.g., living rooms of Buddhist friends). From the perspective of grounded cognition, priors and likelihoods are combined to produce simulations that prepare agents for what is likely to exist in the world, for how to act on the world, and for the mental states likely to result (Barsalou 2009; Barsalou et al. 2007). Because such simulations utilize the modalities for perception, action, and internal states, representations in these modalities become primed, thereby facilitating expected interaction with the environment. Although architectures remain to be developed 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