Behavioral & Brain Sciences, *34*, pp. 212-213
Comment on Jones & Love: "Bayesian Fundamentalism or Enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition"

Bayesian computation and mechanism: Theoretical pluralism drives scientific emergence

David K. Sewell and Daniel R. Little
The University of Melbourne

Stephan Lewandowsky
The University of Western Australia

The breadth-first search adopted by Bayesian researchers to map out the conceptual space and identify what the framework can do is beneficial for science and reflective of its collaborative and incremental nature. Theoretical pluralism among researchers facilitates refinement of models within various levels of analysis, which ultimately enables effective cross-talk between different levels of analysis.

The target article by Jones and Love (henceforth J&L) is another entry to the recent debate contrasting the merits of Bayesian and more mechanistic modeling perspectives (e.g., Griffiths, Chater, Kemp, Perfors, & Tenenbaum, 2010; McClelland et al., 2010). Regrettably, much of this debate has been tainted by a subtext that presupposes the approaches to be adversarial, rather than allied (see e.g., Feldman, 2010; Kruschke, 2010). J&L are correct in asserting that research agendas pitched at different levels of analysis will investigate different research questions that lead to different theoretical solutions (e.g., Dennett, 1987; Marr, 1982 / 2010). However, any complete psychological theory must account for phenomena at multiple levels of analysis, and additionally elucidate the relations *between* levels (e.g., Schall, 2004; Teller, 1984). We also note that the various levels of analysis are causally interrelated and are thus mutually constraining (Rumelhart & McClelland, 1985). It follows that refinement of a model at one level of analysis focuses the search for theoretical solutions at another. We therefore view theoretical pluralism among researchers as an efficient means of developing more complete psychological theories.

We suggest that findings from the "Bayesian Fundamentalist" perspective have highlighted core issues in developing more complete psychological theories, and that discoveries by individual "Bayesian Fundamentalist" researchers may actually facilitate discipline-wide "Enlightenment" by sharpening questions and generating novel insights that stimulate research (e.g., Shiffrin, Lee, Kim, & Wagenmakers, 2008). J&L's admonishment of "Bayesian Fundamentalism", depending on whether it is directed at psychological science as a whole, or to individual researchers, is either a) powerful but directed at a largely non-existent opponent, or b) misguided insofar that the collaborative nature of scientific progress offsets the narrow focus of individual scientists.

Contrary to J&L, we argue the "breadth-first" approach adopted by many Bayesian theorists, rather than stifling theoretical progress, actually facilitates cross-talk *between* levels of analysis. That contemporary Bayesian theorists are aware of, and aspire to resolve this tension, is reflected in recent work that has sought to reconcile rational accounts with more traditional process models. For example, to the extent that models of cognitive processing implement sampling algorithms to approximate full Bayesian inference, models at different levels of analysis can be mutually informative. Shi, Griffiths, Feldman, and Sanborn (2010) illustrate how exemplar models (e.g., Nosofsky, 1986) can be interpreted as an importance sampling algorithm, and similarly, Sanborn, Griffiths, and Navarro (2010) explored the particle filter algorithm as a way of leveraging a process interpretation of Anderson's (1991) rational model. Lewandowsky, Griffiths, and Kalish (2009) used iterated learning (Griffiths & Kalish, 2007; Kalish, Griffiths, & Lewandowsky, 2007), an experimental paradigm motivated by technological advances in sampling techniques used to approximate Bayesian posteriors, to decisively reject a sparse-exemplar model of predicting the future. Kruschke (2006, 2008) contrasted globally and locally Bayesian approaches to associative learning, the latter of which can be construed as carrying very direct process implications concerning selective attention. J&L acknowledge the potential of these approaches for transcending computational level theories but do not acknowledge the role of the computational theories for driving research in this direction.

One area where Bayesian perspectives appear particularly more illuminating than mechanistic approaches is in explaining individual differences. For example, work from within the knowledge partitioning framework has repeatedly found large differences in transfer performance in tasks that can be decomposed into a number of simpler sub-tasks (e.g., Lewandowsky, Kalish, & Ngang, 2002; Lewandowsky, Roberts, & Yang, 2006; Yang & Lewandowsky, 2003). Mechanistic modeling of these results has highlighted the importance of modular architecture (Kalish, Lewandowsky, & Kruschke, 2004; Little & Lewandowsky, 2009), selective attention (Yang & Lewandowsky, 2004), and their interaction (Sewell & Lewandowsky, 2011) in accounting for such individual differences. However, a significant limitation of a mechanistic approach is that the solutions have been built into the models. By contrast, recent Bayesian modeling of knowledge partitioning has showed that many aspects of the individual differences observed empirically emerge naturally if one assumes that people are trying to learn about their environment in a rational manner (Navarro, 2010).

J&L draw uncharitable parallels between "Bayesian Fundamentalism" on the one hand, and behaviorism, connectionism, and evolutionary psychology on the other. In response, we note that theoretical setbacks in those paradigms have clarified our understanding of how the mind does and does not work, emerging with a more refined theoretical toolkit and new, incisive research questions. For behaviorism, a restrictive theoretical stance solidified the need to consider more than just the history of reinforcement in explaining behavior (Neisser, 1967). The inability of the perceptrons to handle nonlinearly separable problems forced connectionists to consider more powerful model architectures (Thomas & McClelland, 2008). Likewise, controversies that have erupted in evolutionary psychology over the propagation of cognitive modules have forced theorists to refine and reevaluate classical notions of modularity (cf. Fodor, 1983, and Barrett & Kurzban, 2006). Thus, the failures of the precedents chosen by J&L actually constitute successes for the field; for example, the cognitive revolution was propelled and accelerated by the spectacular failure of behaviorism.

We close by considering how J&L's critique of "Bayesian Fundamentalism" relates to scientific activity in practice. If they address the scientific community as a whole, their criticism is powerful, but lacks a real target. Alternatively, if J&L's concerns are directed at individual scientists, their plea overlooks the fact that scientific progress, being inherently distributed across multiple research groups, "averages out" individual differences in theoretical dispositions. That is, the aggregate outcomes produced by the scientific community are unlikely to be reflected in the individual outcomes produced by a given scientist (Kuhn, 1970).

Whereas a complete level-spanning theory will always be the goal of science, the approach toward that collective goal will be incremental, and those pursuing it will tend to focus on a particular level of analysis. The important question for any individual researcher is whether an adopted theoretical framework sharpens questions, provides insight and guides new empirical inquiry (Shiffrin et al., 2008); recent Bayesian modeling of cognition undoubtedly fulfills these requirements.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*, 409-429.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*, 628-647.

Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Feldman, J. A. (2010). Cognitive science should be unified: comment on Griffiths et al. and McClelland et al. *Trends in Cognitive Sciences*, *14*, 341.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*, 357-364.

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*, 441-480.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, *14*, 288-294.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review*, *111*, 1072-1099.

Kruschke, J. K. (2006). Locally Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review*, *113*, 677-699.

Kruschke, J. K. (2008). Bayesian approaches to associative learning: From passive to active learning. *Learning & Behavior*, *36*, 210-226.

Kruschke, J. K. (2010). Bridging levels of analysis: comment on McClelland et al. and Griffiths et al. *Trends in Cognitive Sciences*, *14*, 344-345.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago, IL: University of Chicago Press.

Lewandowsky, S., Griffiths, T. L., & Kalish, M. L. (2009). The wisdom of individuals: Exploring people's knowledge about everyday events using iterated learning. *Cognitive Science*, *33*, 969-998.

Lewandowsky, S., Kalish, M., & Ngang, S. K. (2002). Simplified learning in complex situations: Knowledge partitioning in function learning. *Journal of Experimental Psychology: General*, *131*, 163-193.

Lewandowsky, S., Roberts, L., & Yang, L.-X. (2006). Knowledge partitioning in categorization: Boundary conditions. *Memory & Cognition*, *34*, 1676-1688.

Little, D. R., & Lewandowsky, S. (2009). Beyond nonutilization: Irrelevant cues can gate learning in probabilistic categorization. *Journal of Experimental Psychology: Human Perception & Performance*, *35*, 530-550.

Marr, D. (1982 / 2010). *Vision: a computational investigation into the human representation and processing of visual information*. Cambridge: MIT Press.

McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T. T., Seidenberg, M. S., et al. (2010). Letting structure emerge: connectionist and dynamical systems approaches to cognition. *Trends in Cognitive Sciences*, *14*, 348-356.

Navarro, D. J. (2010). Learning the context of a category. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 23, p. 1795-1803).

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39-57.

Rumelhart, D. E., & McClelland, J. L. (1985). Levels indeed! A response to Broadbent. *Journal of Experimental Psychology: General*, *114*, 193-197.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review*, *117*, 1144-1167.

Schall, J. D. (2004). On building a bridge between brain and behavior. *Annual Review of Psychology*, *55*, 23-50.

Sewell, D. K., & Lewandowsky, S. (2011). Restructuring partitioned knowledge: The role of recoordination in category learning. *Cognitive Psychology*, *62*, 81-122.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing Bayesian inference. *Psychonomic Bulletin & Review*, *17*, 443-464.

Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E. J. (2008).

A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, *32*, 1248-1284.

Teller, D. Y. (1984). Linking propositions. *Vision Research*, *10*, 1233-1246.

Thomas, M. S. C., & McClelland, J. L. (2008). Connectionist models of cognition. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (p. 23-58). New York: Cambridge University Press.

Yang, L.-X., & Lewandowsky, S. (2003). Context-gated knowledge partitioning in categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 663-679.

Yang, L.-X., & Lewandowsky, S. (2004). Knowledge partitioning in categorization: Constraints on exemplar models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 1045-1064.