Relevant and robust. A response to Marcus and Davis.

Noah D. Goodman

Department of Psychology, Stanford University

Michael C. Frank

Department of Psychology, Stanford University

Thomas L. Griffiths

Department of Psychology, University of California, Berkeley

Joshua B. Tenenbaum

Department of Brain and Cognitive Sciences, MIT

Peter Battaglia

Department of Brain and Cognitive Sciences, MIT

Jessica Hamrick

Department of Psychology, University of California, Berkeley

Address all correspondence to Noah D. Goodman, Stanford University, Department of Psychology, Jordan Hall, 450 Serra Mall (Bldg. 420), Stanford, CA, 94305 E-mail: ngoodman@stanford.edu

Abstract: Computational models in psychology are precise, fully explicit scientific hypotheses. Over the past 15 years, probabilistic modeling of human cognition has yielded quantitative theories of a wide variety of reasoning and learning phenomena. Recently, Marcus and Davis (2013) critique several examples of this work, using these critiques to question the basic validity of the probabilistic approach. Contra the broad rhetoric of their article, the points made by Marcus and Davis—while useful to consider—do not indicate systematic problems with the probabilistic modeling enterprise.

Computational models in psychology are precise, fully explicit scientific hypotheses. Probabilistic models in particular formalize hypotheses about the beliefs of agents—their knowledge and assumptions about the world—using the structured collection of probabilities referred to as priors, likelihoods, etc. The probability calculus then describes inferences that can be drawn by combining these beliefs with new evidence, without the need to commit to a process-level explanation of how these inferences are performed (Marr, 1982). Over the past 15 years, probabilistic modeling of human cognition has yielded quantitative theories of a wide variety of phenomena (Tenenbaum, Kemp, Griffiths, & Goodman, 2011).

Marcus and Davis (2013, henceforth, M&D) critique several examples of this work, using these critiques to question the basic validity of the probabilistic models approach, based on the existence of alternative models and potentially inconsistent data. Contra the broad rhetoric of their article, the points made by M&D—while useful to consider—do not indicate systematic problems with the probabilistic modeling enterprise.

Several objections stem from a fundamental confusion about the status of optimality in probabilistic modeling, which has been discussed in responses to other critiques (see: Griffiths, Chater, Norris, & Pouget, 2012; Frank, 2013). Briefly: *an* optimal analysis is not *the* optimal analysis for a task or domain. Different probabilistic models instantiate different psychological hypotheses. Optimality provides a bridging assumption between these hypotheses and human behavior; one that can be re-examined or overturned as the data warrant.

Model selection. M&D argue that individual probabilistic models require a host of potentially problematic modeling choices. Indeed, probabilistic models are created via a series of choices concerning priors, likelihoods, response functions, etc. Each of these choices embodies a proposal about cognition, and these proposals will often be wrong. The identification of model assumptions that result in a mismatch to empirical data allows these assumptions to be replaced or refined.

Systematic iteration to achieve a better model is part of the normal progress of science. But if choices are made post-hoc, a model can be *overfit* to the particulars of the empirical data. M&D suggest that certain of our models suffer from this issue. For instance, they show that data on pragmatic inference (Frank & Goodman, 2012) are inconsistent with an alternative variant of the proposed model that uses a hard-max rather than a soft-max function, and ask whether the choice of soft-max was dependent on the data.

The soft-max rule is foundational in economics, decision-theory, and cognitive psychology (Luce, 1959, 1977), and we first selected it for this problem based on a completely independent set of experiments (Frank, Goodman, Lai, & Tenenbaum, 2009). So it's hard to see how a claim of overfitting is warranted here. Modelers must balance unification with exploration of model assumptions across tasks, but this issue is a general one for all computational work, and does not constitute a systematic problem with the probabilistic approach.

Task selection. M&D suggested that probabilistic modelers report results on only the narrow range of tasks on which their models succeed. But their critique focused on a few high-profile, short reports that represented our first attempts to engage with important domains of cognition. Such papers necessarily have less in-depth engagement with empirical data than more extensive and mature work, though they also exemplify the applicability of probabilistic modeling to domains previously viewed as too complex for quantitative approaches.

There is broader empirical adequacy to probabilistic models of cognition than M&D imply. If M&D had surveyed the literature they would have found substantial additional evidence for the models they reviewed—and more has accrued since their critique. For example, M&D critiqued Griffiths and Tenenbaum's (2006) analysis of everyday predictions for failing to provide independent assessments of the contributions of priors and likelihoods, precisely what was done in several later and much longer papers (Griffiths & Tenenbaum, 2011; Lewandowsky, Griffiths, & Kalish, 2009). They similarly critiqued the particular tasks selected by Battaglia, Hamrick, and Tenenbaum (2013) without discussing the growing literature testing similar "noisy Newtonian" models on other phenomena (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2012; Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014; Sanborn, Mansinghka, & Griffiths, 2013; Smith, Dechter, Tenenbaum, & Vul, 2013; Téglás et al., 2011). Smith, Battaglia, and Vul (2013) even directly address exactly the challenge M&D posed regarding classic findings of errors in physical intuitions. In other domains, such as concept learning and inductive inference, where there is an extensive experimental tradition, probabilistic models have engaged with diverse empirical data collected by multiple labs over many years (e.g. Goodman, Tenenbaum, Feldman, & Griffiths, 2008; Kemp & Tenenbaum, 2009).

M&D also insinuate empirical problems that they do not test. For instance, in criticizing the choice of dependent measure used by Frank and Goodman (2012), they posit that a forced-choice task would yield a qualitatively different pattern (discrete rather than graded responding). In fact, a forced-choice version of the task produces graded patterns of responding across a wide variety of conditions (Stiller, Goodman, & Frank, 2011, 2014; Vogel, Emilsson, Frank, Jurafsky, & Potts, 2014).

Conclusions. We agree with M&D that there are real and important challenges for probabilistic models of cognition, as there will be for any approach to modeling a system as complex as the human mind. To us, the most pressing challenges include understanding the relationship to lower levels of psychological analysis and neural implementation, integrating additional formal tools, clarifying the philosophical status of the models, extending to new domains of cognition, and, yes: engaging with additional empirical data in the current domains while unifying specific model choices into broader principles. As M&D state, "ultimately, the Bayesian approach should be seen as a useful tool"—one that we believe has already proven its robustness and relevance by allowing us to form and test quantitatively accurate psychological hypotheses.

References

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences, 110*, 18327–18332.

Box, G. E., & Draper, N. R. (2007). *Response surfaces, mixtures, and ridge analyses.* New York, NY: John Wiley & Sons.

Frank, M. C. (2013). Throwing out the Bayesian baby with the optimal bathwater: Response to Endress (2013). *Cognition, 128*, 417–423.

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science, 336*, 998–998.

Frank, M. C., Goodman, N. D., Lai, P., & Tenenbaum, J. B. (2009). Informative communication in word production and word learning. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 206–211). Austin, TX: Cognitive Science Society.

Gerstenberg, T., Goodman, N., Lagnado, D. A., & Tenenbaum, J. B. (2012). Noisy newtons: Unifying process and dependency accounts of causal attribution. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Gerstenberg, T., & Goodman, N. D. (2012). Ping pong in church: Productive use of concepts in human probabilistic inference. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2014). From counterfactual simulation to causal judgment. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive Science, 32*, 108–15.

Griffiths, T. L., Chater, N., Norris, D., & Pouget, A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). *Psychological Bulletin, 138*, 415–422.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science, 17*, 767–773.

Griffiths, T. L., & Tenenbaum, J. B. (2011). Predicting the future as Bayesian inference: People combine prior knowledge with observations when estimating duration and extent. *Journal of Experimental Psychology: General, 140*, 725.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review, 116*, 20.

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.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology, 15*, 215–233.

Marcus, G. F., & Davis, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychological Science, 24*, 2351–2360.

Marr, D. (1982). *Vision*. Cambridge, MA: MIT Press.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological Review, 120*, 411.

Smith, K., Battaglia, P., & Vul, E. (2013). Consistent physics underlying ballistic motion prediction. In *Proceedings of the 35th Conference of the Cognitive Science Society* (pp. 3426–3431). Austin, TX: Cognitive Science Society.

Smith, K., Dechter, E., Tenenbaum, J., & Vul, E. (2013). Physical predictions over time. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Stiller, A., Goodman, N. D., & Frank, M. C. (2011). Ad-hoc scalar implicature in adults and children. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Stiller, A., Goodman, N. D., & Frank, M. C. (2014). Ad-hoc implicature in preschool children. Language Learning and Development .

Téglás, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure reasoning in 12-month-old infants as probabilistic inference. *Science, 332*, 1054–1059.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. Science , 331 (6022), 1279–1285.

Vogel, A., Emilsson, A. G., Frank, M. C., Jurafsky, D., & Potts, C. (2014). Learning to reason pragmatically with cognitive limitations. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.