Check for updates

Commentary

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2015, Vol. 26(4) 542–544
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614568433
pss.sagepub.com
SAGE

Still Searching for Principles: A Response to Goodman et al. (2015)

Gary F. Marcus and Ernest Davis
New York University

Received 12/10/14; Revision accepted 12/24/14

We appreciate the efforts of Goodman et al. (2015) to address concerns we raised in our prior critique (Marcus & Davis, 2013) of Bayesian accounts of human cognition but remain unconvinced.

Multiple Models

In our original article, we argued that multiple, equally plausible, Bayesian models could be constructed for the tasks under consideration, and that Bayesian theories do not constrain which model applies in any given case. Without a prior theory on how to choose proper models, we suggested, the Bayesian approach to cognition risks becoming an exercise in post hoc modeling. Additionally, we pointed out that the word *optimality* is used with many different meetings (as we discuss later in this Commentary).

The best rejoinder by Goodman et al. is their correct assertion that their use of the soft-max rule in Frank and Goodman (2012) was not arbitrary, as we claimed; rather, they have used it consistently across a number of publications in which they have developed their rational-speech-act (RSA) theory of communication.

But our larger point stands: If we look beyond RSA, Bayesian models in other domains use various choice rules, and no general rule for their selection has been proposed. Battaglia, Hamrick, and Tenenbaum (2013), Kemp and Tenenbaum (2008), and Cain, Vul, Clark, and Mitroff (2012) used hard max; Gweon, Tenenbaum, and Schulz (2010) used probability matching; Griffiths and Tenenbaum (2006, 2011) used the median. Smith and Vul (2013) used a hard-max rule for a task involving predicting the movement of a bouncing ball; for a very similar task, Smith, Dechter, Tenenbaum, and Vul (2013) use two separate soft-max rules, with four parameters tuned to fit the data. Even in the rejoinder by Goodman et al., we see no principle for deciding which rule applies in any given situation.

Moreover, though the choice rule in RSA is fixed, other aspects of the model remain fluid or arbitrary. Frank and Goodman (2012) assumed without justification, for example, that the hearer knows what word choices are available to the speaker; this hardly seems plausible, yet the problem persists in the work by Kao, Wu, Bergen, and Goodman (2014).

These same problems beset other work that we did not include in our original article. For example, Gweon et al. (2010) argued that their squeaky-toy experiment showed that infants compute a posterior probability on hypotheses. Their model posited that the babies chose among four different hypotheses. But there was no principled justification for that particular model. In a more recent analysis (Davis & Marcus, 2014), we found that there are 43 different hypotheses—all about equally plausible a priori—that the babies might have considered, and that there are more than 7,500 different Bayesian models—all equally well motivated—that a theorist might use with these data.

Rips, Asmuth, and Bloomfield (2013) pointed out the same flaw in the Bayesian theory of number learning proposed by Piantadosi, Tenenbaum, and Goodman (2012). The model relies on having a limited vocabulary of primitive concepts under consideration, and it is not explained how the child learner would select the appropriate vocabulary.

Optimality

In our original critique, we noted that strong, unwarranted claims for the optimality of performance are often made in the literature on Bayesian models, and that the notion of optimality varies among reports, with no systematic criterion being offered. Goodman et al. claim that we are suffering from a "fundamental confusion" (p. 539) about what is meant by *optimality*, but the truth is that

Corresponding Author:
Gary F. Marcus, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003
E-mail: gary.marcus@nyu.edu

the confusion stems from the literature that we critiqued, which uses the term too often, in too many different ways—sometimes as a weak guideline, but often as a strong psychological claim, just as we suggested.

Strong claims

The word *optimal* or *optimize* is used, for example, in the titles of articles by Cain et al. (2012); Griffiths and Tenenbaum (2006); Kording, Tenenbaum, and Shadmehr (2007); and Piantadosi, Tiley, and Gibson (2011). Claims of optimality or near optimality are made in Kao et al. (2014), Téglás et al. (2011), and many more reports. Oaksford and Chater (2009) argued that "behavioral predictions [should be] derived from the assumption that the cognitive system is solving this problem, optimally (or, more plausibly, approximately), under . . . constraints" (p. 72). Griffiths and Tenenbaum (2006) stated that "everyday cognitive judgments follow . . . optimal statistical principles" and that there is "close correspondence between people's implicit probabilistic models and the statistics of the world" (p. 767). Sanborn, Mansinghka, and Griffiths (2013) proposed that "people's judgments [about physical events] are based on optimal statistical inference over a Newtonian physical model that incorporates sensory noise and intrinsic uncertainty about the physical properties of the objects being viewed" (p. 411). Frank's (2013) more moderate view is the exception rather than the rule.

Arbitrary criteria

In our original article, we demonstrated that the optimality claims in two of the experiments reported in Griffiths and Tenenbaum (2006) depended on arbitrary assumptions about what information in the problems the subjects were considering and what information they were ignoring. Similarly, the justifications of Oaksford and Chater (2009) and of Tenenbaum and Griffiths (2001) for viewing subjects' nonnormative answers to questions as in fact optimal depends on arbitrary assumptions about how the subjects' interpretations differed from the experimenters'. In the title of the article by Piantadosi et al. (2011), "Word Lengths Are Optimized for Efficient Communication," the word *optimized* means little more than "pretty good." Each of the varying choice rules we referred to earlier is "optimal" in a different sense.

If *optimal* means something different in each report, the overall claim that cognitive processes in general are optimal becomes nearly meaningless. The cleverly worded reply in Goodman et al.—that "*an* optimal analysis is not *the* optimal analysis" (p. 539)—merely sidesteps the problem.

Literature

What is left? Goodman et al. note that we did not cite, well, everything, including a variety of reports that had not come out before our critique went to press. That is true enough, but we were hardly lax. We cited 10 reports by the authors of Goodman et al. (compared with just 1 by ourselves). More important, the additional reports that Goodman et al. mention hardly refute our argument. Goodman et al. also chide us for focusing on work that was not "mature" (p. 540), but in fact we focused primarily on articles in prestigious outlets like *Science* (Frank & Goodman, 2012), *Proceedings of the National Academy of Sciences* (Battaglia et al., 2013), and *Psychological Science* (Griffiths & Tenenbaum, 2006). Given that both Griffiths and Tenenbaum list the latter highly cited article as among their key publications, it hardly seems unfair to focus attention on it.

Conclusion

What is most telling, however, is what is absent. In our original piece, we concluded that in many cases, there are many possible Bayesian models that could be used to characterize high-level cognition, that there are many possible standards of optimality, and that Bayesian theory offers no principled way to choose among them. We see nothing in the response of Goodman et al. that alleviates those concerns.

Author Contributions

This Commentary was jointly written by both authors.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

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, USA*, *110*, 18327–18332.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A Bayesian optimal foraging model of human visual search. *Psychological Science*, *23*, 1047–1054.

Davis, E., & Marcus, G. F. (2014). *The hypothesis space in Gweon, Tenenbaum, and Schulz (2010)*. Retrieved from http://www.cs.nyu.edu/faculty/davise/papers/GTS.html

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

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

Goodman, N. D., Frank, M. C., Griffiths, T. L., Tenenbaum, J. B., Battaglia, P. W., & Hamrick, J. B. (2015). Relevant and robust: A response to Marcus and Davis (2013). *Psychological Science, 26*, 539–541.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal prediction 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 observation when estimating duration and extent. *Journal of Experimental Psychology: General*, *140*, 725–743.

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, USA*, *107*, 9066–9071.

Kao, J., Wu, J., Bergen, N., & Goodman, N. (2014). Nonliteral understanding of number words. *Proceedings of the National Academy of Sciences, USA*, *111*, 12002–12007.

Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences, USA*, *105*, 10687–10692.

Kording, K., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory are a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, *10*, 779–786.

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

Oaksford, M., & Chater, N. (2009). Précis of Bayesian rationality: The probabilistic approach to human reasoning [Target article and commentaries]. *Behavioral & Brain Sciences*, *32*, 69–120.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2012). Bootstrapping in a language of thought: A formal model of numerical concept learning. *Cognition*, *123*, 197–217.

Piantadosi, S. T., Tiley, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences, USA*, *108*, 3526–3529.

Rips, L. J., Asmuth, J., & Bloomfield, A. (2013). Can statistical learning bootstrap the integers? *Cognition*, *128*, 320–330.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review*, *120*, 411–437.

Smith, K. A., Dechter, E., Tenenbaum, J. B., & Vul, E. (2013). Physical predictions over time. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1342–1347). Austin, TX: Cognitive Science Society.

Smith, K. A., & Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in Cognitive Science*, *5*, 185–199.

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., & Griffiths, T. L. (2001). The rational basis of representativeness. In J. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society* (pp. 1036–1041). Mahwah, NJ: Erlbaum.