

HOW IS HAIR GEL QUANTIFIED?

MARK A. PITT

*Department of Psychology, Ohio State University, 1835 Neil Avenue, Columbus, Ohio,
43210, USA*

JAY I. MYUNG

*Department of Psychology, Ohio State University, 1835 Neil Avenue, Columbus, Ohio,
43210, USA*

Shillcock et al suggest that the preoccupation with simplicity in cognitive modeling has been detrimental to the discipline. They propose instead an approach in which the complexity of the real world should be the objective of modelers. We discuss some of the difficulties in achieving this goal from the standpoint of quantitative methods of model selection.

1. Introduction

In the target article [1], Shillcock, Roberts, Kreiner and Obregon (SRKO), review some of the philosophical assumptions in the contemporary computational modeling of cognition, and are troubled by what they perceive as an over-emphasis on simplicity in modeling, which they believe has been detrimental to the advancement of cognitive science.

According to SRKO, the notion of simplification has become too much of a guiding principle in current cognitive modeling, to the point that simplification (obeying Occam's Razor) is routinely taken as a goal of modeling in itself. They opine that this approach is myopic, incomplete, and potentially misleading. Instead, SRKO contend that simplification should be done with much greater care to ensure that the model, even after simplifying abstractions and idealizations, still contains the most basic yet essential kernel of truth. In the end, SRKO adopt the position that an alternative, more productive path would be to model "the fullest possible representation of the real-world complexity of the modeling domain..." SRKO clarify their approach and provide guidance for potential adopters in the form of eight questions for modelers to consider.

Although it is difficult to disagree with the gist of their argument, we believe that the proposal to "reproduce the full, real-world complexity of the

domain under study” may be unrealistic and unrealizable in practice because of the many conceptual and implementational challenges it poses.

2. Models Are Just Tools

We preface our main point with one that sometimes gets forgotten in discussions of cognitive modeling. Models are quantitative stand-ins of psychological theories that are developed from data acquired through experimentation. The idealized, naive goal of modeling is to identify the underlying regularities (truth) from which the data are actually generated. This goal, however, is not achievable, for at least two reasons. First, there are never enough observations (i.e., data) to pin down the truth exactly. Second, the truth may be quite complex, beyond the modeler’s imagination, and thus is likely to be different from any one of the candidate models the modeler may contemplate, as captured by the famous quote “All models are wrong but some are useful” credited to George E. P. Box (1975).

A utilitarian paradigm of cognitive modeling that echoes George Box is to view models as no more than tools with which to study the brain and behavior. Viewed from this perspective, models are used primarily to increase the precision of theoretical predictions, generate novel and experimentally testable hypotheses, provide insights into complex behavior, etc. Samuel Karlin nicely sums up this aspect of modeling when he said, “The purpose of modeling is not to fit the data but to sharpen the questions” [2].

From this perspective, a realistic goal of cognitive modeling is to identify the one model, among a set of candidate models, that represents the closest possible approximation to the cognitive process of interest so that the identified model would capture the real-world complexity of the process under study, not necessarily fully but in some meaningful ways. Although we suspect SRKO would not disagree with this rather uncontroversial statement, the devil is in the implementational details. For example, a key issue is that of how one should select one model, from a set of competing models, that best approximates the cognitive process. This is the topic of model evaluation and comparison. In the remainder of this commentary, we discuss a few of the challenges that will arise when Occam’s Razor must be balanced with other criteria, such as considering real-world complexity.

3. Model Evaluation and Comparison: Where Hair Gel Becomes Sticky

The ever-increasing popularity of modeling in cognitive science has resulted in the introduction of a great number of computational models within and across content areas. Although their purposes vary, from being proof-of-concept demonstrations to modeling a complex cognitive process such as reading, what has lagged considerably are methods for evaluating the quality of the models and comparing between competing models. The nuts and bolts of model evaluation are a critical component in model development, as they assist in justifying a range of choices, whether it be in the design of a model or the selection of one model over another.

Although standard practice is to compare empirical data with model output, this sufficiency test is minimally informative and it would be prudent to develop model analysis tools as well in tandem. With a set of tools that provide keener insight into model performance (answering *how* and *why* a model performs the way it does), the current modeling practice would likely be more productive and suffer less from some of the problems that SRKO rightly highlight, such as a short lifespan.

Cognitive science is a particularly challenging discipline in which to develop formal model comparison methods because of the diversity of the types of models – How does one compare a simulation-based neural network model with an algebraically formulated model? The task is not impossible, and we have suggested methods for doing so [3]. The challenge is to envision its implementation in SRKO's paradigm, where model breadth must also be considered.

In the context of model comparison, to abide by Occam's Razor is not the same as adhering to simplicity. Rather, it is intended as a check on excess, unnecessary model complexity (flexibility in performance). The principle is instantiated in quantitative model selection methods as a means of justifying the additional complexity held by one model over its competitors. A complex model (e.g., one with more parameters, more hidden units, or a more complex functional form) will be preferred, but only as long as it is justified by the complexity of the cognitive process under study. Specifically, one uses Occam's Razor to identify the model that is sufficiently complex to capture the regularities underlying the data but not too complex to enable it can capitalize on ever-present random noise in the data [4]. By relaxing this criterion, an overly complex model could be favored, one that might fit virtually any data

pattern (both regularity and noise), but reveals little of psychological relevance. This would surely drive cognitive science to an impasse.

Looked at in another way, Occam's Razor emphasizes a very strong bottom-up (i.e., data-driven) orientation. This ensures the model maintains maximum contact with what is known about the cognitive process under study. Although SRKO acknowledge a need for parsimony, it would have to be loosened in an effort to accommodate a wider range of phenomena.

Such a rebalancing of parsimony with a model's explanatory breadth would result in a few problems. One is that the source of superior model performance, whether measured as a more exact simulation or a better fit to the data, is difficult to isolate in the model. With a highly complex model, good performance could be attributable to the accuracy of the model (i.e., it is a good approximation of the underlying cognitive process) or to its excess complexity (i.e., its ability to reproduce a wide range of data patterns). Only by some means of simplification, such as turning off unnecessary parts of the model, could this question begin to be answered.

A related problem emerges when models become highly detailed is behavioral tractability. What parts of the model are responsible for its many behaviors? SRKO touch on this point briefly when criticizing the current practice of cognitive modeling, but do not explain how it can be overcome in their approach. Without intimate knowledge of model behavior, it can be very difficult to improve model performance, largely because the sources of performance gains are not transparent or non-unique, with one part of the model able to compensate for deficiencies in other parts. As one can imagine, this type of flexibility stymies modeling.

SRKO cite the abundance of data in the literature as evidence that the current modeling paradigm is not working, and what is needed is a more comprehensive, integrated approach. From the vantage point of model evaluation, the problem is just the opposite: Data that can convincingly discriminate models are in short supply. Models under consideration can often mimic each other so closely that they differ in primarily subtle quantitative predictions. What is more, it can be very difficult to determine how two models actually differ, even when comparing models that generate simple data patterns. It is not surprising, then, that it can be exceedingly difficult to design experiments that stand a good chance of discriminating between models. It is unlikely that this situation will change by trying to model the "whole." The increased complexity that accompanies more detailed models may well result in greater mimicry, not less.

4. Conclusion

The desire to model real-world complexity is probably held by most modelers. This commentary is intended to point out what we believe is a formidable challenge in practicing this type of modeling, how to balance model parsimony with model breadth in model selection. If Occam's Razor "only tells us half the story," how is the other half quantified in a mathematically rigorous way so as to facilitate model selection? It is hard to imagine hair gel ever going out of fashion, but style-conscious modelers have a formidable task ahead of them.

References

1. Shillcock, R., Roberts, M., Kreiner, H., and Obregon, M. Some issues in computational modeling; Occam's Razor and Hegel's hair gel. In E. J. Davelaar (ed.), *Proceedings of the Twelfth Neural Computation and Psychology Workshop* (London, UK: World Scientific), xxx-xxx (2011).
2. Karlin, S. The 11th R. A. Fisher Memorial Lecture. *The Royal Society 20 Meeting* (1983).
3. Myung, J. I., Tang, Y., and Pitt, M. A. Evaluation and comparison of computational models. *Methods in Enzymology*, **454**, 287-304 (2009).
4. Pitt, M. A., Myung, I. J., and Zhang, S. Toward a method of selecting among computational models of cognition. *Psych. Rev.*, **109**, 472-491 (2002).
5. McClelland, J. L., and Elman, J. L. The TRACE model of speech perception. *Cog. Psych.*, **18**, 1-86 (1986).