Title: Blind ignorance or rational inattention? On the history and future of metabolic considerations for cognitive models

Abstract: We build upon Haueis & Colaço's account in two ways. First, we argue that the backgrounding of metabolism is a pragmatic response to a *data granularity problem* faced by traditional cognitive models. Then, we suggest that *equivalence demonstrations* offer a promising pathway for incrementally introducing metabolic considerations into extant modeling practice.

Commentary: Haueis & Colaço's (2026; H&C) account asserts a *normative* role for metabolism in cognitive modeling: in general, models that incorporate metabolic considerations ought to be preferred over those that don't. This raises an important question left unaddressed by the authors: *why* hasn't metabolism played a normative role in cognitive modeling? Our comment offers an answer—the *data granularity problem*—which subsequently informs our suggestion to use *equivalence demonstrations* as a method for integrating metabolic considerations into existing cognitive modeling research.

Cognitive modeling has its roots in mathematical psychology, which emerged as a subdiscipline of psychology in the 1950s. The goal of this field was to build formally-specified theories describing the cognitive constructs and processes that generate human behavior in experimental and real-world settings (Batchelder, 2010). These theories are tested by application to behavioral data. Modelers first define a set of mathematical objects and relationships that are intended to represent cognitive constructs and processes (i.e., they specify a model *structure*). Then, they identify numerical values for model variables that minimize the difference between the model's predictions and measurements of human behavior (i.e., they *fit* free parameters of the model to data). Inferences about *which* constructs and processes best explain behavior are then typically conducted by comparing model performance metrics across competing models (i.e., model comparison) and/or performing statistical tests on fitted model parameters.

Contrasting this description of a conventional cognitive model with H&C's definition of "idealized, abstract systems whose elements stand in a relation to assumed elements of the target capacity" (p. 5) reveals an important conceptual gap. Namely, H&C have in mind the broad space of formal structures that *could* function as models of cognitive processes, rather than the narrower space of formal structures that have *actually functioned* as "cognitive models" for decades. Understanding the motivations behind the historical backgrounding of metabolism is essential for identifying concrete, tractable, and actionable proposals for future research.

We suggest that the "backgrounding" of metabolic considerations emerged as a pragmatic response to empirical constraints on the model-based inference process. For a parameter value to be meaningfully estimated via behavior, it has to stand in for a latent component that plausibly (i) changes as a function of task conditions, (ii) differs across individuals performing the same task, or (iii) is some combination of (i) and (ii). Adding terms that do not contribute to structural variability in behavioral data risks *overfitting* the parameter values to idiosyncrasies of the sample, resulting in poor out-of-sample generalization. Further, as model complexity increases, so too does the probability of two competing models *mimicking* each other: making identical predictions about behavior on the basis of qualitatively distinct model structures (Khodadadi & Townsend, 2015; Zhang et al., 2014). To defend against these perils, modelers relied heavily on the *principle of parsimony* to guide model generation and evaluation: all else being equal, simpler models should be preferred (Vandekerckhove et al., 2015).

These historical considerations thus reveal that a *data granularity problem* figured centrally in the backgrounding of metabolism: the parameters of cognitive models are estimated by application to behavior, but behavioral data are often too coarse-grained to meaningfully constrain model parameters related to metabolism. One solution to this problem can be incorporating biological measurements into the data that cognitive models are tasked with fitting, such that parameter values must be estimated jointly using behavior and brain data (Turner et al., 2019). These approaches have had some success in adjudicating among mimicking models (Kelly et al., 2021), but (1) it is not practical to require that all cognitive model parameters be estimated jointly to neural and behavioral data and (2) "joint" models face their own challenges related to generalizability, interpretability, etc. (Churchland & Kiani, 2016).

As an alternative, we propose that *equivalence demonstrations*—formal derivations or numerical simulations identifying equivalence between two differently-specified models—are a practical and powerful way to incrementally incorporate metabolism into cognitive modeling. This approach is exemplified in a recent study by Wang and Donkin (2024), who demonstrated that relaxing an empirically-supported assumption in a prominent cognitive model increases its similarity to a more complex model that asserts a specific biological mechanism for the data-generating process. While this modification does not change the original model's ability to fit behavior, it increases the validity of interpreting its fits in light of biological processes. The authors essentially *leveraged* model mimicry—or functional equivalence—to their advantage by identifying simple models that mimic more complex, biologically-inspired alternatives.

In this vein, we suggest that the field engage in *re-expressing* existing models in information-theoretic terms. Identifying information-theoretic "conversions" of prominent cognitive models would allow for investigating the metabolic costs of different operations within those models, making the *models themselves* "legible" with respect to metabolic facts rather than making *metabolic knowledge* "legible" or compatible with a given model's formulation (as suggested by H&C in §2.1.5). Our suggestion exemplifies the *generative* and *evaluative* functions of metabolism proposed by H&C: it "populates" the space with new functional forms, which themselves serve to "partition" the space of *existing* models with respect to their metabolic feasibility. Crucially, because these information-theoretic models are *not* tasked with fitting behavior directly, they need not abide by the strict parsimony constraints shaping traditional cognitive models. Instead, modelers can focus their attention on exploring the vast space of functional forms relating metabolic considerations to cognitive functions using the "common currency" of information theory.

In sum, the backgrounding of metabolism in cognitive modeling stems not from blind ignorance or unwarranted dismissal of metabolic considerations, but principled constraints on the types of inference warranted by the historical *explananda* of cognitive models: behavioral data. The norms guiding cognitive modeling structurally disincentivize directly representing metabolic processes, which poses a challenge for implementing H&C's prescription. Building new classes of *equivalent* models that represent relevant biological information at varying levels of detail both offers relief from these structural constraints and directly facilitates generative and evaluative functions of metabolic considerations.

References

Batchelder, W. H. (2010). Mathematical psychology. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(5), 759–765.

Churchland, A. K., & Kiani, R. (2016). Three challenges for connecting model to mechanism in decision-making. *Current Opinion in Behavioral Sciences*, *11*, 74–80.

Kelly, S. P., Corbett, E. A., & O'Connell, R. G. (2021). Neurocomputational mechanisms of prior-informed perceptual decision-making in humans. *Nature Human Behaviour*, *5*(4), 467–481.

Khodadadi, A., & Townsend, J. T. (2015). On mimicry among sequential sampling models. *Journal of Mathematical Psychology*, *68-69*, 37–48.

Turner, B. M., Forstmann, B. U., & Steyvers, M. (2019). *Joint Models of Neural and Behavioral Data* (1st ed.) [PDF]. Springer Nature.

Vandekerckhove, J., Matzke, D., & Wagenmakers, E. (2015). Model Comparison and the Principle of parsimony. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology* (pp. 300–319). Oxford University Press.

Wang, J.-S., & Donkin, C. (2024). The neural implausibility of the diffusion decision model doesn't matter for cognitive psychometrics, but the Ornstein-Uhlenbeck model is better. *Psychonomic Bulletin & Review*, 1–13.

Zhang, S., Lee, M. D., Vandekerckhove, J., Maris, G., & Wagenmakers, E.-J. (2014). Time-varying boundaries for diffusion models of decision making and response time. *Frontiers in Psychology*, *5*, 1364.