test items were to the items they saw at training as well as to the test item's similarity to agent prototypes (Dowty, 1991). It seems that not all agents are equally good agents, a surprising result if there is true role-filler independence.

The authors correctly point out that connectionist models "simulate compositionality, but fail to preserve identity of the original representational elements" (target article, sect. 2, para. 7). The authors do not consider the possibility that human compositionality may be simulated as well (Dekker, Otto, & Summerfield, 2022; Lahav, 1989).

Lastly, *logical operators* such as AND, IF, and OR are a "hallmark of LoT architectures" (target article, sect. 2, para. 10). Yet children under the age of about five have a notoriously difficult time learning categories based on even the simplest logical rules (Rabi, Miles, & Minda, 2015; Rabi & Minda, 2014). Adults are better (and certainly better than other animals!), but arguably rule-based reasoning is far more difficult than it should be if such logical operators actually underlie much of our perception and reasoning (Goldwater, Don, Krusche, & Livesey, 2018; Lupyan, 2013; Mercier & Sperber, 2017).

It is true that at least for stimuli composed of easy-to-verbalize and recombine features such as circles and triangles of various colors used by Piantadosi, Tenenbaum, and Goodman (2016) adults can do well, showing patterns of behavior well-explained by the use of logical operators. However, such behavior is fragile in ways unexpected if these operators underlie our everyday cognition. Formally simple operations like XOR are notoriously difficult for people (Shepard, Hovland, & Jenkins, 1961). Even on simple rules like IF A, performance strongly depends on factors like verbal nameability of the constituents (Zettersten & Lupyan, 2020).

Ironically, Piantadosi, cited in support of hard-coded logical connectives (Piantadosi et al., 2016) was explicit that their data concern adults ("our results are not about children," p. 22) making the claim that logical operators underlie our core cognitive processes suspect. He later went on to argue that "primitives" like AND and OR need not in fact be primitives and can be learned (Piantadosi, 2021). I would add that such learning may be supported in part by natural language (Lupyan & Bergen, 2016).

To be fair, not all the evidence the authors use in support of the LoTH is linguistic. A considerable weight is placed on the construct of object files that are somehow meant to explain perception in terms of LoTH. Although object files may be a useful construct for understanding certain perceptual generalizations, there is good reason why research in perception treats visual representations as analog/iconic representations (Block, forthcoming).

In a town inhabited by highly educated people with a Western philosophical bent, LoTH is a sensible starting point in thinking about how cognition works. In towns inhabited by the rest of us, it is a curious game that some learn to play. The most fun games are often those that transport us to imagined worlds. The world of the LoT hypothesis is likely one of these.

**Financial support.** This study was supported by NSF-PAC 2020969.

**Competing interest.** None.

## References

Block, N. (forthcoming). Let's get rid of the concept of an object file. In B. McLaughlin & J. Cohen (Eds.), *Contemporary debates in philosophy of mind* (2nd ed., pp. 494–516). Wiley. https://philarchive.org/rec/BLOLGR

Dekker, R. B., Otto, F., & Summerfield, C. (2022). Curriculum learning for human compositional generalization. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(41), e2205582119. https://doi.org/10.1073/pnas.2205582119

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*(3), 547–619. https://doi.org/10.2307/415037

Goldwater, M. B., Don, H. J., Krusche, M. J. F., & Livesey, E. J. (2018). Relational discovery in category learning. *Journal of Experimental Psychology. General*, *147*(1), 1–35. https://doi.org/10.1037/xge0000387

Lahav, R. (1989). Against compositionality: The case of adjectives. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *57*(3), 261–279.

Levinson, S. C. (1997). From outer to inner space: Linguistic categories and non-linguistic thinking. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 13–45). Cambridge University Press.

Lupyan, G. (2013). The difficulties of executing simple algorithms: Why brains make mistakes computers don't. *Cognition*, *129*(3), 615–636. https://doi.org/10.1016/j.cognition.2013.08.015

Lupyan, G. (2015). The paradox of the universal triangle: Concepts, language, and prototypes. *Quarterly Journal of Experimental Psychology*, *70*(3), 389–412. https://doi.org/10.1080/17470218.2015.1130730

Lupyan, G. (2016). The centrality of language in human cognition. *Language Learning*, *66* (3), 516–553. https://doi.org/10.1111/lang.12155

Lupyan, G., & Bergen, B. (2016). How language programs the mind. *Topics in Cognitive Science*, *8*(2), 408–424. https://doi.org/10.1111/tops.12155

Lupyan, G., & Zettersten, M. (2021). Does vocabulary help structure the mind?. In M. D. Sera & M. A. Koenig (Eds.), *Minnesota symposia on child psychology* (pp. 160–199). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119684527.ch6

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2023). *Dissociating language and thought in large language models: A cognitive perspective*. arXiv: 2301.06627. https://doi.org/10.48550/arXiv.2301.06627

Malt, B. C., Gennari, S., Imai, M., Ameel, E., Saji, N., & Majid, A. (2015). Where are the concepts? What words can and can't reveal. In E. Margolis & S. Laurence (Eds.), *Concepts: New directions* (pp. 291–326). MIT Press.

Malt, B. C., & Majid, A. (2013). How thought is mapped into words. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(6), 583–597. https://doi.org/10.1002/wcs.1251

Matute, E., Montiel, T., Pinto, N., Rosselli, M., Ardila, A., & Zarabozo, D. (2012). Comparing cognitive performance in illiterate and literate children. *International Review of Education*, *58*(1), 109–127. https://doi.org/10.1007/s11159-012-9273-9

Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Olson, D. R. (2002). What writing does to the mind. In E. Amsel & J. P. Byrnes (Eds.), *Language, literacy, and cognitive development* (pp. 153–165). Erlbaum.

Piantadosi, S. T. (2021). The computational origin of representation. *Minds and Machines*, *31*(1), 1–58.

Piantadosi, S. T., Tenenbaum, J., & Goodman, N. (2016). The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological Review*, *123*(4), 392–424.

Rabi, R., Miles, S. J., & Minda, J. P. (2015). Learning categories via rules and similarity: Comparing adults and children. *Journal of Experimental Child Psychology*, *131*, 149–169. https://doi.org/10.1016/j.jecp.2014.10.007

Rabi, R., & Minda, J. P. (2014). Rule-based category learning in children: The role of age and executive functioning. *PLoS ONE*, *9*(1), e85316. https://doi.org/10.1371/journal.pone.0085316

Rissman, L., & Lupyan, G. (2022). A dissociation between conceptual prominence and explicit category learning: Evidence from agent and patient event roles. *Journal of Experimental Psychology. General*, *151*(7), 1707–1732. https://doi.org/10.1037/xge0001146

Rissman, L., & Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychonomic Bulletin & Review*, *26*(6), 1850–1869. https://doi.org/10.3758/s13423-019-01634-5

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42. https://doi.org/10.1037/h0093825

Zettersten, M., & Lupyan, G. (2020). Finding categories through words: More nameable features improve category learning. *Cognition*, *196*, 104135. https://doi.org/10.1016/j.cognition.2019.104135

# Stop me if you've heard this one before: The Chomskyan hammer and the Skinnerian nail

Alex Madva

California State Polytechnic University, Pomona, CA, USA
alexmadva@gmail.com
https://alexmadva.com/

doi:10.1017/S0140525X23001851, e282

**Abstract**

The target article signal boosts important ongoing work across the cognitive sciences. However, its theoretical claims, generative value, and purported contributions are – where not simply restatements of arguments extensively explored elsewhere – imprecise, noncommittal, and underdeveloped to a degree that makes them difficult to evaluate. The article's apparent force results from engaging with straw rather than steel opponents.

> Batman: Then why do you want to kill me?
> Joker: Kill you? I don't want to kill you. What would I do without you? Go back to ripping off mob dealers? No, no, no. No, you… you complete me. (Nolan, 2008)

For many a hammer, everything is a nail. For many a philosopher of mind, everything is a chance to rehearse Kant's criticism of Hume, Chomsky's criticism of Skinner, and Fodor's criticism of every empiricist, holist, or, in the pages of *BBS* (1985), relativist who rubbed him the wrong way. Thus the target article repeats, again and again across different domains, a well-worn argumentative maneuver in psychology and philosophy: "this mental phenomenon you're trying to explain in terms of a simpler process – be it associations, model-free learning, neural nets, or icons – must instead be explained by a more complex process, which performs language-like computations."

The theoretical payoff of this selective tour through case studies is remarkably modest. For Quilty-Dunn et al. do not deny that the simpler, nonlinguistic processes exist and have real effects. They deny that the nonlinguistic processes explain *everything*. Repurposing the old joke, we've established what kind of theorist you are – a pluralist – and now we're just haggling over the details. The haggling, in this case, recalls trench warfare. On some fronts (like artificial intelligence), neural nets make stunning advances, even as the language-of-thought hypothesis (LoTH) plants its flag on other patches of cognitive terrain hitherto claimed by nonlinguistic theories. The broader import of this unsystematic assemblage of localized skirmishes is unclear. Is LoTH "the best game in town," or one game among others, which the mind perhaps plays somewhat more often than some think?

The authors deny that so-called "system 1" (target article, sect. 6, note 11) is *purely* associative. It's true that associative interpretations of implicit bias continue to hold sway in pop-psych discourse, but hardly anyone paying attention to what the authors rightly call a "near-deluge" (target article, sect. 6.2, para. 5) of research on propositional effects in implicit social cognition continues to defend the extreme associationist views targeted by the authors. Many of us never did (Brownstein & Madva, 2012; Del Pinal, Madva, & Reuter, 2017; Gawronski & Bodenhausen, 2006, pp. 706–707; Madva, 2016, p. 2681, 2019; see also Brownstein, 2018, Chs. 2–3). This is not to say all our predictions panned out, but to question the ease with which other pluralist approaches are pigeonholed into the dreaded empiricist/associationist/behaviorist position in these recurring debates (e.g., Kurdi, Morris, & Cushman, 2022b, p. 3). Indeed, according to Mandelbaum (2022, sect. 8), we represent "a revival of associationist theories in philosophy," citing a paper that is explicitly orthogonal to the association–proposition debate (Madva & Brownstein, 2018, sect. 6.1; see also Kurdi, Mann, Charlesworth, & Banaji, 2019; Phills, Hahn, & Gawronski, 2020). With apologies to Voltaire, one senses that if modern-day associationists did not exist, modern-day Fodorians would have to reinvent them. With apologies to Taylor Swift, I would very much like to be excluded from this narrative.

In any case, the downfall of pure-associative models has not occasioned the uncontested reign of propositional alternatives. Leading propositional theorists continue to uncover effects more naturally explained by nonpropositional processes, or at least uneasily assimilated into prevailing propositional theories (e.g., Van Dessel, De Houwer, Gast, Roets, & Smith, 2020; see also Byrd, 2021). As a recent meta-analysis by Kurdi, Morehouse, and Dunham (2023, p. 1) explains, no current theory is well-poised to predict and explain the disorienting array of findings, and the time for "existence proof demonstrations" of propositional effects has passed. Yet in lieu of synthesizing the disarray, the target article consists in just such a grab bag of existence proofs, trumpeting all and only recent successes for propositional approaches – while ignoring evidence of their shortcomings and boundary conditions, and deferring long-standing concerns about how LoTs are implemented in the brain and integrated with other processes.

The authors nevertheless advertise LoTH's "unificatory power" (target article, sect. 1, para. 7), specifically its provision of a lingua franca mediating between psychological domains (perception, higher-order thinking, so-called "system 1," etc.). But if each of these domains involves proprietary LoTs *and who knows how many other* representational formats, the question still remains how these diverse representational formats interact with each other (within each psychological domain, rather than between domains). If non-LoTs interface with LoTs after all, what explanatory traction is gained by noting how LoTs pop up in lots of distinct psychological locales? And if a thousand other representational formats are abloom across the mind (target article, sect. 2, para. 2), why couldn't some of them mediate between domains, too? What we have here is not unification but proliferation, not explanation but more to explain.

No doubt the authors would cite their six core LoT properties as significant theoretical contributions. But the conceptual and causal interrelations of these properties (which are invoked in seemingly random combinations from one case study to the next) are muddled at best. Do they represent a homeostatic property cluster, as the authors claim, or are they tied more tightly together? The authors stress that representations involving discrete constituents need not be structured like sentences, but they "usually interpret" sentence-like representations "as requiring" discrete constituents (target article, sect. 2, para. 8). They then grant that successive properties on their list necessitate others, for example, predicate–argument structures and logical operators "requiring role-filler independence" (target article, sect. 2, para. 9). To the extent that property B *requires* property A, it is completely trivial to predict that A will show up wherever B does, and only slightly less trivial to predict that B will appear alongside A above chance. The mere prediction that properties "should tend to cooccur" (target article, sect. 2, para. 12) is weak, vague, and unconstrained, allowing theorists to underscore cooccurrences and ignore (or explain away) noncooccurrences. We "usually require" fewer degrees of freedom from our theoretical frameworks. We are also compelled to ask whether the six properties offer anything substantively novel or illuminating, or simply stick new labels on the analytic entailments contained in the original LoT view.

The target article at times positions itself as a lone voice of logic in an associationist wilderness, fighting the good fight for a nearly forgotten rationalist cause while flanked on all sides by zombie empiricisms that refuse to stay dead. Yet the article's principal value consists in signal-boosting others' important ongoing work. The question, then, is what it would mean to take up the authors' proposals over and above what the exemplary researchers being cited are already doing.

https://doi.org/10.1017/S0140525X22002849 Published online by Cambridge University Press

**Financial support.** This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

**Competing interest.** None.

## References

Brownstein, M. (2018). *The implicit mind: Cognitive architecture, the self, and ethics*. Oxford University Press.

Brownstein, M., & Madva, A. (2012). The normativity of automaticity. *Mind & Language*, *27*(4), 410–434. https://doi.org/10.1111/j.1468-0017.2012.01450.x

Byrd, N. (2021). What we can (and can't) infer about implicit bias from debiasing experiments. *Synthese*, *198*(2), 1427–1455. https://doi.org/10.1007/s11229-019-02128-6

Del Pinal, G., Madva, A., & Reuter, K. (2017). Stereotypes, conceptual centrality and gender bias: An empirical investigation. *Ratio*, *30*(4), 384–410. https://doi.org/10.1111/rati.12170

Fodor, J. A. (1985). Précis of the modularity of mind. *Behavioral and Brain Sciences*, *8*(1), 1–5. https://doi.org/10.1017/S0140525X0001921X

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Kurdi, B., Mann, T. C., Charlesworth, T. E. S., & Banaji, M. R. (2019). The relationship between implicit intergroup attitudes and beliefs. *Proceedings of the National Academy of Sciences*, *116*(13), 5862–5871. https://doi.org/10.1073/pnas.1820240116

Kurdi, B., Morehouse, K. N., & Dunham, Y. (2023). How do explicit and implicit evaluations shift? A preregistered meta-analysis of the effects of co-occurrence and relational information. *Journal of Personality and Social Psychology*, *124*(6), 1174–1202. https://doi.org/10.1037/pspa0000329

Kurdi, B., Morris, A., & Cushman, F. A. (2022b). The role of causal structure in implicit evaluation. *Cognition*, *225*, 105116. https://doi.org/10.1016/j.cognition.2022.105116

Madva, A. (2016). Why implicit attitudes are (probably) not beliefs. *Synthese*, *193*(8), 2659–2684. https://doi.org/10.1007/s11229-015-0874-2

Madva, A. (2019). Social psychology, phenomenology, and the indeterminate content of unreflective racial bias. In E. S. Lee (Ed.), *Race as phenomena: Between phenomenology and philosophy of race* (pp. 87–106). Rowman & Littlefield.

Madva, A., & Brownstein, M. (2018). Stereotypes, prejudice, and the taxonomy of the implicit social mind. *Noûs*, *52*(3), 611–644. https://doi.org/10.1111/nous.12182

Mandelbaum, E. (2022). Associationist theories of thought. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (winter). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2022/entries/associationist-thought/

Nolan, C. (2008). *The dark knight*. Warner Bros.

Phills, C. E., Hahn, A., & Gawronski, B. (2020). The bidirectional causal relation between implicit stereotypes and implicit prejudice. *Personality and Social Psychology Bulletin*, *46*(9), 1318–1330. https://doi.org/10.1177/0146167219899234

Van Dessel, P., De Houwer, J., Gast, A., Roets, A., & Smith, C. T. (2020). On the effectiveness of approach-avoidance instructions and training for changing evaluations of social groups. *Journal of Personality and Social Psychology*, *119*, e1–e14. https://doi.org/10.1037/pspa0000189

# A language of episodic thought?

Johannes B. Mahr and Daniel L. Schacter

Department of Psychology, Harvard University, Cambridge, MA, USA
jmahr@fas.harvard.edu;
dls@wjh.harvard.edu

doi:10.1017/S0140525X2300198X, e283

**Abstract**

We propose that episodic thought (i.e., episodic memory and imagination) is a domain where the language-of-thought hypothesis (LoTH) could be fruitfully applied. On the one hand, LoTH could explain the structure of what is encoded into and retrieved from long-term memory. On the other, LoTH can help make sense of how episodic contents come to play such a large variety of different cognitive roles after they have been retrieved.

Quilty-Dunn et al. convincingly show that language-of-thought hypothesis (LoTH) is alive and kicking in contemporary cognitive science. One domain they do not discuss, however, is episodic memory and imagination (i.e., episodic thought). This is not surprising: Traditionally, episodic thought has been widely viewed in terms of iconic forms of representation. Nonetheless, we believe that episodic thought is rife for being theorized in terms of LoTH. Most importantly, LoTH generates novel perspectives on how humans achieve such remarkable productivity and flexibility when thinking about other places and times.

Recent research on episodic memory and imagination suggests that different kinds of episodic thoughts (past memories, future imaginations, counterfactual imaginations, etc.) are cognitively not individuated through their contents (Addis, 2020, 2018; Mahr, 2020; Schacter et al., 2012); that is, episodic contents are "taxonomically neutral" with respect to their cognitive role as imaginations or memories. For example, Mahr, Greene, and Schacter (2021; see also De Brigard, Gessell, Yang, Stewart, & Marsh, 2020) found that participants' ability to recall the contents of a previously imagined event only weakly predicts their ability to recall whether this event was about the future or the past. This finding suggests that whether a given episode is taken to represent the past or the future (say) is not determined by what is retrieved from memory (i.e., episodic content) but by processes that occur before or after such retrieval (Mahr, 2020).

With this in mind, we propose that there are two main ways in which LoTH can be cashed out in episodic thought. On the one hand, LoTH can help to conceptualize the structure of episodic contents: What is *encoded* into long-term memory and how these contents are later *retrieved* in the service of the construction of both episodic memories and imaginations. According to the "constructive episodic simulation" hypothesis (Schacter & Addis, 2007), episodic retrieval consists in the flexible recombination of the elements of previously encoded experiences. Although there is good evidence to support this idea (see, e.g., Schacter & Addis, 2020, for a review), it remains unclear what mechanisms allow such flexible recombination of episodic elements in the service of episodic simulation. These processes are most commonly thought of in terms of associative inference (Addis, 2020; Carpenter & Schacter, 2017; Horner & Burgess, 2013) even though – as Quilty-Dunn et al. point out – LoT-style representations like scene-grammars (Võ, 2021), object files (Zimmer & Ecker, 2010), and event files (Hommel, 2004) play a role in structuring the information encoded into long-term memory. Similarly, these representations might play a role in structuring what content is retrieved and how it is composed. The fact that episodic contents could thus exhibit LoT properties – contributing to the flexibility and productivity of episodic simulation – has so far been underexplored. For example, evidence for the influence of "schemas" in episodic encoding and retrieval (Irish & Piguet, 2013; Renoult, Irish, Moscovitch, & Rugg, 2019), which also play a role in episodic simulation of future scenarios (Wynn, van Genugten, Sheldon, & Schacter, 2022), might be understood in this light (e.g., Draschkow, Wolfe, & Vo, 2014; Võ & Wolfe, 2013).

On the other hand, LoTH can help to understand how episodic contents come to play their respective cognitive roles. In the minds of adult humans, episodic contents can fill a variety of different roles – for example, as imaginations of past counterfactuals (De Brigard, Addis, Ford, Schacter, & Giovanello, 2013) or representations of event types (Addis, Pan, Vu, Laiser, & Schacter, 2009). A complete theory of episodic simulation requires an account of how "taxonomically neutral" episodic

https://doi.org/10.1017/S0140525X22002849 Published online by Cambridge University Press