



Edmunds, C.E.R., Wills, A.J. and Milton, F. (2025) 'Structure-based dissociations provide agnostic evidence to the multiple-systems debate', *Nature Reviews Psychology*, 4 (66).

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's [AM terms of use](#), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <https://doi.org/10.1038/s44159-024-00395-w>

ResearchSPAce

<http://researchspace.bathspa.ac.uk/>

This pre-published version is made available in accordance with publisher policies.

Please cite only the published version using the reference above.

Your access and use of this document is based on your acceptance of the ResearchSPAce Metadata and Data Policies, as well as applicable law:-

<https://researchspace.bathspa.ac.uk/policies.html>

Unless you accept the terms of these Policies in full, you do not have permission to download this document.

This cover sheet may not be removed from the document.

Please scroll down to view the document.

Structure-based dissociations provide agnostic evidence to the multiple systems debate.

C. E. R. Edmunds^{1†}, Andy J. Wills² & Fraser Milton³

¹ Psychology, Bath Spa University, Bath, UK.

² School of Psychology, University of Plymouth, Plymouth, UK.

³ School of Psychology, University of Exeter, Exeter, UK.

[†]email: ceremunds@gmail.com

Competing interests: The authors declare no competing interests.

Minda and colleagues review the literature in favour of multiple systems accounts of category learning (Minda, J. P., Roark, C. L., Kalra, P., & Cruz, A. Single and multiple systems in categorization and category learning. *Nat Rev Psych*, 1-16 (2024).)¹, focusing on the COVIS model². This Review provides a comprehensive and nuanced account of the evidence, mostly focusing on experiments that report dissociations between learning in rule-based category structures (in which optimal categorisation can be achieved using simple verbalizable rules) and information-integration structures (in which optimal learning requires participants to pre-decisionally combine information from two or more incommensurate stimulus dimensions). However, there is a substantial issue in the interpretation of these results which might undermine the authors' conclusions.

COVIS hypothesises that there are two learning systems: an 'explicit system' that implements verbalizable rule strategies using working memory and an 'implicit system' (or procedural system) that uses a strategy of combining stimulus features and associating them with a response.² Critically, COVIS predicts that participants sometimes switch learning systems. In rule-based conditions, participants start learning using the explicit system with its easily verbalizable rules, find that it results in high accuracy, and continue using that system. By contrast, in information-integration conditions, participants start learning using the explicit system but discover that it does not result in optimal responding because the structure is difficult to verbalise and applying rules results in poor accuracy, which motivates a switch to the implicit system. Thus, COVIS predicts that in information-integration conditions, there is always a portion of the experiment in which participants use the non-optimal explicit system. For a researcher to draw conclusions about the properties of the implicit and explicit systems from experimental data, they first need to determine which system the participant used during the experiment.

The typical method to determine which system is used is decision-bound modelling.³ This method finds the decision bound (most typically a straight line through stimulus space) that best separates the stimuli that the participant judges as one category from the other. In the COVIS literature, some decision bounds are assumed to correspond to rule-based strategies and—as only the explicit system can implement rule-based strategies—participants classified in this way are assumed to be using the explicit system. Other sorts of decision bounds are

associated with non-rule-use and—as only the implicit system could implement those strategies—these participants are assumed to be using the implicit system. Thus, unlike what Minda and colleagues suggest, the most important contribution of this analysis is not determining what strategy participants are using per se. Rather, it confirms that participants are using a strategy that would be consistent with the system they are purportedly using.⁴⁵ As each system is hypothesised to implement a distinct set of strategies, by identifying the individual strategy a participant uses, one can infer which system was in control of responding. Thus, decision-bound modelling in this context is a manipulation check. For instance, if participants' responses are best described by a diagonal boundary through stimulus space, this strategy is consistent with the predictions of the implicit system, but not the explicit system. Only if the participant's strategy could be implemented by that system can the effect of the manipulation be used to infer properties or features of that system.

However, previous work has shown that decision-bound modelling results are biased by the category structure of the task.⁵ Rule-based decision-bound models are more likely to be found in rule-based tasks, and information-integration models in information-integration tasks, regardless of what the participant was actually doing. Thus, decision-bound modelling cannot reliably determine participants' strategies and therefore cannot be used to infer which system the participant is using.

As a consequence of this bias, decision-bound modelling cannot be used as a manipulation check. Without a method to prove that participants use a certain system, one cannot assume that any dissociations in task performance are due to the manipulation selectively interacting with a specific system. However, it is precisely this type of evidence that Minda and colleagues review in support of the COVIS model. Without a clear determination of the system in use, the evidence for the existence of multiple categorization systems is at best ambiguous. Participants in those experiments might have been using multiple systems optimally (according to COVIS), but we cannot tell using the available evidence.

References

1. Minda, J. P., Roark, C. L., Kalra, P., & Cruz, A. Single and multiple systems in categorization and category learning. *Nat Rev Psych*, 1-16 (2024).
2. Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U. & Waldron, E. M. A neuropsychological theory of multiple systems in category learning. *Psychol. Rev.* **105**, 442–481 (1998).
3. Ashby, F. G., & Gott, R. E. Decision rules in the perception and categorization of multidimensional stimuli. *JEP:LMC*, *14*(1), 33, (1988).
4. Ashby, F. G., & Crossley, M. J. (2010). Interactions between declarative and procedural-learning categorization systems. *Neurobiology of learning and memory*, *94*(1), 1-12.
5. Edmunds, C. E., Milton, F., & Wills, A. J. Due process in dual process: Model-recovery simulations of decision-bound strategy analysis in category learning. *Cog. Sci.* **42**, 833-860 (2018).