THEORIA

Editors' introduction

(Introducción de los editores)

Marc Artiga
Universitat de València
https://orcid.org/0000-0003-1247-2809

Manolo Martínez
Universitat de Barcelona
https://orcid.org/0000-0002-6194-7121

The concept of representation enjoys the curious (if perhaps not uncommon) status of being both widely perceived as foundational to cognitive science and artificial intelligence, while there being nothing close to a consensus as to what it refers to, or even if it refers to anything at all. Still, very prominent research programs in cognitive science and AI are predicated on the existence of representations, and explicitly framed in representational terms: it seems that many of our insights into cognition centrally depend on a sound understanding of the nature of representations, the explanatory roles they play, their make-up and structure. This special issue seeks to contribute to this ongoing debate by gathering cutting-edge work produced by what the editors believe to be five of the most interesting established and early-career researchers currently working on representation in cognitive science.

The articles gathered in this special issue can be classified into two main groups. The first one addresses representational models in cognitive science and AI, and discusses how they should inform our views on representation. An initial, pressing question for any representationalist theory of cognition is: should we take representational models in cognitive science at face value? Do the results of our best scientific theories commit us to the existence of mental representations? Discussions of these questions have tended to assume that representations are theoretical posits, so the question turned on whether they are required for our best scientific theories and therefore, in more or less Quinean fashion, whether we are required to accept, or at least warranted in accepting, their existence. Recently, however, some authors have tried to shift the frame of the debate focusing on scientific experimentation, rather than scientific theorizing (Thomson and Piccinini, 2018; Nanay, 2022). In a nutshell, the idea is that the way scientists experiment and manipulate representations can provide the strongest argument for representation realism. Drayson's (2025) contribution argues that such attempts to reframe the debate fail: the question about representation realism inevitably concerns the role of representations as theoretical posits. Whether one actually counts as observing or manipulating a representation still depends on one's theoretical assumptions: the focus on experimentation, enlightening as it is, falls short of providing a theory-independent realist vindication of representations.

Coelho Mollo's (2025) contribution explores a related set of questions from the perspective of recent work on AI. What can we learn about representations, computation and intelligence, by relying on AI models? Building on his original, behavioural characterisation of intelligence, Coelho Mollo argues that AI can not only help us understand better how human and other biological cognitive systems actually work or can be employed to recreate biological intelligence in artificial systems, but it can also play a fundamental role in helping us explore "intelligence space", i.e. different ways of being intelligent that may significantly differ from the ways human or other biological systems display intelligence. He illustrates this role with a case study centred on recent large language models (LLMs) that are able to pass difficult

tests involving conceptual combination. According to Coelho Mollo, these results suggest a way of being intelligent that is distinctly non-human. In consequence, they also support computational and representational multiple realizability, i.e. the idea that a capacity or function can be realized by multiple computational processes and can involve different kinds of representations with different content and formats. In particular, LLMs perform conceptual combinations without apparently possessing discrete and stable conceptual representations, but relatively unstructured ones. Moreover, Coelho Mollo argues, LLMs employ a non-rule-based, possibly non-compositional procedure. AI-as-exploration provides an original perspective that suggests a range of questions, and possible answers, on the nature and implementational details of computation and representation.

Despite their different perspectives, both contributions support a qualifiedly deflationary take on representations: whether, e.g., neural activity or vector transformations count as genuine forms of representation depends on the truth of central assumptions involving the nature of vehicles, content, explanation, and the like. The mere fact that one can observe, manipulate or create a system that is able to perform some difficult task does not irrevocably vindicate a representationalist explanation, in the absence of additional theoretical work. This, we think, is an important lesson of the first part of the special issue.

The second block focuses on one of the key topics in recent debates in the philosophy of cognitive science: structural representation. The main idea is that a better understanding of the structural relationship between vehicles and the target system of the representations they support should shed light on some pressing issues, such as the criteria for distinguishing genuine representations from mere causal relays; or the usual concerns and anxieties surrounding content determination. On the other hand, in recent literature the concept of structural representation is still very much in flux, and it is far from clear that it can meet those high expectations. The real value of this notion is still undecided.

First, Mann (2025) employs the formal tools of information and computational complexity theory to illuminate the following question: does structure, understood as certain relational properties between vehicles, really provide new information to be used by cognitive systems? In a sense it does: for any set of vehicles, this set plus the way it is organized can carry more information about a phenomenon of interest than the unstructured set alone. But in another sense it doesn't: another, unstructured set of vehicles that carries the same amount of information is always possible. Mann then suggests that it is complexity theory, rather than information theory, that has the better chance of identifying key differences between structural and non-structural representations (for instance, in terms of a particular problem being tractable or intractable). Thus, the paper can be regarded as an invitation to explore more deeply specific areas of information and computational complexity theory to provide a better characterization (and, eventually, vindication) of structural representations.

The next article of this second block, Mamak and Miłkowski (2025), can be understood both as an attempt to meet Mann's challenge and as a contribution to the debate on the value of representational models addressed in the first block of the special issue. The paper provides a specific and original understanding of a model based on structural representations and clarifies its nature by employing some tools of information theory. More precisely, Mamak and Milkowski introduce what they call the Correspondence Network Framework: a formal approach to cognitive representations that, among other things, aims to vindicate a representationalist approach to cognition by revealing the fundamental explanatory value of informational correspondences provided by structural similarities between different systems. This approach assumes and further develops a correspondence-based account of semantic information (Miłkowski 2023), according to which semantic information relies on structural similarity between informational structures. Based on this framework, the paper shows how these informational correspondences are exploited by cognitive systems, in a way that warrants attribution of genuine representations. To illustrate their approach, they discuss in detail models based on Representational Similarity Analysis (RSA), where the concept of representational geometry plays a central role.

Whereas contributions to the second block differ in their assessment of the explanatory value of structural representation, they both vindicate the use of formal tools to better grasp and develop the concept of structural representation.

Thus, they join an emergent trend that has delivered remarkable results in the context of mental representations and, more generally, the nature of signaling. We think the papers in this second part of the special issue and the formal tools they employ open a wide range of interesting questions to be explored in future research.

All in all, we are delighted to present this special issue, including what we think are excellent contributions to difficult debates on the concept of representation in cognition science. We would like to thank the authors for writing these thought-provoking contributions and submitting their work to this special issue. Finally, we thank the journal editors for accepting our proposal, for their confidence and support throughout the whole process and, more generally, for contributing with their work to a more open society, where ideas and arguments can be exchanged without paywalls.

REFERENCES

Coelho Mollo, D. (2025). AI-as-exploration: Navigating intelligence space. THEORIA. *An International Journal for Theory, History and Foundations of Science*. https://doi.org/10.1387/theoria.25837

Drayson, Z. (2025). Representations are (still) theoretical posits. THEORIA. *An International Journal for Theory, History and Foundations of Science*. https://doi.org/10.1387/theoria.27209

Mamak, W , & Miłkowski, M (2025). Meaning from matching: How representational mechanisms exploit structural similarity. THEORIA. *An International Journal for Theory, History and Foundations of Science*. https://doi.org/10.1387/theoria.25171

Mann, S. (2025). Quantifying information in structural representations. THEORIA. *An International Journal for Theory, History and Foundations of Science*. https://doi.org/10.1387/theoria.25212

Miłkowski, M. (2023). Correspondence theory of semantic information. *The British Journal for the Philosophy of Science*, 74(2), 485-510. doi: 10.1086/714804

Nanay, B. (2022). Entity realism about mental representations. *Erkenntnis*, 87(1), 75-91. https://doi.org/10.1007/s10670-019-00185-4

Thomson, C., & Piccinini, G. (2018). Neural representations observed. *Minds and Machines*, 28(1), 191-235. https://doi.org/10.1007/s11023-018-9459-4

Marc Artiga is Full Professor at the Department of Philosophy at the University of Valencia. He works in philosophy mind, philosophy of biology, philosophy of language and cognitive science.

Address: Departament de Filosofia, Facultat de Filosofia i CC. de l'Educació, Universitat de València, Avda. Blasco Ibáñez, 30, 46010 València, Spain. E-mail: marc.artiga@uv.es – ORCID: 0000-0003-1247-2809

Manolo Martínez is Associate Professor (professor agregat) at the Department of Philosophy of the Universitat de Barcelona. He works on the philosophies of cognitive science, mind, and biology.

Address: Facultat de Filosofia, Universitat de Barcelona, Montalegre, 6, 08001 Barcelona, Spain. E-mail: manolomartinez@ub.edu – ORCID: 0000-0002-6194-7121