

MIT Press • Open Encyclopedia of Cognitive Science

Multi-Representational Learning

Shaaron E Ainsworth¹

¹University of Nottingham

MIT Press

Published on: Nov 20, 2025

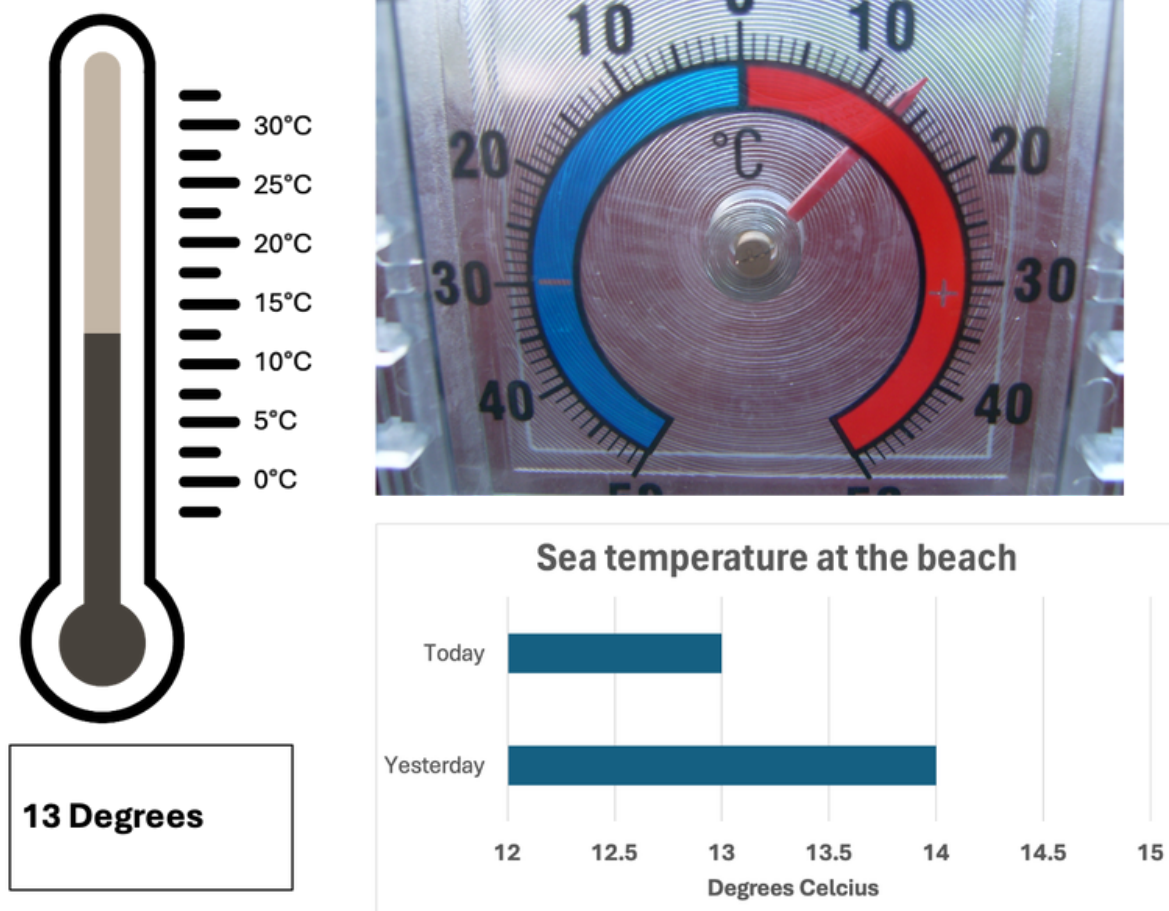
DOI: <https://doi.org/10.21428/e2759450.5cc1206d>

License: [Creative Commons Attribution-NonCommercial 4.0 International License \(CC-BY-NC 4.0\)](#).

Multi-representational learning refers to the way that people learn when they interpret more than one representation that others, often experts, have created or when they construct multiple representations for themselves. Examples include reading books replete with text, photographs, diagrams, or participating in an augmented reality–assisted visit to an archaeological site and later sketching its plan. Multiple representations have shaped human learning for millennia, and as new technologies are invented (e.g., paintings on cave walls, paper, printing presses, multimedia, and haptic interfaces) and their forms and modes of interaction continue to evolve. Multi-representational learning has drawn interest from different theoretical traditions, including cognitive, sociocultural, and semiotic approaches. Although there are key differences between these accounts, all argue for the importance of multi-representational learning, with applied cognitive scientists considering how best to take advantage of its ever-changing possibilities.

History

Although the discussion of representational learning can be argued to date to Plato, research on the topic within cognitive science emerged in the 1970s. [Palmer \(1978\)](#) suggested that to understand representations, two things are key: the represented world (the concept, phenomena, etc.) and the representing world, which reflects the former. Consider the represented world of sea temperature; there are different possible ways of representing the same phenomena (e.g., the sea temperature at a local beach, see [Figure 1](#)) and different possible represented worlds (different days) but in the same form (e.g., a picture of the thermometer in January, June, and October). Therefore, to be counted as a multi-representational system, the representations may be of different representing worlds, different representing worlds, or both (as is frequently the case).

**Figure 1**

A drawing, photograph (from

https://commons.wikimedia.org/wiki/File:Thermometer_13_Degree_Celsius.JPG), bar chart, and words (clockwise from top left) illustrating sea temperature at my local beach.

Semiotic definitions (e.g., [Peirce, 1992](#)) make it more apparent that a representation cannot be understood without considering what meaning someone makes from it. This approach illustrates that the thermometer picture could represent 13°C, but equally, it could be intended to signal “the last swim of the holiday” or “the impact of climate change on sea temperature.” What the creators intend may not be the interpretation that someone else makes. Consequently, multi-representational learning fundamentally depends on the interpretations made of the representations, which makes salient the need to understand the underpinning cognitive processes.

Accounts of cognitive processes underlying multi-representational learning initially followed an information-processing approach. For example, the cognitive theory of multimedia learning (see [Mayer, 2021](#) for a recent account) describes three stages. To be successful, selected aspects of the text and pictures from sensory memory must be actively processed (rather than passively received) by limited-capacity, modality-specific systems in working memory to create verbal and pictorial models [see [Working Memory](#)]. These are then integrated with knowledge stored in long-term memory. It follows that

when multi-representational learning involves different processing systems, combining them enhances an individual's capacity to learn. A distributed cognition account (e.g., [Zhang & Norman, 1994](#)) considers multi-representational learning as a function of how the task draws on the structures of the internal mind and external environment. Consequently, representations not only support memory but they also structure cognition and even change the nature of the task. Finally, an embodied cognition account of multi-representational learning (e.g., [Nathan, 2021](#)) postulates direct bidirectional links between the sensory processing of representations and the perceptual action schemas in long-term memory that directly encode them. This makes the choice of representations fundamental in embodied theories of multi-representational learning, as it is action with representations that is the knowledge to be learnt.

Core concepts

Important concepts in multi-representational learning include *informational* and *computational equivalence* (e.g., [Palmer, 1978](#)). Two representations are said to be *informationally equivalent* if the same information is inferable from both. They are said to be *computationally nonequivalent* if this information is more easily drawn from one of them. For example, diagrams utilize location and size in space ([Tversky, 2011](#)) to reduce the costs of search and recognition compared to sentential, symbolic representations ([Larkin & Simon, 1987](#)). Sentential forms are more able to handle abstractions or complex negations ([Schnotz, 2002](#)). Consequently, a fundamental rationale for multi-representational learning is to take advantage of the complementary computational differences between representations.

Another core concept is *ontological commitment* ([Kress, 2009](#)) or *specificity* ([Stenning & Oberlander, 1995](#)). Although the terms differ, the fundamental idea is that certain forms of representation (typically those that are iconic and more closely resemble what they are representing) compel certain decisions by the creator [see [Iconicity](#)]. They are not arbitrary, and even if they privilege certain features or delete or exaggerate their form, the representation and its referent must overlap in some manner. For example, you can write “the knife is by the fork” without indicating the number of prongs on the fork, but this cannot be drawn without including such details (and many more). This can be useful in multi-representational learning, as more specific representations can help constrain the (mis)interpretations that learners may hold of another less specific form. For example, Celsius is not stated in the text in [Figure 1](#), but you were unlikely to infer Fahrenheit due to the other representations presented alongside it.

Multi-representational learning can also be beneficial when learners abstract over multiple representations to help understand what the invariant properties of the represented world are and what features are reflected ([Ainsworth, 2006](#)). For example, explaining how quadratic equations presented symbolically describe the area of squares presented, as pictures can provide learners with new insights.

These three functions of multi-representational learning (to complement one another, to constrain misinterpretation, and to support the construction of new knowledge) underpin its central role as a *social* and *cultural tool*. It supports social learning by externalizing knowledge, allowing collaborators to more easily understand one another ([Suthers, 2014](#)). For example, learners talking about a diagram can point to

it and gesture over it, even if they do not understand all of the words. External representation extends communication across time and space without the need for reproduction (as argued by [Morin, 2023](#)). Consequently, people can learn from others through their representations, even when the individuals are absent, hence developing cumulative culture [see [Cultural Evolution](#)]. From this perspective, multi-representational learning involves learning representational practices through action to perform culturally valued activities ([Säljö, 1999](#)). This is also true of specific subcultures (e.g., see the analysis by [Nersessian, 2024](#) of scientific practices or the account by [Goodwin, 1994](#) of professional vision in archaeology). In these accounts, multi-representational learning is not simply a helpful way to learn but the target of learning.

Questions, controversies, and new developments

The cognitive science of multi-representational learning draws on very diverse traditions whose assumptions, research questions, and methods differ markedly. For example, consider how people learn through relating language and pictures to each other. Researchers may explore the neuronal processes of text–picture integration that happen over milliseconds (e.g., [Li et al., 2020](#)) to establish how distinctive text and picture processing is and when semantic integration occurs. They may instead focus on whether learners’ outcomes are enhanced when they can integrate text and pictures in study material and whether learners must have sufficient prior knowledge to do so successfully (e.g., [Seufert, 2003](#)). However, other researchers ask questions about learning with text and pictures that look at how the phenomenon has evolved over centuries, such as the central role of social interaction in promoting abstraction in the emergence of written language from pictorial forms ([Garrod et al., 2007](#)). Others have even longer time frames in mind when they address how human cognition evolved to take advantage of multiple representations over many millennia ([Donald, 1993](#)). These differences make developing an integrated account of multi-representational learning challenging, and consequently, understandings of different aspects of the topic can develop in isolation from one another.

One long-standing controversy is the nature of the relationship between internal and external representations. For example, does successful multi-representational learning result in a single integrated internal representation in long-term memory whose form may not reflect the forms of the external representation (as predicted by some of the information processing accounts, e.g., [Mayer, 2021](#)), or does it result in multiple internal modality-specific representations that are drawn upon and recreated as needed (such as suggested in [Nathan, 2021](#))?

Applied cognitive scientists focus on issues such as how to support learners to master the complex demands of multi-representational learning. For example, how does a learner know what the envisioned interpretation is; the representations in [Figure 1](#) might be intended to show 13°C, but they also include additional information that may not be relevant (such as the change of color at 0°C or yesterday’s values). This is a very simple example, but in more complex scenarios, multi-representational learning is as much about learning about multiple representations as it is learning with and through them ([Ainsworth, 2006](#)).

Another key issue is to understand how forms of representation that result from technological development change multi-representational learning. For example, early use of generative artificial intelligence in education focused on creating text output from text input. This is no longer true with research now expanding to include text-to-image and other modalities ([Heilala et al., 2025](#)). There seems to be little doubt that the world is moving to a future with even more multi-representational learning, with an expanding diversity of forms and increased interactive possibilities.

Broader connections

Multi-representational learning receives explicit attention within the learning sciences (such as [White & Pea, 2011](#)), developmental and educational psychology (e.g., [Uttal & Doherty, 2008](#) and [Rau, 2017](#)), and science education (e.g., [Treagust & Tsui, 2013](#)). The field of semiotics is also very closely connected (e.g., see [Chandler, 2002](#) for an introduction). Research about drawing to learn is also deeply connected to multi-representational learning ([Ainsworth et al., 2011](#)), as learners frequently construct one representation from existing ones. Finally, one increasingly important area that implicitly, and occasionally explicitly, draws on multiple representations is data visualization and data literacy ([Binali et al., 2024](#)), and this provides an important context to explore the future of multi-representational learning.

Further reading

- Fan, J. E., Bainbridge, W. A., Chamberlain, R., & Wammes, J. D. (2023). Drawing as a versatile cognitive tool. *Nature Reviews Psychology*, 2(9), 556-568. <https://doi.org/10.1038/s44159-023-00212-w>
- Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The science of visual data communication: What works. *Psychological Science in the Public Interest*, 22(3), 110-161. <https://doi.org/10.1177/15291006211051956>
- Hegarty, M. (2011). The cognitive science of visual-spatial displays: Implications for design. *Topics in Cognitive Science*, 3(3), 446-474. <https://doi.org/10.1111/j.1756-8765.2011.01150.x>

References

- Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333(6046), 1096-1097. <https://doi.org/10.1126/science.1204153>

←

←

- Binali, T., Chang, C.-H., Chang, Y.-J., & Chang, H.-Y. (2024). High school and college students' graph-interpretation competence in scientific and daily contexts of data visualization. *Science & Education*, 33(3), 763-785. <https://doi.org/10.1007/s11191-022-00406-3>

↩

- Chandler, D. (2002). *Semiotics: The basics*. Routledge.

↩

- Donald, M. (1993). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press

↩

- Garrod, S., Fay, N., Lee, J., Oberlander J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, 31(6), 961-987. <https://doi.org/10.1080/03640210701703659>

↩

- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96(3), 606-633. <https://doi.org/10.1525/aa.1994.96.3.02a00100>

↩

- Heilala, V., Araya, R., & Hämäläinen, R. (2025). Beyond text-to-text: An overview of multimodal and generative artificial intelligence for education using topic modeling. *arXiv*. <https://doi.org/10.1145/3672608.3707764>

↩

- Kress, G. (2009). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

↩

- Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11(1), 65-100. <https://doi.org/10.1111/j.1551-6708.1987.tb00863.x>

↩

- Li, S., Chen, S., Zhang, H., Zhao, Q., Zhou, Z., Huang, F., Sui, D., Wang, F., & Hong, J. (2020). Dynamic cognitive processes of text-picture integration revealed by event-related potentials. *Brain Research*, 1726, 146513. <https://doi.org/j.brainres.2019.146513>

↩

- Mayer, R. E. (2021). Cognitive theory of multimedia learning. In L. Fiorella & R. E. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (3 ed., pp. 57-72). Cambridge University Press. <https://doi.org/10.1017/9781108894333.008>

↩

- Morin, O. (2023). The puzzle of ideography. *Behavioral and Brain Sciences*, 46, e233. <https://doi.org/10.1017/S0140525X22002801>

↩

- Nathan, M. J. (2021). *Foundations of embodied learning: A paradigm for education*. Routledge.

↩

- Nersessian, N. J. (2024). How do scientists think? Contributions toward a cognitive science of science. *Topics in Cognitive Science*, 17(1), 7-33. <https://doi.org/10.1111/tops.12777>

↩

- Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 259-303). Lawrence Elbaum Associates.

↩

- Peirce, C. S. (1992). *The essential Peirce, selected philosophical writings, volume 1 (1867–1893)*. Indiana University Press.

↩

- Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, 29(4), 717-761. <https://doi.org/10.1007/s10648-016-9365-3>

↩

- Säljö, R. (1999). Learning as the use of tools. In K. Littleton & P. Light (Eds.), *Learning with computers: Analysing productive interaction* (pp. 144-161). Routledge.

↩

- Schnotz, W. (2002). Commentary - Towards an integrated view of learning from text and visual displays. *Educational Psychology Review*, 14(1), 101-120. <https://doi.org/10.1023/A:1013136727916>

↩

- Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction*, 13(2), 227-237. [https://doi.org/10.1016/S0959-4752\(02\)00022-1](https://doi.org/10.1016/S0959-4752(02)00022-1)

↩

- Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19(1), 97-140. https://doi.org/10.1207/s15516709cog1901_3

↩

- Suthers, D. D. (2014). Empirical studies of the value of conceptually explicit notations in collaborative learning. In A. Okada, S. J. Buckingham Shum, & T. Sherborne (Eds.), *Knowledge cartography: Software tools and mapping techniques* (pp. 1-22). Springer. https://doi.org/10.1007/978-1-4471-6470-8_1

↩

- Treagust, D. F., & Tsui, C.-Y. (2013). Contributions of multiple representations to biology education. In D. F. Treagust & C.-Y. Tsui (Eds.), *Multiple representations in biology education* (pp. 349-367). Springer. https://doi.org/10.1007/978-94-007-41928_19

↩

- Tversky, B. (2011). Visualizing thought. *Topics in Cognitive Science*, 3(3), 499-535. <https://doi.org/10.1111/j.1756-8765.2010.01113.x>

↩

- Uttal, D. H., & Doherty, K. O. (2008). Comprehending and learning from “visualizations”: A developmental perspective. In J. K. Gilbert, M. Reiner, & M. Nakhlel (Eds.), *Visualization: Theory and practice in science education*. (pp. 53-72). Springer.

↩

- White, T., & Pea, R. (2011). Distributed by design: On the promises and pitfalls of collaborative learning with multiple representations. *Journal of the Learning Sciences*, 20(3), 489-547. <https://doi.org/10.1080/10508406.2010.542700>

↩

- Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, 18(1), 87-122. https://doi.org/10.1207/s15516709cog1801_3

↩