

The heterogeneity of mental representation: Ending the imagery debate

Author(s): Joel Pearson and Stephen M. Kosslyn

Source: *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 112, No. 33 (August 18, 2015), pp. 10089-10092

Published by: National Academy of Sciences

Stable URL: <https://www.jstor.org/stable/10.2307/26464860>

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/10.2307/26464860?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/10.2307/26464860?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



National Academy of Sciences is collaborating with JSTOR to digitize, preserve and extend access to *Proceedings of the National Academy of Sciences of the United States of America*

JSTOR

# The heterogeneity of mental representation: Ending the imagery debate

Joel Pearson<sup>a,1</sup> and Stephen M. Kosslyn<sup>b</sup>

<sup>a</sup>*School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia; and* <sup>b</sup>*Minerva Schools at Keck Graduate Institute, San Francisco, CA 94103*

Edited by Daniel L. Schacter, Harvard University, Cambridge, MA, and approved June 25, 2015 (received for review March 21, 2015)

**The possible ways that information can be represented mentally have been discussed often over the past thousand years. However, this issue could not be addressed rigorously until late in the 20th century. Initial empirical findings spurred a debate about the heterogeneity of mental representation: Is all information stored in propositional, language-like, symbolic internal representations, or can humans use at least two different types of representations (and possibly many more)? Here, in historical context, we describe recent evidence that humans do not always rely on propositional internal representations but, instead, can also rely on at least one other format: depictive representation. We propose that the debate should now move on to characterizing all of the different forms of human mental representation.**

mental imagery | imagery debate | working memory | mental codes | artificial intelligence

How do we humans represent information internally? One answer to this question is that the storage and manipulation of information rely on billions or trillions of electrochemical reactions in the brain. However, this answer is a lot like saying “What are buildings made of? Bricks, boards, steel, and concrete.” Such an answer is, of course, correct, but misses the point. Simply knowing the constituent components does not tell us much, if anything, about architecture. Simply knowing about the components does not tell us much about the function of the building, whether it is a house or a fire station.

To answer the question “How do we humans represent information internally,” we need to focus on the level of analysis that specifies how information is represented and processed. That is, we can analyze brain function not only at the level of individual neural events but also at the level of large neural ensembles that function to store and process information in specific ways. Following the convention in cognitive psychology, we take the “mind” to be the functional level of description of brain activity that specifies information processing.

For several decades, a debate about the nature of mental representation has raged, spanning many fields: notably, cognitive science, artificial intelligence, philosophy, and neuroscience. This debate arose during the period of “classical artificial intelligence” research in the 1970s, and hinged on ideas about how one could program computers to mimic mental events. That is, if one were building a mind, how would one represent information in it? Here, we use the term “mental representation” to refer to a physical

state that functions to store mental content, and in some cases this state can then be operated on flexibly in working memory or during mental imagery.

To be clear, the debate was not about what kind of information can be stored but rather about how the information is stored. In this context, we must distinguish between the content, which is the information being conveyed, and the format, which is the nature of the code used to represent it. The same content can be represented in many different formats; for example, the information in this sentence could be conveyed in Morse code, oral English, or written French. Similarly, visual information (which is a kind of content) could be described in words or depicted in a photograph, among many other formats. Many researchers, such as David Marr (20), believed that visual content is represented in a “symbolic” (i.e., propositional) format very early in the processing sequence. The “imagery debate” was not about whether visual information is stored; no one disputes that such content is stored. The issue was about how that information is stored.

One side of the mental representation debate argued that all information is stored in a symbolic, language-like, descriptive format, regardless of the content. There is no dispute that humans sometimes rely on language-like “propositional” representations; such propositional representations convey the gist of what is expressed in a verbal statement. The dispute is whether all mental representations rely on such an internal monolog. Another camp has argued that information can be stored in numerous different formats.

The initial debate focused on just two formats: propositional vs. depictive. Specifically, the debate was about whether, in addition to a descriptive format of the sort used in language, information can be stored in a depictive, pictorial format. In a depiction, each part of the representation corresponds to a part of the represented object such that the distances among the parts in the representation correspond to the actual distances among the parts. Thus, a depiction requires a functional space (e.g., an actual page or XY coordinate space).

We argue here that recent empirical findings have resolved this debate. Although the researchers may not have always conceived of their results in this context, recent empirical evidence now strongly supports the claim that we humans can represent information in multiple ways, and that such representations can be used flexibly in working memory or during mental imagery. This conclusion opens the next chapter for empirical research, namely, characterizing all of the different possible formats of mental representation, as well as discovering when and how they are used during cognition.

## Historical Roots of the Debate

Many philosophers have argued that depictive mental images play a key role in mental representation, but many others have argued to the contrary (reviewed in ref. 1). Being

Author contributions: J.P. and S.M.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. Email: jpearson@unsw.edu.au.

limited to logical analysis and synthesis, philosophers could not resolve the issue.

The behaviorist revolution in psychological research, which began in the early 20th century and peaked in the early 1950s, led psychologists to deny the existence of mental images. Behaviorists proposed that what is mistaken for “mental images” is, in fact, merely “inner speech,” and that any opposition to this view was simply inaccurate opinion. This strong antirepresentational view set the stage for many years to come, and, perhaps even today, the legacy of such strong views continues to influence how scientists regard the study of mental events.

It was not until the 1960s and 1970s that researchers began rigorously to study the nature of the mental representations used in learning, memory, and reasoning. For example, Shepard and Metzler (2) demonstrated that objects in visual mental images could be rotated, and that the farther they are rotated, the more time is required. Similarly, Kosslyn (3) showed that objects in visual mental images could be scanned, and that the farther one scans, the more time is required. The problem here, however, was that such behavioral data were too unconstraining. Response time differences typically can be explained in many ways, many of which do not require positing depictive representations. For example instead of a depictive image, an object could be internally represented as a description, for instance, as a set of linked propositions, and the apparent effects of scanning greater distances could be explained as simply having to transition between more links in such a propositional structure (reviewed in ref. 4). According to this view, the conscious experience of “seeing” a mental picture during visual mental imagery is epiphenomenal: It is like the heat thrown off by a light bulb when one is reading; the heat plays no functional role in the reading process. It was during the late 1970s that the imagery debate began (5).

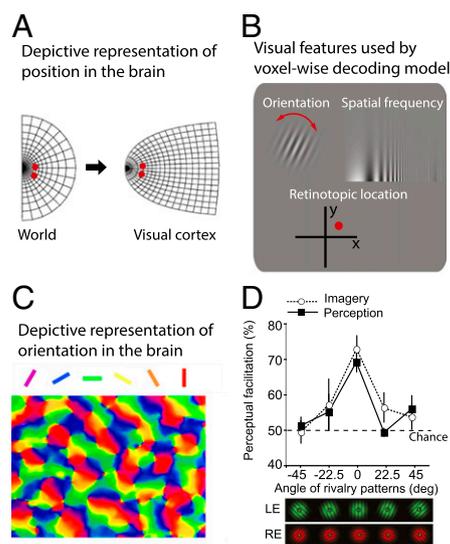
The debate took on a new form in the early 1990s, when neuroimaging became available. Turning to the brain made sense because many cortical areas are organized into map-like structures. The most striking example of such an area (and there are many) is area V1 (primary visual cortex), the first cortical area to receive visual signals from the eyes. This area is physically organized such that nearby neurons in the cortex have receptive fields that register stimuli at nearby locations in space, with additional dimensions embedded in these maps, such as orientation, eye of origin, spatial frequency, and color (6). Such cortical field maps are retinotopic; they preserve the spatial layout

of the retina. Damage to a local portion of area V1 produces a blind spot in the corresponding part of space; two lesions that are near each other on the cortex produce two blind spots that are near each other in space (Fig. 1A). Thus, these areas represent information, in part, by depicting it.

There is now strong evidence that when one visualizes (i.e., forms a mental image of) how something looks in darkness or with eyes closed, there is activity in area V1 (7–10). Because area V1 is depictive, these findings alone suggest that visual mental images involve depictive representations. However, the evidence from recent neuroimaging goes further: Researchers have been able to “read” or “decode” a mental image from patterns of activation in area V1. That is, just based on brain activity, researchers can learn what an individual is visualizing. Critically, overlapping activity patterns occur in retinotopically organized visual areas during imagery and visual perception. We know this because the algorithm used to decode visual mental images can be trained not on imagery, but on the pattern of activation in area V1 during visual perception (7, 8), and then used to decode activation during visual mental imagery (6, 11, 12). Because the algorithm was trained on depictive sensory representations in area V1 during perception (not imagery), imagery decoding can only work if the visualized stimulus involves some of the same patterns of activity in area V1 as the afferent sensory stimulus.

Although such research tells us that there is indeed substantial overlap between afferent sensory perception and imagery-induced activity, the source of such similarity remains unclear. For example, it is possible that a nondepictive commonality (e.g., focus of attention, expectation of reward) between imagery and perception is driving the algorithm’s decoding success. However, for the first time to our knowledge, a recent study was able to overcome this ambiguity by explicitly using a sensory multifeature-based encoding model (13). The authors fit a voxel-wise Gabor wavelet model to each voxel based on the blood oxygen level-dependent activity triggered by complex perceptual images. This model is specifically based on each voxel’s response to the retinotopic location, spatial frequency, and orientation of perceptual stimuli (Fig. 1B). After fitting the model to the perceptual data, the researchers found that the same model could successfully predict which pictures participants were visualizing.

This type of model-based decoding can only work if the depictive features embedded in the model are also evoked during visual mental



**Fig. 1.** Depictive visual features in the brain. (A) Position in visual space in the world is processed depictively in the brain. (B) Depictive visual features used by a voxel-wise model to decode mental images in a study by Naselaris et al. (13). (C) Spatial orientation is processed in a depictive manner in primate visual cortex, modified from a study by Blasdel and Salama (15). (D) Behavioral data from a study by Pearson et al. (16) showing how the content of a mental image biases later perception in an orientation (depictive)-specific manner. deg, degree; LE, left eye; RE, right eye.

imagery. Because the model was based on the depictive building blocks of afferent sensory perception (retinotopic location, spatial frequency, and orientation), the successful decoding confirms that these depictive features also characterize visual mental imagery (13).

Do such low-level visual features play a functional role in mental representation? Transcranial magnetic stimulation has been used to investigate this issue. Researchers showed that magnetic pulses delivered to the medial occipital lobe (the location of human area V1), compared with another location, impaired both visual mental imagery and visual perception, and did so to a comparable extent (14).

In addition, convergent evidence that relies on very different methodologies implicates depictive representations in area V1 during visual mental imagery. This area processes visual features such as spatial orientation in depictive cluster maps, in which neighboring spatial orientations are processed by neighboring neurons, with smooth transitions from one orientation to the next, as evident in the smooth color transitions in Fig. 1C (17). Thus, an orientation-contingent interaction between mental imagery content and a perceptual stimulus would constitute behavioral evidence that the mental image is being processed depictively. Behavioral research

has indeed demonstrated that visual mental imagery induces graded depictive orientation-specific effects on subsequent visual perception (16). This work demonstrated an orientation tuning function for visual mental imagery much like the tuning function found for visual perception. The effect of mental images on subsequent perception traverses through functional orientation space smoothly, just like depictive sensory stimulation (16, 17) (Fig. 1D). In these studies, participants were randomly cued to visualize either a vertical or horizontal pattern for several seconds, followed by a brief presentation of a binocular rivalry stimulus. On different trials, the binocular rivalry stimulus was parametrically presented through a range of different spatial orientations, which allowed mapping of the facilitative effect of imagery on subsequent rivalry through orientation space. The strongest bias effect occurred when the imagery and rivalry both contained the same orientation: a peak in the tuning function. Because the primary neurons selective for spatial orientation in this manner are in the depictive cluster maps in early visual cortex, these data support the claim that visualized orientation is represented depictively.

### Functional Role of Depictive Representations

Why might the brain produce depictive representations when we visualize how something looks? First, a depictive visual code respects the wiring optimization principle originally discussed by Ramon y Cajal (18). Because visual information in the environment tends to be highly correlated over space and time, the related neural activity also tends to be highly correlated based on retinal space and time. Neurons that fire together tend to wire together. Hence, to minimize the length of connections between neurons, cells with neighboring receptive fields should be as close as possible (19). Wiring based on correlated activity will lead to depictive maps of sensory-based activity, and these depictive cluster maps are used by the brain to represent visual information.

Second, depictive formats are useful for memory. Depictions promote the “principle of least commitment” (20); that is, they allow the brain to avoid throwing away potentially useful information. By their nature, images contain much implicit information that can be recovered retrospectively. For example, answer this question: What shape are a cat’s ears? Most people report visualizing the ears to answer. The shape information was implicit in the mental depiction, even though it was not explicitly considered at the time of encoding.

Third, depictive formats are useful in reasoning. They allow one to “simulate mentally” interactions that could occur in the real world, seeing the possible results of such interactions (21). Because depictions preserve incidental characteristics of objects (e.g., specific aspects of their shapes), they are useful when considering how objects will interact. For example, when packing luggage into a car’s trunk, it is useful to conduct a mental simulation at the outset, saving the physical work required to arrange the heavy objects into the most efficient configuration.

Fourth, if one of the functions of mental representations is to help us interact successfully with the world, and we represent the visual world depictively during the early stages of visual perception, then the brain might at times need to relate propositional language-like thoughts to these low-level depictively coded sensory experiences. Depictive mental representations might functionally bridge propositional information to depictive perception, allowing stored depictive information to change how we experience the world.

### Next Chapter

Given the preponderance of recent data, from many different sources, one can reasonably conclude that humans can use at least two forms of mental representation. This opens the door to a perhaps more interesting question: How many formats can the brain use? For example, do we have separate formats for motor, auditory, kinesthetic, and tactile information?

In addition, some theorists propose that all cognition involves grounded representation across all of the senses or modalities (22). Grounded or embodied cognition posits that all cognition, even abstract concepts such as justice and love, involve bodily or sensory representations. This theory proposes that perceptual symbols are extracted from sensory representations and stored in long-term memory, and it is these symbols that are used during cognition (23). Currently, the format of such proposed symbols is unclear; they might indeed involve depictive representations,

but symbolic content might also be represented in a propositional format early in the processing sequence, much like the propositional format proposed by David Marr (20). The new methods we discuss here could be applied directly to test the nature and format of representations involved in embodied cognition.

Another recent proposal focuses on the sensory representations induced by classical conditioning or by visual illusions that involve filled-in nonretinal vision. It is possible that such representations overlap with voluntarily induced mental representation like those that arise during mental imagery (24). Currently, it is unknown whether voluntary and involuntary visual representations share the same format.

Moreover, Paivio (25) long ago proposed dual coding theory, which rests on the idea that learned associations between two different types of information, verbal and nonverbal (mental imagery), are stored separately in long-term memory. To our knowledge, researchers have never studied the format of the nonverbal representations used in such learning.

There is now evidence that many cognitive processes, such as moral reasoning (26), language comprehension (27), autobiographical memory, imagining the future (28), dreams (29), and expectations about upcoming tasks (30), all involve sensory representations. What remains unclear, however, is the precise representational format of these visual mental representations.

Many other questions about mental representation beg to be addressed using the new techniques now available. For example: Why do some individuals have strong mental images when they think about how something looks, whereas others have weak images? What might be the consequences of such differences? It is up to future research to probe the formats of mental representation across the cognition and the senses, and the relationship between brain mechanisms and the sensory strength of such representations.

- Kosslyn SM (1980) *Image and Mind* (Harvard Univ Press, Cambridge, MA).
- Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. *Science* 171(3972):701–703.
- Kosslyn SM (1973) Scanning visual images: Some structural implications. *Percept Psychophys* 14(1):90–94.
- Kosslyn SM, Ganis G, Thompson WL (2001) Neural foundations of imagery. *Nat Rev Neurosci* 2(9):635–642.
- Pylshyn ZW (1973) What the mind’s eye tells the mind’s brain: A critique of mental imagery. *Psych Bull* 80(1):1–24.
- Mountcastle VB (1997) The columnar organization of the neocortex. *Brain* 120(Pt 4):701–722.
- Albers AM, Kok P, Toni I, Dijkerman HC, de Lange FP (2013) Shared representations for working memory and mental imagery in early visual cortex. *Curr Biol* 23(15):1427–1431.
- Stokes M, Thompson R, Cusack R, Duncan J (2009) Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J Neurosci* 29(5):1565–1572.
- Slotnick SD, Thompson WL, Kosslyn SM (2005) Visual mental imagery induces retinotopically organized activation of early visual areas. *Cereb Cortex* 15(10):1570–1583.
- Kosslyn SM, Thompson WL, Ganis G (2006) *The Case for Mental Imagery* (Oxford Univ Press, New York).
- Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458(7238):632–635.
- Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-specific delay activity in human primary visual cortex. *Psychol Sci* 20(2):207–214.
- Naselaris T, Olman CA, Stansbury DE, Ugurbil K, Gallant JL (2015) A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *Neuroimage* 105:215–228.

- 14** Kosslyn SM, et al. (1999) The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science* 284(5411): 167–170.
- 15** Blasdel GG, Salama G (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* 321(6070): 579–585.
- 16** Pearson J, Clifford CWG, Tong F (2008) The functional impact of mental imagery on conscious perception. *Curr Biol* 18(13):982–986.
- 17** Ohki K, et al. (2006) Highly ordered arrangement of single neurons in orientation pinwheels. *Nature* 442(7105):925–928.
- 18** Ramón y Cajal S (1899) *Textura del Sistema Nervioso del Hombre y de los Vertebrados trans* (Nicolas Moya, Madrid).
- 19** Chklovskii DB, Koulakov AA (2004) Maps in the brain: What can we learn from them? *Annu Rev Neurosci* 27:369–392.
- 20** Marr D (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, New York).
- 21** Moulton ST, Kosslyn SM (2009) Imagining predictions: Mental imagery as mental emulation. *Philos Trans R Soc Lond B Biol Sci* 364(1521):1273–1280.
- 22** Barsalou LW (2008) Grounded cognition. *Annu Rev Psychol* 59:617–645.
- 23** Barsalou LW (1999) Perceptual symbol systems. *Behav Brain Sci* 22(4):577–609, discussion 610–660.
- 24** Pearson J, Westbrook F (2015) Phantom perception: Voluntary and involuntary nonretinal vision. *Trends Cogn Sci* 19(5):278–284.
- 25** Paivio A (1991) Dual coding theory: Retrospect and current status. *Can J Psychol* 45(3):255–287.
- 26** Amit E, Greene JD (2012) You see, the ends don't justify the means: Visual imagery and moral judgment. *Psychol Sci* 23(8):861–868.
- 27** Dils AT, Boroditsky L (2010) Visual motion aftereffect from understanding motion language. *Proc Natl Acad Sci USA* 107(37): 16396–16400.
- 28** Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: The prospective brain. *Nat Rev Neurosci* 8(9): 657–661.
- 29** Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y (2013) Neural decoding of visual imagery during sleep. *Science* 340(6132): 639–642.
- 30** Kok P, Failing MF, de Lange FP (2014) Prior expectations evoke stimulus templates in the primary visual cortex. *J Cogn Neurosci* 26(7):1546–1554.