MIT Press • Open Encyclopedia of Cognitive Science

# Concepts

**Nicholas Shea**[1,2]

[1]Institute of Philosophy, School of Advanced Study, University of London, London, UK,
[2]Faculty of Philosophy, University of Oxford, Oxford, UK.

**MIT Press**

**Published on:** Jul 24, 2024

**DOI:** https://doi.org/10.21428/e2759450.1145a437

Concepts are recombinable elements of deliberate conscious thoughts. When I think *birds fly*, I use my concept of birds and my concept of flying. We think about the world by categorizing things under concepts. This allows us to use existing knowledge (the *bird* may well fly off). And when we learn something new (*birds* have feathers), concepts store that information systematically. As studied in cognitive science, concepts are mental representations: physical particulars (in the brain, and perhaps body, of the thinker) that refer to things in the world. A prominent version of the representational approach argues that concepts combine and behave like words of natural language—that we think in a language of thought. As information floods in from the world, concepts are a key way we make sense of what we perceive. They play a central role in thought, language, communication, and learning. They are powerful tools for organizing information, making inferences, and planning for the future.

## History

The most important advance in the scientific and theoretical understanding of concepts was the representational theory of mind. The fruitful reliance on the representational theory of mind and its accompanying *cognitive revolution* in the cognitive sciences included work on concepts (Margolis & Laurence, 1999), which are now standardly taken to be a type of mental representation (rather than treating concepts as abilities, say). The representational theory of mind shows how mental states and processes can be physically realized. Like words on the page or data packets in a fiber-optic cable, concepts, too, are physical entities that carry meaning or content. When one concept-involving thought follows another in a chain of thought, that is a causal process occurring between physical particulars. One set of representations causes the next. When an individual thinks about the same entity on different occasions, *birds*, say, they use different instances of their BIRD concept (different tokens of the same representation type).

As well as concepts' paradigmatic use in conscious thought, the same mental representations may show up elsewhere—for example, in long-term memory. They may figure in perception, for example when a person tracks track an object visually, and in other special-purpose systems—for instance, as labels in a cognitive map of spatial locations. It could also turn out that concepts occur in non-conscious processing.

Concepts sort objects in the world into categories. But not all categorical representations can figure as elements in deliberate thinking, and so count as concepts. Humans hear phonemes categorically; for example, English speakers hear a categorical difference between /r/ and /l/. Hearing the phoneme /r/ does not require that the hearer should deploy a representation of /r/ that can figure in deliberate thinking (or even that they possess such a representation). Furthermore, not all concepts are categorical. We have concepts of individuals (Ursula von der Leyen), of properties (length), and of abstract entities (numbers, beauty), to give some examples.

When non-human animal researchers use the term "concept," often the question is whether there are behavioral responses that generalize at the level of a category, or, sometimes, whether an animal can learn about abstract

categories like *same/different*.[1] Researchers are interested in whether behavior is driven by representations at all (contra behaviorism) and, if so, whether there are representations involved that are abstract in the sense of going beyond what the animal can represent perceptually. Some, but by no means all, would also require concepts to be processed cognitively in some way that goes beyond the merely perceptual.

Work on concepts in psychology and philosophy tends to make that commitment more explicit, holding that to count as a concept, a representation must be of a type that can figure in cognition (thus in conscious deliberation, for those that can deliberate). According to many theorists, a concept must also exhibit free recombinability. This is the influential *generality constraint* on what it takes to be a concept (Evans, 1982). The generality constraint says that someone who has the concepts ANJI, BRYCE, and LOVES can thereby think *Anji loves Bryce* or the converse and indeed can think *a loves b* for any singular concepts A and B that they possess.[2]

A different approach takes a concept to be not just a single mental representation about some X in the world, but a body of information about Xs: information the thinker uses to categorize Xs and draw inferences about them. Psychologists have worked hard to find out what kind of information is involved, focusing mainly on categorization.

The first idea was that a concept is a definition, a set of individually necessary and jointly sufficient conditions for being an X. For example, the concept BACHELOR might consist of the definition *unmarried eligible male human*. In this view, a thinker categorizes something under BACHELOR just in case they judge that it satisfies the definition. It turns out that it is hard to find satisfying definitions for all but a few concepts. Furthermore, the experimental evidence comes out against the putative definitional structure playing any psychological role in categorization (Murphy, 2002).

Inspired by philosopher Ludwig Wittgenstein, Eleanor Rosch proposed instead that we categorize the world using prototypes: a set of properties that Xs typically exhibit (Rosch, 1978). Birds typically fly, but a penguin has enough prototypical features that it is categorized as a bird, too, albeit an atypical one. The typicality effects Rosch discovered proved to be robust. Atypical instances are judged as such and are categorized more slowly and less accurately.

An alternative is that we store a small number of exemplars of the category (Smith & Medin, 1981), for example, a metal spoon and a wooden spoon, as examples of what it takes to fall under the concept SPOON. Category membership is judged on the basis of closeness to one or more exemplars. Exemplar effects prove to be prominent when people learn new, artificial categories, which are often based on geometrical figures or cartoon creatures. There is debate about whether prototypes or exemplars alone can account for all the experimental data, but both may well be needed.

A different advance, driven by work in developmental psychology, was the discovery that thinkers use theoretical knowledge to categorize objects and reason about them. For example, we may know that someone with flu has a virus and a runny nose, but we also know that the virus causes the nasal symptoms. An important place where theoretical knowledge is at work is with natural kinds, e.g., *skunk*. People presuppose that members of a kind share some essential property that is causally responsible for their surface features: children reliably say that an animal born as a skunk, but which acquires the look and behavior of a cat, nevertheless remains a skunk (Keil, 1989) [see Intuitive Theories].

The debate between prototype-, exemplar-, and theory-based theories of concepts gradually gave way to acceptance that all three kinds of psychological structure play a role in categorization and inference. Even for the same subject matter, different structures are used on different occasions. Although some have argued that each structure—prototype, exemplars, mini-theory—is a different concept (Weiskopf, 2009), a more common view is that concepts have a more heterogeneous character involving all three different kinds of information (*concept hybridism*: Vicente & Martínez Manrique, 2016), and perhaps more [see Core concepts section below].

One lesson is that researchers need to distinguish carefully between two uses of the term "concept." On the one hand, there is a representation that refers to things in the world and that people use to categorize and think about those things. On the other hand, there is a body of information that is used in categorization and inference. For example, it could have turned out that in thinking about any category X, people always started with a definition of Xs: categorizing using the definition and making inferences by starting with the definition, proceeding onwards from there. However, the experimental findings show that there is no single body of information through which all further cognition proceeds. Different kinds of stored information are used depending on the context.

One could conclude that there are, therefore, no concepts (Machery, 2009). A more constructive conclusion is that scholars can hang onto concepts in the mental representation sense, but we should accept that a single representation-type gives the thinker access to a diverse store of information, with different items of information being used in categorization and inference on different occasions. This entry disambiguates by reserving "concept" for the representation type rather than the wider body of information.

# Core concepts

## *What is a concept?*

A concept is a type of representation—a mental representation of Xs—that refers to things and the world and that figures as a recombinable element of deliberate conscious thoughts. It is helpful to have a separate term for the information that is deployed to categorize and think about Xs: conceptions (Millikan, 2000). Conceptions include semantic memories like *Paris is the capital of France*. According to some, conceptions encompass all

the semantic memories the thinker has about Xs: for example, *the Eiffel Tower is in Paris*, not just information that is central to Paris being the kind of thing it is. Conceptions include prototypes, exemplars, theoretical information, and functional information: what a thing is for and its affordances for action and use.

Conceptions also extend more widely. People often think about a thing in holistic and context-sensitive ways, relying on stereotypes, sensory images, and emotions. This wider collection of information serves to characterize the thing people are thinking about. An influential case has been made for including information of this kind in theorizing about concepts, dubbed *characterizations* (Camp, 2015). Thus, the conceptions associated with a concept include characterizations that are sensory and imagistic, functional and action-directed, and affective and evaluative.

Different disciplines within the cognitive sciences have focused on conceptions of different kinds. For example, anthropology has done much work to analyze the features underpinning kinship terms, to catalogue the structure of folk-biological taxonomies, and to understand how important concepts, like the concept of ownership, have evolved (Boyer, 2015). In linguistics, conceptual metaphor theory seeks to understand how characterizations are involved in the way hearers make sense of metaphorical language (Kövecses, 2021). In psychology, as well as a large body of work on categorization, empirical work has elucidated how conceptions are deployed in forming hierarchical taxonomies, combining categories, reasoning with theoretical information, and making inductive inferences about members of a category or between categories (Murphy, 2002). The primary concern is to understand the psychological structures and processes that are responsible for concept use. Philosophy has a different interest (Fodor, 1990; Margolis & Laurence, 1999; Millikan, 2000): to understand the semantic content of concepts—what a concept refers to so as to contribute to the truth conditions of thoughts, and why (and whether there are other aspects to concept meaning). It was initially hoped (or, worse, assumed) that there would be a common answer to these two questions. A definition, for example, could both be used in categorization and inference, and might determine reference—the set of things satisfying the definition. With the much richer picture researchers now have of the conceptions involved in categorization and inference, it has become clear that conceptions do not determine reference in any straightforward way. The philosopher's question needs to be sharply distinguished from the psychologist's question.

## *Externalist theories of concepts*

Philosophical theories of reference now appeal to other considerations to found the semantic value of concepts. According to most theories, some special mind–world relation between a concept and things in the world establishes reference. Since facts outside the thinker play a major role in determining reference, these theories of content are *externalist*: in two intrinsically indiscernible thinkers, the mental representation that refers to $H_2O$ in one thinker would refer to a different substance, XYZ, in the other, in a world where it was XYZ that filled rivers, lakes, and drinking glasses (Putnam, 1973).

A leading externalist theory starts with the observation that natural selection or learning has established certain patterns in which a thinker acts on their judgement *that is an X* to produce certain kinds of behavior. The concept they use refers to X in virtue of the fact that acting on or in relation to Xs has stabilized that pattern of behavior (Millikan, 1984, 2000); or there may be some kind of privileged informational connection between the concept and Xs (Dretske, 1988; Fodor, 1990). A thinker's conceptions play the causal role of mediating these connections, but they do not play a constitutive role in establishing reference.

When researchers say a thinker uses a concept to "pick out" certain things in the world, that phrase is ambiguous. In the semantic sense, the concept refers to a property, kind, individual, or set of things. In the psychological sense, the thinker uses the concept to categorize things (*that is an X*). When "picking out" in the second sense, the thinker can be mistaken (it's not actually an X). Applications of a concept to the world (categorization) do not straightforwardly determine what the concept *correctly* refers to. Externalism means that thinkers might even be quite radically mistaken about the property their concept refers to. They might, for example, think that arthritis occurs in muscles rather than joints. That does not stop their concept ARTHRITIS referring to inflammation of the joints (Burge, 1979). Externalism makes it straightforward to explain why the layperson is making an error, since the expert and the layperson can have concepts that refer to the same disease despite having different conceptions of it.

## Concept combination

The difference between psychological and philosophical concerns is also evident in debates about concept combination. Concepts combine to form truth-evaluable thoughts, e.g., ANJI LOVES BRYCE. They also combine to form new sub-propositional representations. A notorious example is PET FISH. Semantically, this might refer to the property of being *both a pet and a fish*. At the level of reference, the semantic value of the compound is plausibly fixed by the semantic values of the constituents and their mode of combination. But, psychologically, the story is not so simple. People have prototypes associated with PET and with FISH: something like a dog or a cat, and a salmon or a trout, respectively. They also have a prototype for PET FISH, perhaps like a goldfish. The prototype of the compound is not a simple function of the prototypes of the constituents. This leaves psychologists with the task of explaining where the prototype for PET FISH comes from, perhaps as some more complex function of the prototypes of PET and FISH (Hampton, 1997), maybe also involving real-world knowledge about pet fish (which could be stored with the PET concept, or the FISH concept, or both). The philosophical account of the referent of the compound is more complex, too. The free recombinability of concepts is a widely agreed datum, but the philosophical and psychological questions it raises are far from resolved.

Thinkers can also put together conceptions in a way that is tailored specifically to a certain occasion of use: what theorists have called an *ad hoc* concept. Ad hoc concepts can be complex. When moving house, I start thinking about the category *things to sell at a garage sale* (Barsalou, 1983). Simple concepts might also be ad hoc in the sense that people think about the category on each occasion in a somewhat different, context-specific

way. A person thinking about going hunting for birds will make use of different conceptions than someone thinking about keeping birds as pets. While these thoughts may deploy the same concept, qua representational vehicle, these cases highlight that the conceptions that are brought to bear are variable and occasion-specific.

## *Sense and reference*

A thinker can have two different concepts that refer to the same entity. A classic example is the ancient Babylonian astronomers who, without realizing it, had two concepts of Venus: HESPERUS for the morning star and PHOSPHORUS for the evening star. Researchers characterize the difference between these two concepts in terms of their associated conceptions. The first concept is associated with a conception of a star that rises in the morning, the second with a conception of a star that rises in the evening. If a concept comes with a stable set of conceptions shared between concept users, then there may be merit in picking out the concept in terms of those conceptions (in terms of its *sense*) in addition to doing so by what the concept refers to (its reference).

Developmental psychologists have discovered cases in which a small group of conceptions arrive together in development and explain a lot of the inferences that the thinker makes with a concept. For example, there is evidence that children acquire a small group of conceptions that together constitute a proper grasp of the concept of division, the key one being to understand that repeated division, though reducing a quantity, never takes you to zero (Carey, 2009). Similarly, there is a principle about counting and number that is argued to be constitutive of grasping the concept of natural number (Carey, 2009).

The prospects for concept individuation at the level of sense have recently been given a boost by an empirical finding in the world of machine learning. Computational linguistics has long attempted to track and individuate word meanings using only facts about the statistics of word distribution in sentences. The distributional hypothesis is that the meaning of a word can be individuated by the company it keeps—by the words that tend to occur near to it in ordinary sentences. That is the basic principle at work in large language models (LLMs): deep neural networks that are trained to produce sensible and intelligible text. The surprising ability of these systems to output seemingly intelligent language, and to translate between languages, probably depends on them representing text in a distributional state space (Piantadosi & Hill, 2022). This state space may in some cases reflect structure in the world, like the visual relations between colors, and not just the structure of language (Patel & Pavlick, 2022). While these results do not imply that LLMs understand the semantic spaces with which they represent the world (Mitchell & Krakauer, 2023), the work does make it more plausible that there are sets of conceptions, or of inferential roles, that are broadly shared between language users. If so, these are candidates for individuating concepts at the level of sense.

Whether or not there are senses shared between concept users, something must be shared when people communicate. For example, according to some, successful communication requires sharing thoughts with the same referent and some of the same (contextually-relevant) conceptions.

# Questions, controversies, and new developments

## *Types of conceptions*

There are two basic ways in which information can be stored with a concept: two kinds of conceptions. On the one hand, a concept may give the thinker access to further information. The semantic memory, *Paris is the capital of France* is like that. On the other hand, a concept may dispose the thinker to make certain inferences. The thinker may be disposed, for any singular term x, when thinking *x is a whale* to infer *x is a mammal*, with the inferential disposition triggered by tokening their WHALE concept in thoughts of that form. The concept comes with certain inference rules (Carston, 2010). These two kinds of conception are referred to as two different forms of conceptual structure. The set of characterizations to which a concept affords access is even richer but even for information that is factual in an everyday sense, there are two quite different ways of encoding it.

Concepts afford the thinker access to conceptions of various kinds. Some researchers argue for the importance of experiential characterizations (Camp, 2015). Often an affective response is involved. For example, there is a sense of awe that goes with the concept CATHEDRAL, and an affective evaluation is prompted when someone is categorized under the concept LIAR. In the latter case, the negatively valenced response is part of the content that is stored with the concept. For QUARTERBACK, the concept may involve a stereotype that includes not just properties like being a natural leader or being affable and a bit shallow, but also a sensory representation of a certain sort of square, clean-shaven jaw, gleaming teeth, and a ready smile (Camp, 2015).

These sensory, motoric, affective, and evaluative characterizations play an important role in concept-involving thought processes. There is now a large body of data showing effects of sensorimotor processing on the way conceptual thinking unfolds (Barsalou, 1999, 2003). Neuroscientific evidence points the same way (Pulvermüller, 2005, 2013). Some go further and argue that concepts just are sensorimotor simulations (Barsalou, 1999, 2003, see also Prinz, 2002). The claim is that concepts are a form of embodied cognition. The strong form of this hypothesis has been undermined by neuropsychological data showing that patients who lack motor abilities and motor activations exhibited by neurotypical subjects using a concept (e.g., “hammer”) nevertheless preserve standard conceptual abilities, like the ability to name pictures and say what the objects pictured are used for (Mahon & Caramazza, 2009). Controversy remains about whether embodiment, in some form, is nevertheless constitutive of conceptual thought (Calzavarini, 2021), or whether sensory information has relatively little impact on how concepts are organized (Bedny & Saxe, 2012), but in any case, the sensorimotor system is at least an important aspect of the kind of information to which a concept typically affords access. This is particularly apparent in the way concepts are involved in imagining or simulating possibilities in order to learn something new (Aronowitz & Lombrozo, 2020). Simulating situations and events is also a central part of problem-solving in science (Nersessian, 2018).

## Nativism or empiricism?

A second current debate concerns concept nativism. Does the developing child come to the world with some conceptual knowledge, so that important concepts can come on-stream in cognition without learning? Developmental psychologists have performed decades of ingenious experiments to investigate how different concepts develop. For example, an explicit, verbalizable concept of belief seems not to arise until between three and four years old. Other abilities trace back much earlier. Assumptions about the nature of objects, agents, and stuffs are at work in the way infants track things in perception even in the very first year of life (Spelke & Kinzler, 2007). Newborn infants have a disposition to track objects that have the characteristic configuration of a face, so it seems likely that some representational capacities, at least, are not learned. It is not clear that these early representations are concepts. For concepts, controversy remains: various core concepts may be not learned, or learning may be involved in every new concept the child acquires (Carey, 2015; Gopnik & Wellman, 2012; Karmiloff-Smith, 1994).

A philosophical question in the nativism debate is how learning can ever give rise to genuinely novel concepts. If concept learning were always a matter of putting together existing concepts in new configurations, then learning could never increase the thinker's expressive resources. To answer this worry, philosophers have shown how it is possible to introduce or *introject* a new representation into a psychological structure that endows it with a genuinely novel content (Laurence & Margolis, 2002; Strevens, 2012). On the psychological side, there is an influential and empirically-detailed account of how the child learns the concept of natural number in a way that expands their expressive resources (Carey, 2009).

## Cross-cultural variation or universality?

A third controversy concerns cross-cultural variation in concepts. This is related to the nativism question: if a psychological capacity or representation type is found to be cross-culturally universal, that offers some evidence that its development is canalized across normal variation in human physical and social environments, and hence that it may not depend on learning. Studying concepts across diverse populations is logistically more challenging than drawing subjects from researchers' local community—which, given the locations of most psychology departments, leads to biased samples. However, there is now a growing body of work investigating conceptual cognition in a wider range of cultures. For example, the seemingly intuitive distinction between cutting and breaking is not cross-culturally universal (Majid, Boster, & Bowerman, 2008), whereas there is evidence that there is a cross-culturally shared set of dimensions for organizing concepts of locomotion (Malt et al., 2015).

A further question is whether an evolutionary process of cultural selection shapes the form of some concepts. Concepts are tools used in thinking, so it is possible that they are *cognitive gadgets* that have been designed by cultural evolution (Heyes, 2018). While there are some suggestive case studies (Boyer, 2015), rigorous investigation of the kind of demanding evidence that would be needed to establish this claim is in its infancy.

The question is philosophically significant, since one of the most promising theories of content for concepts appeals to selective processes acting over lineages of concepts as the basis for their representational content (Millikan, 1984, 2000). Furthermore, these are processes cognitive scientists will need to understand if researchers are going to succeed in large-scale deliberate *conceptual engineering* with the aim of improving the way people conceptualize the world (Burgess, Cappelen, & Plunkett, 2020).

## Broader connections

Work on concepts in cognitive psychology is closely connected to work in linguistics on lexical semantics. Indeed, according to many, word meanings just are concepts. Questions remain about the priority between thought and language. While language can doubtless shape human thinking, the strong thesis that thinking just consists of subvocal use of language is not widely accepted. For instance, it seems that we can have concepts for which we do not have a word and which we cannot readily express in language.

An interesting case is polysemy, in which a single word seems to have several different but closely-related meanings. For example, “school” can be used to talk about a physical building (“the school burned down”) or an educational institution (“discipline at that school is strict”). Polysemy is different from homonymy, in which one word has two unrelated meanings (“bank” for river or money). Resolving the ambiguity of a homonymous word slows down the reader. Polysemy does not (if it really exists [Brody & Feiman, 2024]), but it does call for an account of how the hearer manages to understand the sentence in the right way. A theory of concepts can then explain how the thinker has access to information that provides the appropriate understanding of the word in both sentences (Quilty-Dunn, 2021). Polysemy resolution may be very like the phenomenon linguists have described in terms of ad hoc concepts (Carston, 2019).

Concepts research connects to broader questions in psychology about the nature of mental processing. A prominent theory argues that humans have two cognitive systems, System 1 and System 2 (thinking fast and thinking slow). Even if these are not really two different systems, there is good evidence that some types of processing draw heavily on working memory, making tasks prone to interfere with one another. Other cognitive processes proceed more automatically and can operate in parallel with little interference. This raises the question of which type of processing is involved when we think with concepts. The answer seems to be that it can be either. Work on category learning, for example, shows that people can learn a new concept effortfully and deliberately, inferring category membership based on logical operations on one or two features (spiky hair and no spots), or they can learn a category more automatically that weighs a whole array of features in parallel (Ashby & Valentin, 2017). More broadly, in making downstream inferences, as well as when categorizing objects, it seems that people can use concepts in both kinds of processing: fast and automatic, or slow and deliberative.

A final connection to mention is that with neuroscience. Theorizing about concepts is now deeply embedded in knowledge of what is going on in the brain. There is a long history of studying people with brain lesions to

understand how conceptual abilities fractionate and how they dissociate from other capacities, like the use of language. Damage to distinct areas of the brain produces semantic dementia, which is an impairment in the storage of information about a category, and semantic aphasia, which is an impairment of the capacity to manipulate and integrate information about a category (Lambon Ralph et al., 2017). Both are distinct from anomia, the inability to name an object (while still understanding what it is and what it is used for).

Techniques like functional magnetic resonance imaging (fMRI) provide further information about how areas of the brain support conceptual processing in neurotypical adults. Newer, powerful techniques for analyzing brain data (e.g., from fMRI or ECoG, electrocorticography) are beginning to throw light on the way patterns of neural activation are structured in conceptual thought. When people listen to radio programs, for example, activity in many parts of the cortex reflects the semantic structure of the words they are processing (Huth et al., 2016). The researchers can even train a decoder that, taking the fMRI signal as input, generates intelligible word sequences that roughly recover the meaning of what the subject is listening to (Tang et al., 2023). Other studies elucidate the neural structures involved in combining concepts into structured thoughts so as to differentiate between, for example, *dog bites man* and *man bites dog* (Frankland & Greene, 2020). Furthermore, there is some evidence that, as well as representing locations and their spatial relations, the hippocampus can form a more abstract relational structure that encodes relations between concepts (Schuck & Niv, 2019; Whittington et al., 2020). Similar neuronal codes, in which cells respond at regular intervals across a feature space forming a hexagonal pattern, have also been found outside the hippocampus, including in the prefrontal cortex. While this line of work is in its early stages, it raises the prospect that cognitive neuroscience will be able to complement behavioral psychology and neuropsychology and together build up a picture of the neural computations that are responsible for combining and processing concepts.

There is currently active research on concepts in many other interesting areas: normative concepts, abstract concepts (and their potential grounding in experiential processing), Bayesian models of concept acquisition and concept-based inference, concepts in perception and the putative perception–cognition distinction, vagueness and borderline cases, metacognition directed at concepts, the prospects for conceptual engineering, and the status of the philosophical method of conceptual analysis in the light of the best theories of the nature of concepts (Strevens, 2019).

While the decades since the cognitive revolution in the 1950s have brought a huge increase in the understanding of concepts and concept-involving psychological processes, deep questions remain. Most foundational is the question of how concepts come to be about or refer to things in the world. A theory of content should ultimately tell us how it is that this ubiquitous but puzzling feature of our mental life—that it is imbued with meaning—arises in the physical world.

## Acknowledgements

The author would like to thank Michael Strevens and the two editors for helpful feedback. This work was supported by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No. 681422 (MetCogCon).

## Further reading

- Carey, S. (2009). *The origin of concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001
- Margolis, M., & Laurence, S. (Eds.) (1999). *Concepts: Core readings*. MIT Press.
- Millikan, R. G. (2000). *On clear and confused ideas*. Cambridge University Press. https://doi.org/10.1017/CBO9780511613296
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.

## Footnotes

1. Italics are used to give the content of a thought (e.g. *birds fly*) or to refer to a property (e.g. *same / different*). Quotation marks are used to refer to a word or phrase, qua linguistic item (i.e. to mention rather than use the term). ↩

2. Small caps are used to refer to concepts. ↩

## References

- Aronowitz, S., & Lombrozo, T. (2020). Learning through simulation. *Philosophers' Imprint*, *20*(1), 1–18. ↩
- Ashby, F. G., & Valentin, V. V. (2017). Multiple systems of perceptual category learning: Theory and cognitive tests. In Cohen, H. & Lefebvre, C. (Eds.), *Handbook of categorization in cognitive science* (2nd ed., pp. 157–188). Elsevier. https://doi.org/10.1016/B978-0-08-101107-2.00007-5 ↩
- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227. https://doi.org/10.3758/BF03196968 ↩
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–609. https://doi.org/10.1017/S0140525X99002149 ↩
-

↩

- Bedny, M., & Saxe, R. (2012). Insights into the origins of knowledge from the cognitive neuroscience of blindness. *Cognitive Neuropsychology*, *29*(1-2), 56–84. https://doi.org/10.1080/02643294.2012.713342

↩

- Boyer, P. (2015). How natural selection shapes conceptual structure: human intuitions and concepts of ownership. In E. Margolis & S. Laurence (Eds.), *The conceptual mind*: *New directions in the study of concepts* (pp. 185–200). MIT Press. https://doi.org/10.7551/mitpress/9383.003.0013

↩

- Brody, G., & Feiman, R. (2024). Polysemy does not exist, at least not in the relevant sense. *Mind & Language*, *39*(2), 179–200. https://doi.org/10.1111/mila.12474

↩

- Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, *4*(1), 73–121. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

↩

- Burgess, A., Cappelen, H., & Plunkett, D. (Eds.). (2020). *Conceptual engineering and conceptual ethics*. Oxford University Press. https://doi.org/10.1093/oso/9780198801856.001.0001

↩

- Calzavarini, F. (2021). The conceptual format debate and the challenge from (global) supramodality. *The British Journal for the Philosophy of Science*. Advance online publication. https://doi.org/10.1086/717564

↩

- Camp, E. (2015). Logical concepts and associative characterizations. In E. Margolis & S. Laurence (Eds.), *Conceptual mind: New directions in the study of concepts* (pp. 591–622). MIT Press. https://doi.org/10.7551/mitpress/9383.003.0034

↩

- Carey, S. (2009). *The origin of concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195367638.001.0001

↩

- Carey, S. (2015). Why theories of concepts should not ignore the problem of acquisition. *Disputatio*, *7*(41), 113–163. https://doi.org/10.2478/disp-2015-0008

↩

- Carston, R. (2010). Explicit communication and ‘free’ pragmatic enrichment. In B. Soria & E. Romero (Eds.), *Explicit communication* (pp. 217–285). Palgrave Macmillan.

↩

- Carston, R. (2019). Ad hoc concepts, polysemy and the lexicon. In K. Scott, B. Clark, & R. Carston (Eds.), *Relevance, pragmatics and interpretation* (pp. 150–162). Cambridge University Press. https://doi.org/10.1017/9781108290593.014

↩

- Dretske, F. (1988). *Explaining behavior: reasons in a world of causes*. MIT Press.

↩

- Evans, G. (1982). *The Varieties of Reference*. Oxford: O.U.P.

↩

- Fodor, J. A. (1990). *A theory of content and other essays*. MIT Press.

↩

- Frankland, S. M., & Greene, J. D. (2020). Concepts and compositionality: In search of the brain's language of thought. *Annual Review of Psychology, 71*(1), 273–303. https://doi.org/10.1146/annurev-psych-122216-011829

↩

- Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085–1108. https://doi.org/10.1037/a0028044

↩

- Hampton, J. A. (1997). Conceptual combination: Conjunction and negation of natural concepts. *Memory & Cognition, 25*(6), 888–909.

↩

- Heyes, C. M. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

↩

- Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature, 532*(7600), 453–458. https://doi.org/10.1038/nature17637

↩

- Karmiloff-Smith, A. (1994). Précis of beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences, 17*(4), 693–707. https://doi.org/10.1017/S0140525X00036621

↩

- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. MIT Press.

  ↩
- Kövecses, Z. (2021). Standard and extended conceptual metaphor theory. In The Routledge handbook of cognitive linguistics (pp. 191-203). Routledge.

  ↩
- Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

  ↩
- Laurence, S., & Margolis, E. (2002). Radical concept nativism. *Cognition*, *86*(1), 25–55. https://doi.org/10.1016/S0010-0277(02)00127-0

  ↩
- Machery, E. (2009). *Doing without concepts*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195306880.001.0001

  ↩
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, *60*(1), 27–51. https://doi.org/10.1146/annurev.psych.60.110707.163532

  ↩
- Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, *109*(2), 235–250. https://doi.org/10.1016/j.cognition.2008.08.009

  ↩
- Malt, B. C., Gennari, S. P., 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.), *The conceptual mind: New directions in the study of concepts.* MIT Press. https://doi.org/10.7551/mitpress/9383.003.0019

  ↩
- Margolis, M., & Laurence, S. (Eds.). (1999). *Concepts: Core readings*. MIT Press.

  ↩
- Millikan, R. G. (1984). *Language, thought, and other biological categories*. MIT Press. https://doi.org/10.7551/mitpress/4124.001.0001

  ↩

- Millikan, R. G. (2000). *On clear and confused ideas*. Cambridge University Press. https://doi.org/10.1017/CBO9780511613296 ↩
- Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences of the United States of America, 120*(13), e2215907120. https://doi.org/10.1073/pnas.2215907120 ↩
- mith, E., & Medin, D. (1981). *Categories and concepts*. Harvard University Press. https://doi.org/10.4159/harvard.9780674866270 ↩
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press. ↩
- Nersessian, N. J. (2018). Cognitive science, mental modeling, and thought experiments. In M.T. Stuart, Y. Fehige, & J.R. Brown (Eds.), *The Routledge companion to thought experiments* (pp. 309–326). Routledge. ↩
- Patel, R., & Pavlick, E. (2022). Mapping language models to grounded conceptual spaces. *International Conference on Learning Representations*. ↩
- Piantadosi, S. T., & Hill, F. (2022). *Meaning without reference in large language models*. arXiv. https://arxiv.org/abs/2208.02957 ↩
- Prinz, J. (2002). *Furnishing the mind*. MIT Press. https://doi.org/10.7551/mitpress/3169.001.0001 ↩
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/nrn1706 ↩
- Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, *17*(9), 458–470. https://doi.org/10.1016/j.tics.2013.06.004 ↩
- Putnam, H. (1973). Meaning and reference. *The Journal of Philosophy, 70*(19), 699–711. https://doi.org/10.2307/2025079

↩

- Quilty-Dunn, J. (2021). Polysemy and thought: Toward a generative theory of concepts. *Mind & Language*, *36*(1), 158–185. https://doi.org/10.1111/mila.12328

↩

- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Laurence Erlbaum.

↩

- Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447), eaaw5181. https://doi.org/10.1126/science.aaw5181

↩

- Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

↩

- Strevens, M. (2012). Theoretical terms without analytic truths. *Philosophical Studies*, *160*(1), 167–190. https://doi.org/10.1007/s11098-012-9907-6

↩

- Strevens, M. (2019). *Thinking off your feet: How empirical psychology vindicates armchair philosophy*. Harvard University Press. https://doi.org/10.4159/9780674986473

↩

- Tang, J., LeBel, A., Jain, S., & Huth, A. G. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience*, *26*(5), 858–866. https://doi.org/10.1038/s41593-023-01304-9

↩

- Vicente, A., & Martínez Manrique, F. (2016). The big concepts paper: A defence of hybridism. *The British Journal for the Philosophy of Science*, *67*(1), 59–88. https://doi.org/10.1093/bjps/axu022

↩

- Weiskopf, D. A. (2009). Atomism, pluralism, and conceptual content. *Philosophy and Phenomenological Research*, *79*(1), 131–163. https://doi.org/10.1111/j.1933-1592.2009.00269.x

↩

- Whittington, J. C., Muller, T. H., Mark, S., Chen, G., Barry, C., Burgess, N., & Behrens, T. E. (2020). The Tolman-Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, *183*(5), 1249–1263.e23. https://doi.org/10.1016/j.cell.2020.10.024

↩