MIT Press • Open Encyclopedia of Cognitive Science

Categorization

Gary Lupyan[1]

[1]**University of Wisconsin-Madison**

MIT Press

Published on: Jul 16, 2026

URL: https://oecs.mit.edu/pub/mrj0pnj5

Categorization involves treating different things as the same. One common function of categorization is recognition. When an animal recognizes a new substance as being edible, it is categorizing it as a member of the category of edible things. Another function is inference. Categorizing an object as an apple licenses inferences about its likely taste and that it will bruise if dropped. For people, a common use of categorization is naming. Although the differences between a kitchen chair and an office chair are readily apparent, in referring to them both as "chairs," we are explicitly categorizing them as members of the category of chairs. Because categorization is entangled with perception, memory, language, inference, and decision-making, understanding any of these processes requires grappling with how categorization works, making it a foundational topic in the cognitive sciences. The study of categorization includes the computational processes that enable it, the study of factors that influence the ease of categorization, and the relationship between our concepts (the products of categorization) and the structure of the world—whether concepts reflect a pre-existing structure or help impose one, rendering the world comprehensible in the first place.

History

The study of categories and categorization—the manner and mechanisms by which we place things into categories—ranks among the oldest of psychological inquiries. In the fourth century BCE, Plato distinguished forms accessible to us by our senses from eternal and unchanging forms that can only be grasped by the intellect. According to Plato, the beds and horses we see are only shadows or imperfect reflections of their true forms. In his *Categories*, Aristotle rejected the idea of immaterial essences, suggesting instead that essences existed within the objects themselves. Aristotle further drew a distinction between *natural kinds* (simple substances, plants, and animals) and *artifacts* constructed by people. Natural kinds were defined by their essences. Artifacts like beds lacked natural essences and were instead defined in terms of their intended functions; the definition of a bed is an object constructed for sleeping.

For both Plato and Aristotle, category membership was absolute, determined by essences or by necessary and sufficient conditions. These ideas formed the *classical view*, which became a starting point for modern research on categorization (see Murphy, 2002). A key philosophical challenge came from Ludwig Wittgenstein, who observed that everyday categories like games lacked defining conditions. Instead, category members shared *family resemblance*. Importantly, Wittgenstein's argument was not that categories are defined by family resemblance (i.e., that something is a game if it shares sufficiently many features with other games) but rather that when we use a category, we impose on it a definition suitable for certain goals (Ramscar & Hahn, 1998).

The classical Western tradition can be contrasted with some classical non-Western conceptions of categories. In the Buddhist text *Questions of King Milinda*, King Milinda asks the Buddhist monk Nagasena, "What is your name?" Nagasena responds that he is known as Nagasena but that it is merely a conventional label. Pressed by the King, Nagasena makes the point by analogy: What makes the King's chariot a chariot? Not the wheels, the yoke, or the chassis on their own, nor the parts laid out flat, but the parts assembled and functioning together. Nagasena argues that this is also true of the person called Nagasena—the name refers to a temporary arrangement of parts. Neither artifacts nor natural kinds have essences or eternal definitions; categories are defined by convention and habitual use rather than by necessary and sufficient conditions. This contrast between Western and Eastern classical traditions foreshadows many key debates in the empirical study of categorization.

One of the first detailed explorations of the mental process of categorization is found in William James's (1890) *Principles of Psychology*. For James, categorization is at its core about detecting sameness among unlike things—a capacity he treated as foundational to recognition, memory, and reasoning. Rather than relying on essences, rules, or definitions, James placed association at the center of the mind's workings, treating it as the mind's most powerful organizing force. Drawing on thinkers like James Mill and John Stuart Mill, James described categorization as the process of learning to attend to particular associations in the service of various goals. For example, James speculated that when people learn to attach distinct names to tastes that initially seem similar (like different wines), the names help them notice differences between the tastes more easily—making the tastes feel more distinct and easier to tell apart. He also proposed a principle called *dissociation by varying concomitants*: when a feature shows up paired sometimes with one thing and sometimes with another, its link to any particular thing weakens, allowing the feature itself to stand out as something we can notice on its own.

We then see a curious split. James's focus on associations anticipated Pavlov's work on classical conditioning and went on to influence foundational work in behaviorist psychology. For example, his ideas on associating stimuli with more versus less distinctive cues were turned into the doctrines of acquired equivalence and acquired distinctiveness of cues within behaviorist psychology (Miller & Dollard, 1941). Although behaviorists avoided "mentalistic" terms like categories, concepts, and recognition, a large part of the behaviorist research program can be viewed as a study of categorization. For example, in John B. Watson's *Little Albert* study (Watson & Rayner, 1920), 11-month-old Albert was exposed to a white rat that was repeatedly paired with a loud sound. This caused Albert to show distress several months later not only to the rat but also to a rabbit, a fur coat, and a bearded Santa Claus mask. This is *generalization*. Many subsequent studies went on to investigate the *generalization profile* of reinforced stimuli. For example, Guttman and Kalish (1956) trained pigeons to associate a green light with food and found that pigeons pecked most at lights of the trained color, gradually decreasing their pecking as the light's color became less similar to what was reinforced.

Despite such evidence of *category gradedness* and *fuzzy boundaries*, early research on categorization in cognitive psychology remained firmly rooted in the Western classical view, with its emphasis on sharp category boundaries and categories defined by necessary and sufficient conditions. Many category learning experiments, beginning with Bruner et al. (1956), constructed categories based on logical combinations of discrete features such as "all red triangles" or "all figures with two borders," with success measured by how quickly learners could discover the rule governing category membership.

Subsequent work, described in more detail below, went on to challenge the idea of well-defined category membership and necessary and sufficient conditions as well as the idea that generalization is determined by similarity or feature overlap. Contemporary research has focused increasingly on the role of context, goals, and characterizing categories in terms of theories. This shift can be viewed as an echo of James's pragmatic view, according to which the essence of a category is simply whichever properties matter most for one's current purposes (James, 1890).

Core concepts

The functions of categorization

Categorization allows organisms to respond appropriately to classes of stimuli. The process of detecting functional equivalence—which perceivable differences matter and which do not—is central to so many aspects of cognition that some have argued that cognition is categorization (Harnad, 2005). Categorization serves two broad functions: generalization and inference.

Generalization

Consider a basic goal like locating food. If all food had a unique sensory signature, an animal would need only a sense organ tuned to that signature. The reality, especially for animals with varied diets, is that the class of edible things does not have a unique sensory signature. Reliably identifying stimuli as edible therefore requires representing edible things as sufficiently alike despite their varied appearance. More generally, because no two instances of anything are exactly alike, applying what one has learned from past encounters to a novel object requires treating that object as relevantly similar to those past instances.

Inference

Categorization also supports inference about properties that are not directly observed. If Mark likes to eat apples and categorizes a new thing as an apple, he can infer that he will also like eating this one. If he knows that apples are a fruit, he can conclude that this new thing is also a fruit. Note that drawing such inferences does not require placing an object into a discrete class. One might be unsure whether a fruit is an apple but judge it sufficiently similar to one to infer that whatever it is, it is edible. In other cases, however, appropriate behaviors require committing to a category. Approaching a street sign, a driver must decide whether it is a stop sign, which determines the appropriate action. Another domain that requires commitment to categories is language [see Language Production]. Languages instantiate systems of categories, and when we name something, we commit to the category denoted by that word (see The relationship between categorization and language below).

Categorical perception

In a world where members of a given category have unique sensory signatures, organisms could simply use a sensory apparatus tuned to that signature. One example of such a system is "bug detectors" in some insect-hunting frogs (Lettvin et al., 1959), which selectively respond to small moving dark objects. The "cleverness" of nature to evolve such specially tuned receptors should not be underestimated. For example, mantis shrimp have evolved detectors sensitive to patterns of circularly polarized light reflected by mantis shrimp shells, which they appear to use for detecting other mantis shrimp (Gagnon et al., 2015).

As an animal's behavioral repertoire expands and there is a greater need for recognizing a larger number of ecologically relevant categories, a different solution begins to emerge: *categorical perception*. Our ability to categorize certain patterns of light as corresponding to a face, for example, comes not from having a sensory face detector but from our ability to combine more basic visual features—like edges, colors, and shapes—into category detectors. This phenomenon, called *categorical perception,* was first studied in the domain of speech perception.

To recognize (i.e., categorize) that someone said "bin" rather than "pin" requires distinguishing /b/ from /p/. It turns out that these consonants can be distinguished on the basis of voice onset time: the time between the release of air produced by stop consonants like /b/ and /p/ and the subsequent onset of vocal cord vibration. For English speakers, shorter voice onset times are perceived as /b/ and longer ones as /p/. If an artificial continuum of sounds is created, ranging from a clear /b/ to a clear /p/, and people are asked to indicate what syllable they perceive, there is an inflection point at which responses sharply change from /b/ to /p/. It is also generally easier for people to distinguish a /b/ from a /p/ than it is to distinguish one /p/ from another /p/, even when the difference in voice onset time between each pair of sounds is similar (Liberman et al., 1957; for an example, see this demo). This evidence has been taken by some to mean that speech perception involves perceptual categorization mechanisms specifically tuned to speech sounds. Evidence that newborns already show such categorical effects was taken to mean that there exist innate phoneme detectors (Eimas et al., 1971) [see Hearing].

Subsequent research substantially complicated this picture. First, the claim of evolved specialization was challenged by findings that other animals like chinchillas show similar categorical perception of speech sounds (Kuhl & Miller, 1975). This suggests that whatever processes support categorical perception of phonemes reflect more general properties of mammalian auditory systems that speech takes advantage of (Smith & Lewicki, 2006). Subsequent findings of cross-linguistic differences in categorical boundaries (e.g., Werker & Tees, 1984) provided further evidence that categorical speech perception is a joint outcome of more general auditory biases and learning the phonemic contrasts of one's native language.

Second, categorical perception appears to be widespread. Instead of being limited to domains plausibly shaped by natural selection, like speech and face perception, it is also found for artifacts like cups and bowls (Labov, 1973), categories that are clearly culturally constructed. Taken together, these findings suggest that categorical perception can be understood as a domain-general mechanism of *feature* *reweighting*—a process by which perceptual systems learn (or evolve) to increase within-category similarity and between-category differences. Subsequent work further complicates the story by demonstrating that whether a perceptual input is categorized in the first place depends on the current task and context (see Effects of context on categorization).

Categorizing at different levels

All categorization involves some degree of *abstraction*—the process of setting aside certain features while emphasizing others, allowing distinct items to be grouped by shared properties. In categorizing something as red, we might abstract over its shape. In categorizing something as a circle, we might instead abstract over its color. The degree of abstraction depends on the level of categorization. Categorizing something as an animal requires more abstraction than categorizing something as a dog, which in turn requires more abstraction than categorizing something as a Schnauzer. Seminal work on the topic showed that categorizing at the *basic* level—categorizing something as a dog—is easier, or more natural, than categorizing at a *superordinate* level (categorizing something as an animal) or a *subordinate* level (such as categorizing something as a terrier; Rosch, 1978). The intermediate *basic level* reflected a compromise between informativeness and cognitive effort. Compared to the basic level, a subordinate level of categorization is more informative but requires more effort; it typically takes more time to categorize something as a Schnauzer than to categorize something as a dog. What makes a particular level of categorization "basic," however, is not intrinsic to the category but depends on what one knows about the domain. Because typical Americans know more about mammals than birds and plants (e.g., Medin & Atran, 2004), "dog," "bird," and "tree" are at the basic level despite denoting very different taxonomic levels. As one gains expertise in a domain, categorizing at more specific levels becomes relatively easier. For bird experts, identifying the species of bird is as easy as identifying something as a bird (Tanaka & Taylor, 1991).

Typicality and category goodness

Recall that despite behaviorist research on generalization gradients, early cognitive psychology emphasized categorization in discrete terms, with category membership being determined by whether an object possesses the features necessary and sufficient to be a member of the category. On this view, all members of a category are equally good members; there is no sense in which one bird could be more of a bird than another because each either meets the relevant criteria or does not. A challenge to this view came from work showing that people's categorization of both natural kinds and artifacts is graded. Unlike earlier behaviorist work on generalization gradients, this work used behavioral evidence of gradedness to understand the nature of mental representations, leveraging the ideas of *category prototypes* and *typicality*.

Just as some levels of categorization feel more natural than others, some members of a category feel more representative than others. A robin strikes most people as a better example of a bird than a penguin and a chair as a better example of furniture than a beanbag—a phenomenon known as typicality or category goodness. In a now classic study, Posner and Keele (1968) had participants learn to classify dot patterns into categories. For each category, the experimenters first generated a prototype—a particular arrangement of dots, such as a triangle—and then produced training items by randomly displacing each dot from its original position. Participants learned to classify the training items into categories—but never saw the prototypes during training. At test, they classified the never-seen prototypes better than new distortions and only slightly less accurately than previously seen distortions. These results were taken to mean that participants abstracted the central tendency—the *prototype*—of each category.

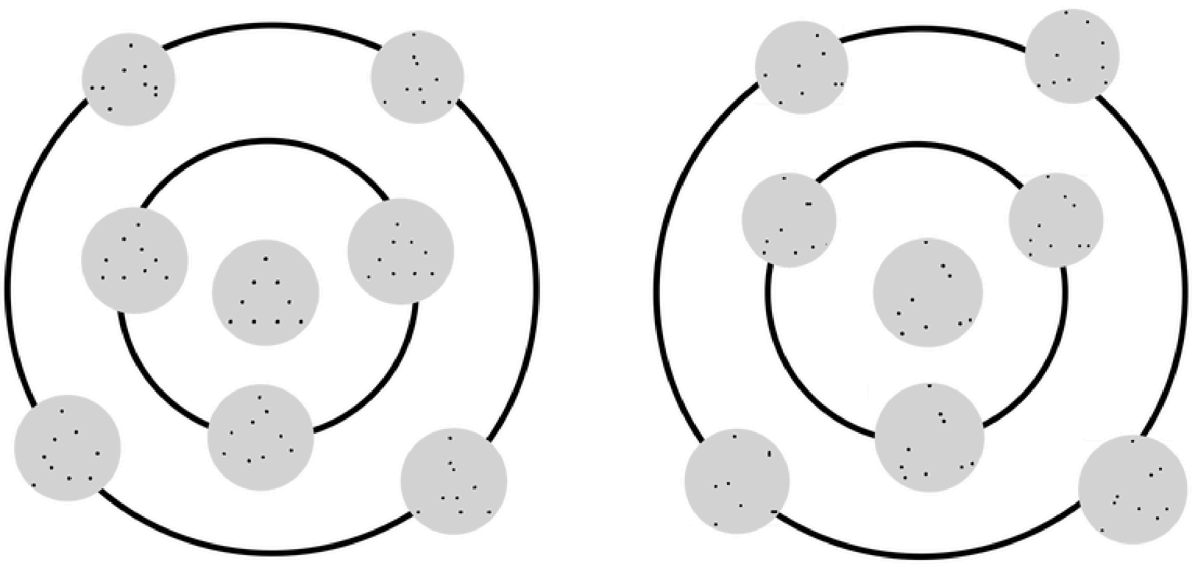

Figure 1
An adaptation of stimuli used by **Posner and Keele (1968)** to study generalization effects in visual learning. The key result shows that exposure to distortions of either familiar (left) or unfamiliar (right) prototypes sensitizes participants to the unseen prototype such that it is recognized/classified as easily as previously seen low-distortion items.

Later work extended the idea of prototypes to real-world categories, showing that typicality and gradedness effects pervade both natural kinds and artifacts (Rosch, 1978; Rosch et al., 1976). For example, when asked to list examples of furniture, people are much more likely to list chairs than bookshelves. The more frequently mentioned items tend to be judged as more typical and, if presented in a category verification task ("Is X furniture?"), are verified as category members more quickly and with more consensus. Typicality also influences what inferences people make. If told that robins (a typical bird) have higher potassium concentration in their blood than humans, people more readily infer that this is true of all birds than if they are told that penguins have higher potassium concentration in their blood than humans (e.g., Osherson et al., 1990). Although there may be technical criteria for determining whether something is a bird, psychologically speaking, category membership is a matter of degree.

The graded nature of categories is not limited to *conventional* categories like natural kinds and artifacts. Robust typicality effects are also observed for goal-derived categories. In such cases, the graded category structure emerges dynamically and appears to be organized around dimensions defined by the goal. For example, the category "things to take from one's home in case of a fire" is organized around value and irreplaceability; the category "things to not eat on a diet" is organized around caloric content (Barsalou, 1985). Typicality effects are also observed for *definitional* categories that have clear necessary and sufficient conditions. A triangle with three sides of different lengths is not only rated as less typical than a triangle with all equal sides, but people are also much more likely to think it is not a "real" triangle and to insist that its angles do not add up to 180 degrees (Lupyan, 2013).

Prototypes, exemplars, and theories

For stimuli like the dot patterns shown in Figure 1 (Posner & Keele, 1968), a prototype is a central tendency and can be computed by simply averaging the observed dot patterns. However, this characterization only holds for very specific contexts. Robins are judged to be very typical birds but are hardly average (Figure 2).

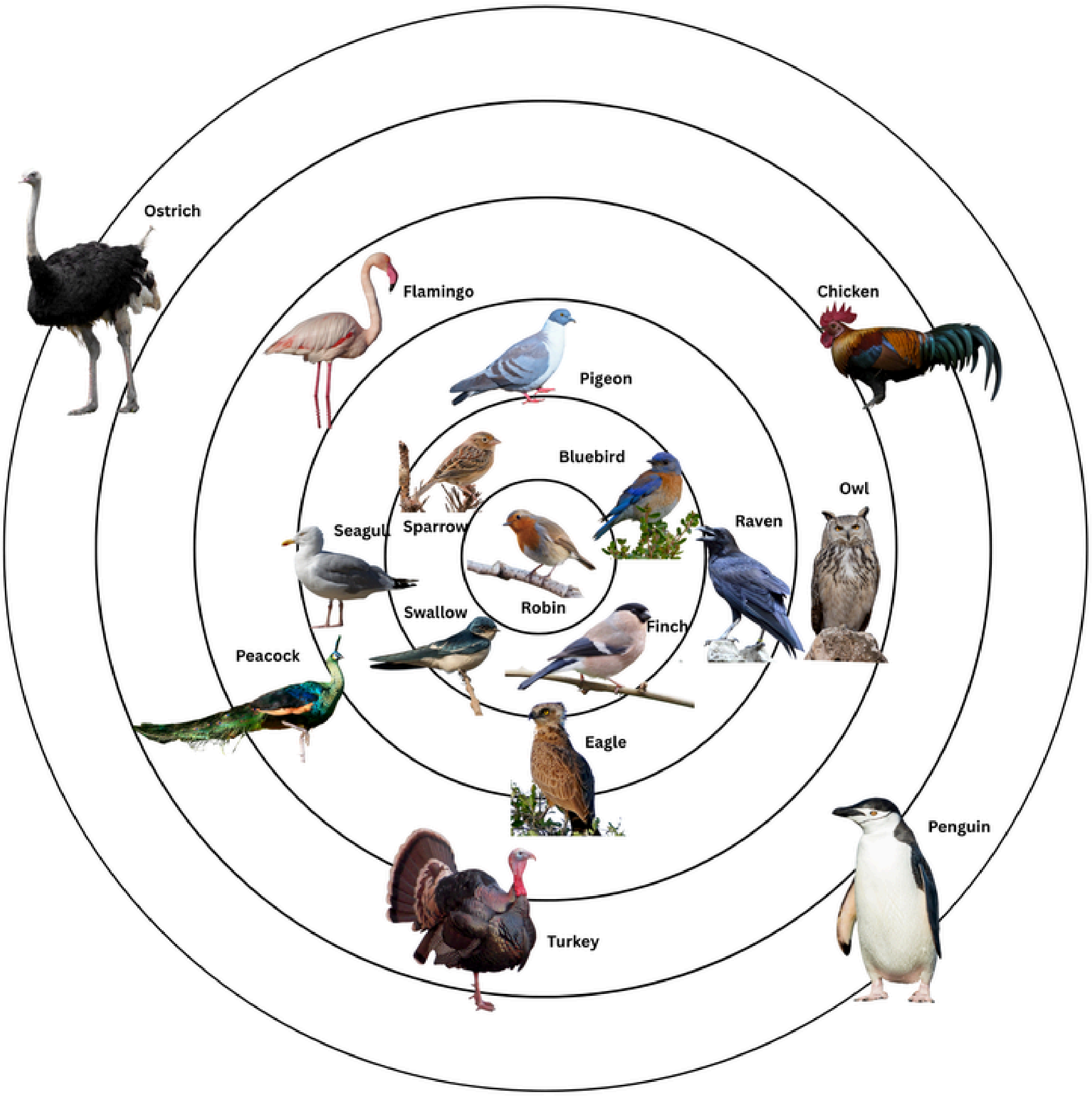

Figure 2

Birds arranged based on their rated typicality, with the most typical birds closer to the center. Figure adapted from **Aitchison (2012)** by Xiaoran Sun.

One alternative is to define typicality as the extent to which an item possesses properties that are common to the category. Robins, on this view, are typical because they have many properties shared by other birds (colorful, small, can fly, and can sing). A bird like a penguin is atypical to the extent that it lacks these properties. This operationalization works well for explaining learning and generalization of artificial stimuli composed of discrete features frequently used in categorization experiments (Figure 3).

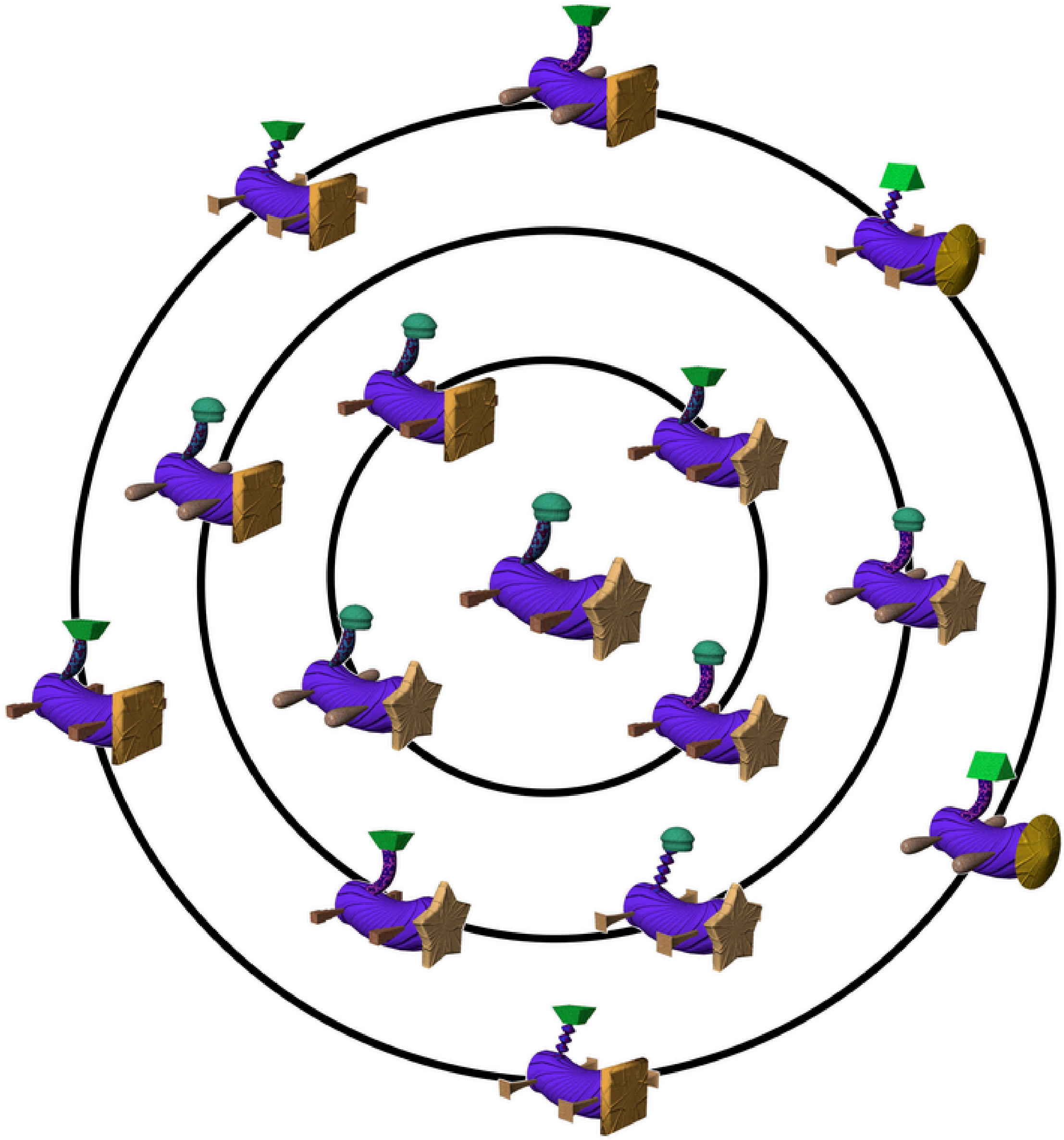

Figure 3
An example of feature-based artificial categories (Fribbles created by the tarrlab) arranged around a prototype (center) that has all category-typical features. Fribbles farther from the center share fewer features with the central prototype.

If a category is represented by its prototype, one might expect that the prototype (e.g., the triangle used in Posner & Keele, 1968) would be recognized better than previously seen but less-typical category members. Yet participants tend to classify and recognize previously seen items more accurately than the prototype itself, suggesting that people's judgments are strongly affected by memory for specific exemplars (e.g., Medin & Shoben, 1988; Nosofsky et al., 1992).

Traditionally, prototype and exemplar accounts of categorization have tended to assume that the dimensions or features along which similarity is being computed are available to the learner. Yet the feature space itself often has to be learned (see Categorical perception). A salient example is learning to read. Initially, children may represent an "M" as a specific configuration of vertical and diagonal lines. As they become experts, the M becomes a single object. People trained to categorize patterns defined by specific combinations of line segments began to see the entire configuration as one perceptual *chunk*—the configurations most diagnostic of category membership became *unitized* (Goldstone, 2000).

Unitization blurs the boundary between prototype and exemplar theories. As repeated exposure fuses co-occurring features into larger chunks, learners can preserve detailed memories for exemplars while simultaneously extracting a more abstract representation that distinguishes category members from nonmembers (Schyns et al., 1998). However, unitization alone does not explain why some combinations of features get bound together while others do not. Wings, beaks, and feathers get bound into "bird," whereas eyes, feet, and skin—features equally present in birds—do not. Something beyond mere co-occurrence must determine which features cohere into a category.

One answer as to why it seems natural to link wings, beaks, and feathers with bird-ness is that these features help distinguish birds from things they are often contrasted with, such as other animals; wings are diagnostic of being a bird, whereas having eyes is not (see Hampton, 2006 for discussion). A more general form of this idea is to consider categories as *theories* (Carey, 1985; Murphy & Medin, 1985). As children learn about the natural world—for example, that mammals share many anatomical and behavioral traits—they begin to weigh certain dimensions as more relevant than others. Categorization on this view depends on causal and functional relations among features; knowing that wings enable flight, or that engines make cars move, helps explain why some properties are more central to category membership and why others are more incidental. Thinking of categorization in terms of the deployment of learned causal models helps to make sense of why variation along some dimensions is relevant for category membership but why variation along other perceptually salient dimensions is not [see Ad Hoc Categories]. Viewing categories as theories also helps make sense of people's ability to understand that a chocolate cake is made with chocolate, but a birthday cake is not made of birthdays, as well as to understand novel compounds such as apartment dog (Murphy, 1988).

Questions, controversies, and new developments

Effects of context on categorization

The section on "Categorical perception" challenged its domain specificity and innateness, but may give the impression that categorizing perceptual inputs operates by applying rule-like decision boundaries to incoming features. For example, if the voice-onset time is less than 20 ms, categorize as /b/, otherwise as /p/. If a visual stimulus fits a face template, categorize it as a face, otherwise do not. The phenomena considered next pose a deeper challenge; they show that the features entering into such a decision—and indeed whether categorization occurs at all—depend on the surrounding context and on the demands of the current task.

Perceptual categorization is profoundly affected by context. A sound that is in between /is/ and /ish/ sounds like /is/ if preceded by "malpracti_" and like /ish/ if preceded by "aboli_." The different contexts appear to change perceptual processing of the ambiguous sound at a relatively low level as revealed by the different effects of the now disambiguated percept on subsequent speech perception (Samuel, 2001). Similar effects of context are observed in visual processing. For example, the fusiform face area, a cortical region thought to be responsible for recognizing faces, responds robustly to an image of a face but minimally to a highly blurred face or a torso by itself [see Face Perception]. Yet, when presented with a torso with a blurred face in the expected position, the fusiform face area again shows a robust response (Cox et al., 2004). These effects of context on neural responses are mirrored in behavior. A blur that is unrecognizable on its own becomes easily recognized as a person or a car depending on its position within a scene (Oliva & Torralba, 2007). Such results suggest that rather than thinking of perceptual categorization as a bottom-up process of reweighing category-relevant features, categorization—of even ecologically important stimuli like faces and speech sounds—involves rich integration of bottom-up and top-down information.

Context effects also shape similarity judgments, revealing that context shapes which features matter and how much. For example, on a simple similarity account, gray lies in between black and white. Yet people judge gray to be more similar to white than black in the context of hair but more similar to black than white in the context of clouds (Medin & Shoben, 1988). Such results show that categorizing a stimulus engages background knowledge that determines which dimensions of variation are diagnostic in the moment.

Finally, the extent to which sensory inputs are categorized in the first place appears to be strongly task dependent. Tasks designed for studying categorical perception effects have tended to require participants to make discrete responses (such as, "Do you hear a /b/ or a /p/?"). It turns out that when tasks do not require such responses, people tend to experience the sounds in a more graded fashion (Pisoni & Lazarus, 1974; see McMurray, 2022 for review). A similar pattern is observed in color perception. When the task structure encourages a categorical encoding of colors, participants tend to show robust categorical perception effects, with better discrimination of perceptual differences that span category boundaries (e.g., a blue and a green) than of psychophysically equal differences that fall within the same category (e.g., two shades of blue; Gilbert et al., 2006). In contrast, when the task calls on frequent within-category discrimination (Witzel & Gegenfurtner, 2015) or when responses require choosing colors from a continuous color space rather than choosing between different discrete color choices (Uchikawa & Shinoda, 1996), categorical color perception is diminished or eliminated entirely.

Controversies remain about the extent to which categorization affects perception and at what level. There is ample evidence that perceptual learning can shape in-the-moment visual representations at low levels, including in primary visual cortex (e.g., Li et al., 2004), and simply hearing category labels (e.g., color words like "blue" and "green") strongly affects the accuracy with which people discriminate colors immediately following the label—a result consistent with language temporarily inducing more categorical color representations (Forder & Lupyan, 2019). However, such effects on discrimination are distinct from claims about lasting changes to the perceptual system itself. Evidence that category learning permanently "warps" perceptual representations is much weaker (McMurray, 2022). One working hypothesis is that categories shape perception when task performance would benefit from deploying a more categorical representation. A task distinguishing greens from blues benefits from a more categorical color representation; a task requiring discrimination among shades of blue does not. It is rarely optimal to permanently abstract over a given dimension, lest the information be needed for another task.

The role of contrast in category learning and category use

Empirical studies of categorization have not only challenged assumptions of the Western classical view in that categories have crisp definitions but have called into question seemingly self-evident principles of logic as applied to categorization, such as the principle of the excluded middle: something is either a member of a given category, or it is not. If one learns that someone is tall, it implies that this person is not short. Yet, when shown an image like Figure 4, 45% agreed that person 2 was both tall and not tall, and 54% agreed that he was neither tall nor not tall (Alxatib & Pelletier, 2011; see Hampton, 2012 for discussion).

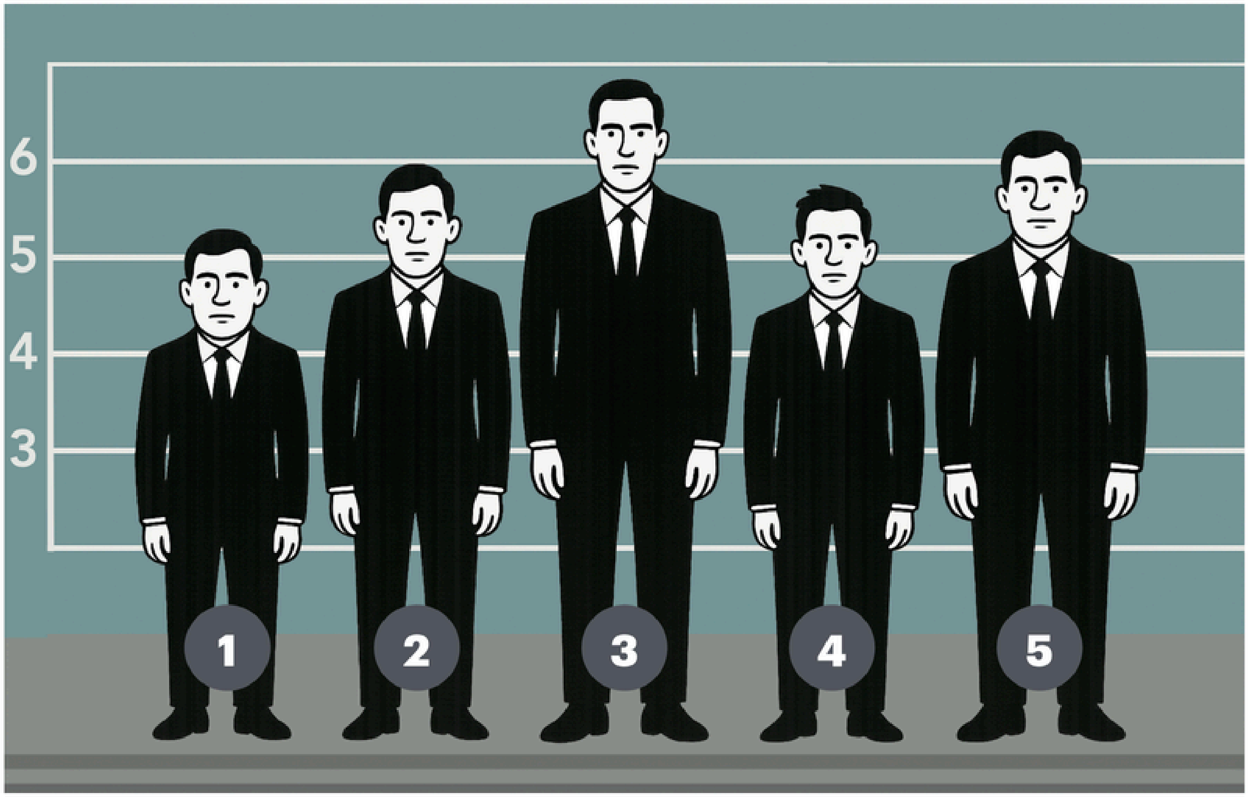

Figure 4
An adaptation of an image from **Alxatib and Pelletier (2011)**. Participants were asked whether each person was tall, not tall, tall and not tall, or neither tall nor not tall, indicating whether each option was true, false, or if they could not tell.

It is not simply that categories like *tall* are vague. Rather, their meaning is strongly dependent on the invoked contrast (tall compared to whom?). Sometimes the contrast is explicit (a "large ant" is large compared to typical ants; a "short basketball player" is short relative to typical basketball players). However, often, the contrast is more implicit and subject to rapid shifts. In Figure 4, person 2 is judged not tall (compared to person 3 and 5) but also not short (compared to person 1 and 4). People appear to readily invoke implicit contrasts in predictable ways, for example, when asked what objects are never found in a kitchen, many will respond with "toilet" and "bed" (Greene, 2016), suggesting that rather than drawing a contrast between a kitchen and the (near-infinite) set of non-kitchens, thinking of kitchens implicitly invokes an implicit contrast with the much more restricted set of non-kitchens, namely, other rooms of the house. People's use of such implicit contrasts and the role that these contrasts (what something is *not*) play in category learning is not well understood.

The relationship between categorization and language

Most words denote categories. A longstanding question in the study of categorization has been whether these categories exist in the mind prior to language. On one view, our categories are determined by the inherent structure of the world, and we attach names to categories that are important to communicate about (Gleitman & Papafragou, 2012). Alternatively, languages actively help us construct the categories. A relatively uncontroversial example of the second position is that it is through language that we learn many concepts such as "twenty," "marsupial," and "unicorn." Many categories that are, in principle, learnable through direct experience are, in practice, learned through language (Harnad, 2005; Levinson, 2003). Language is also a powerful system for making inferences. It is not perceptually obvious that whales are mammals, but learning this allows people to make multiple inferences about a whale's physiology and behavior. In short, language often provides much of the information required to create the causal theories that are then deployed in categorization.

More controversially, language has been argued to serve as a kind of "glue" for unifying perceptually disparate items into coherent categories. Already in early childhood, children use words to infer category membership and override perceptual similarities in favor of linguistic ones (Gelman & Coley, 1990; Nazzi & Gopnik, 2001; Waxman & Markow, 1995). For example, 2.5-year-old children shown a picture of a typical bird and told that it lives in a nest often generalize this property to other animals on the basis of appearance, judging that a visually similar stegosaurus also lives in a nest but an atypical bird like a dodo does not. When the pictures are labeled ("Look, this bird lives in a nest…"), children are more likely to generalize on the basis of category than appearance. Labels continue to play a role in adult category learning; labeled categories have been found to be easier to learn than unlabeled ones (Lupyan et al., 2007). A consistent finding is that verbal labels are particularly effective at amplifying the features most diagnostic of category membership, sharpening the distinction between category members and nonmembers (e.g., Lupyan, 2015).

The effects of labels on categorization appear to be particularly important in learning (and using) categories like "above," "barrier," and "carnivore," which are characterized by relations among entities (Gentner, 2005; Gentner & Loewenstein, 2002). Because such relational categories are less tightly anchored to perceptual regularities than more concrete object-based categories, verbal labels do more of the work of picking them out [see Language].

These effects of labels on categorization do not imply that labels shape how people categorize across different contexts. For example, speakers of English, Spanish, and Chinese differ in how they name containers (i.e., the generalization profiles of words like "box", "jar", "glass," and "carton"; see also Majid et al., 2018). However, these naming differences do not predict what participants do when asked to sort the containers according to their similarity (Malt et al., 1999). Differences in naming do not prevent speakers of different languages from forming similar categories in the context of other tasks.

A related question is whether people from different linguistic and cultural backgrounds differ in the kinds of categories they are drawn to forming. Some cross-cultural comparisons suggest this to be the case. East Asian participants have been reported to group objects thematically, on the basis of co-occurrence or functional relations (e.g., "cow" with "milk"), more often than Western participants, who more often group taxonomically (e.g., "cow" with "goat"; Ji et al., 2004). However, both modes of organization can be readily elicited with small changes to the materials or instructions (Lin & Murphy, 2001), suggesting that cross-cultural differences likely stem from differences in how people construe the task they are presented with.

Broader connections

Social categories and psychological essentialism

Some of our most consequential acts of categorization involve categorizing people based on gender, race, nationality, political affiliation, and social class. The very same psychological machinery that allows us to deploy sophisticated causal models rather than rely on superficial perceptual features can manifest as *psychological essentialism* (Medin & Ortony, 1989), wherein people justify category membership on the basis of unchangeable essences. This tendency toward essentialism can be reinforced by language, namely *generic statements* such as, “Boys are good at math” or “People from X are hardworking” (Gelman & Roberts, 2017). Such statements may cause people to extend causal models that are useful in some domains (e.g., natural kinds) to domains where they are less appropriate (e.g., social categories).

Categorization in the age of large language models

The study of categorization has long emphasized the importance of perceptual experience, and theories of categorization have often been grounded in certain biological mechanisms argued to underlie similarity-based versus rule-based categories (Ashby & Maddox, 2011). Yet, large language models (LLMs) like ChatGPT reproduce many classic categorization effects even though they are trained almost entirely on language and lack not only other perceptual input but any specialized machinery for abstraction or inductive inference [see Large Language Models]. For instance, LLMs reproduce human-like typicality ratings (e.g., a robin is a better “bird” than a penguin; Bhatia & Richie, 2024), show classic asymmetries in category-based induction (generalizing more from robins than penguins; Han et al., 2024), and are able to competently interpret novel conceptual combinations in ways that align with human causal theories (Ciapparelli et al., 2025). Understanding how these effects emerge in LLMs can act as a stress test on theories of human categorization.

Further reading

- Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *WIREs: Cognitive Science*, *1*(1), 69–78. https://doi.org/10.1002/wcs.26
- Hampton, J. A. (2012). Thinking intuitively: The rich (and at times illogical) world of concepts. *Current Directions in Psychological Science*, *21*(6), 398–402. https://doi.org/10.1177/0963721412457364
- Medin, D. L., & Rips, L. J. (2005). Concepts and categories: Memory, meaning, and metaphysics. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 37–72). Cambridge University Press.
- Wierzbicka, A. (2013). *Imprisoned in English: The hazards of English as a default language* (1st ed.). Oxford University Press.

References

- Alxatib, S., & Pelletier, F. J. (2011). The psychology of vagueness: Borderline cases and contradictions. *Mind & Language*, *26*(3), 287–326. https://doi.org/10.1111/j.1468-0017.2011.01419.x

  ↩
- Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161. https://doi.org/10.1111/j.1749-6632.2010.05874.x

  ↩
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *11*(4), 629–654. https://doi.org/10.1037//0278-7393.11.1-4.629

  ↩
- Bhatia, S., & Richie, R. (2024). Transformer networks of human conceptual knowledge. *Psychological Review*, *131*(1), 271–306. https://doi.org/10.1037/rev0000319

  ↩
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. Wiley.

  ↩
- Carey, S. (1985). *Conceptual change in childhood*. MIT Press.

  ↩
- Ciapparelli, M., Zarbo, C., & Marelli, M. (2025). Conceptual combination in large language models: Uncovering implicit relational interpretations in compound words with contextualized word embeddings. *Cognitive Science*, *49*(3), e70048. https://doi.org/10.1111/cogs.70048

  ↩
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*(5667), 115–117. https://doi.org/10.1126/science.1093110

  ↩
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in Infants. *Science*, *171*(3968), 303–306. https://doi.org/10.1126/science.171.3968.303

  ↩
- Forder, L., & Lupyan, G. (2019). Hearing words changes color perception: Facilitation of color discrimination by verbal and visual cues. *Journal of Experimental Psychology: General*, *148*(7), 1105–1123. https://doi.org/10.1037/xge0000560

  ↩

- Gagnon, Y. L., Templin, R. M., How, M. J., & Marshall, N. J. (2015). Circularly polarized light as a communication signal in mantis shrimps. *Current Biology*, *25*(23), 3074–3078. https://doi.org/10.1016/j.cub.2015.10.047

  ↩
- Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, *26*(5), 796–804. https://doi.org/10.1037/0012-1649.26.5.796

  ↩
- Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proceedings of the National Academy of Sciences*, *114*(30), 7900–7907. https://doi.org/10.1073/pnas.1621073114

  ↩
- Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 245–275). Lawrence Erlbaum Associates.

  ↩
- Gentner, D., & Loewenstein, J. (2002). Relational language and relational thought. In J. Byrnes & E. Amsel (Eds.), *Language, literacy, and cognitive development* (pp. 87–120). Lawrence Erlbaum Associates.

  ↩
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *Proceedings of the National Academy of Sciences*, *103*(2), 489–494. https://doi.org/10.1073/pnas.0509868103

  ↩
- Gleitman, L., & Papafragou, A. (2012). New perspectives on language and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 543–568). Oxford University Press.

  ↩
- Goldstone, R. L. (2000). Unitization during category learning. *Journal of Experimental Psychology-Human Perception and Performance*, *26*(1), 86–112. https://doi.org/10.1037//0096-1523.26.1.86

  ↩
- Greene, M. R. (2016). Estimations of object frequency are frequently overestimated. *Cognition*, *149*, 6–10. https://doi.org/10.1016/j.cognition.2015.12.011

  ↩

- Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, *51*(1), 79–88. https://doi.org/10.1037/h0046219

  ↩
- Hampton, J. A. (2006). Concepts as prototypes. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 46, pp. 79–113). Elsevier.

  ↩
- Hampton, J. A. (2012). Thinking intuitively: The rich (and at times illogical) world of concepts. *Current Directions in Psychological Science*, *21*(6), 398–402. https://doi.org/10.1177/0963721412457364

  ↩
- Han, S. J., Ransom, K. J., Perfors, A., & Kemp, C. (2024). Inductive reasoning in humans and large language models. *Cognitive Systems Research*, *83*, 101155. https://doi.org/10.1016/j.cogsys.2023.101155

  ↩
- Harnad, S. (2005). Cognition is categorization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 20–45). Elsevier Science.

  ↩
- James, W. (1890). *Principles of psychology*. Holt.

  ↩
- Ji, L.-J., Zhang, Z., & Nisbett, R. E. (2004). Is It culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, *87*(1), 57–65. https://doi.org/10.1037/0022-3514.87.1.57

  ↩
- Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, *190*(4209), 69–72. https://doi.org/10.1126/science.1166301

  ↩
- Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Bailey & R. W. Shuys (Eds.), *New ways of analyzing variation in English* (pp. 340–373). Georgetown University Press.

  ↩
- Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE*, *47*(11), 1940–1951. https://doi.org/10.1109/JRPROC.1959.287207

  ↩

- Levinson, S. C. (2003). Language and mind: Let's get the issues straight! In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 25–46). MIT Press.

↩

- Li, W., Piech, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature Neuroscience*, *7*(6), 651–657. https://doi.org/10.1038/nn1255

↩

- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358–368. https://doi.org/10.1037/h0044417

↩

- Lin, L. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, *130*(1), 3–28. https://doi.org/10.1037/0096-3445.130.1.3

↩

- Lupyan, G. (2013). The difficulties of executing simple algorithms: Why brains make mistakes computers don't. *Cognition*, *129*(3), 615–636. https://doi.org/10.1016/j.cognition.2013.08.015

↩

- Lupyan, G. (2015). The paradox of the universal triangle: Concepts, language, and prototypes. *Quarterly Journal of Experimental Psychology*, *70*(3), 389–412. https://doi.org/10.1080/17470218.2015.1130730

↩

- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Labels facilitate learning of novel categories. *Psychological Science*, *18*(12), 1077–1082. https://doi.org/10.1111/j.1467-9280.2007.02028.x

↩

- Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., O'Grady, L., Woll, B., LeLan, B., Sousa, H. de, Cansler, B. L., Shayan, S., Vos, C. de, Senft, G., Enfield, N. J., Razak, R. A., Fedden, S., Tufvesson, S., Dingemanse, M., Ozturk, O., ... Levinson, S. C. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences*, *115*(45), 11369–11376. https://doi.org/10.1073/pnas.1720419115

↩

- Malt, B. C., Sloman, S. A., Gennari, S., Shi, M., & Wang, Y. (1999). Knowing versus naming: Similarity and the linguistic categorization of artifacts. *Journal of Memory and Language*, *40*, 230–262. https://doi.org/10.1006/jmla.1998.2593

↩

- McMurray, B. (2022). The myth of categorical perception. *The Journal of the Acoustical Society of America, 152*(6), 3819–3842. https://doi.org/10.1121/10.0016614

  ↩
- Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review*, *111*(4), 960–983. https://doi.org/10.1037/0033-295X.111.4.960

  ↩
- Medin, D. L., & Ortony, A. (1989). Comments on part I: Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). Cambridge University Press.

  ↩
- Medin, D. L., & Shoben, E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*(2), 158–190. https://doi.org/10.1016/0010-0285(88)90018-7

  ↩
- Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. Yale University Press.

  ↩
- Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science*, *12*(4), 529–562. https://doi.org/10.1207/s15516709cog1204_2

  ↩
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.

  ↩
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*(3), 289–316. https://doi.org/10.1037/0033-295X.92.3.289

  ↩
- Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition*, *80*(3), B11–B20. https://doi.org/10.1016/s0010-0277(01)00112-3

  ↩
- Nosofsky, R. M., Kruschke, J. K., & McKinley, S. C. (1992). Combining exemplar-based category representations and connectionist learning rules. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 211–233. https://doi.org/10.1037/0278-7393.18.2.211

  ↩

- Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences*, *11*(12), 520–527. https://doi.org/10.1016/j.tics.2007.09.009

  ↩

- Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*(2), 185–200. https://doi.org/10.1037/0033-295X.97.2.185

  ↩

- Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America*, *55*(2), 328–333. https://doi.org/10.1121/1.1914506

  ↩

- Posner, M. I., & Keele, S. W. (1968). On genesis of abstract ideas. *Journal of Experimental Psychology*, *77*(3 Pt. 1), 353–363. https://doi.org/10.1037/h0025953

  ↩

- Ramscar, M., & Hahn, U. (1998). What family resemblances are not: The continuing relevance of Wittgenstein to the study of concepts and categories. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 865–870). Lawrence Erlbaum.

  ↩

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

  ↩

- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

  ↩

- Samuel, A. G. (2001). Knowing a word affects the fundamental perception of the sounds within it. *Psychological Science*, *12*(4), 348–351. https://doi.org/10.1111/1467-9280.00364

  ↩

- Schyns, P. G., Goldstone, R. L., & Thibaut, J. P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences*, *21*(1), 1–54. https://doi.org/10.1017/s0140525x98000107

  ↩

- Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, *439*(7079), 978–982. https://doi.org/10.1038/nature04485

  ↩

- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*(3), 457–482. https://doi.org/10.1016/0010-0285(91)90016-H

↩

- Uchikawa, K., & Shinoda, H. (1996). Influence of basic color categories on color memory discrimination. *Color Research & Application*, *21*(6), 430–439. https://doi.org/10.1002/(SICI)1520-6378(199612)21:6%253C430::AID-COL5%253E3.0.CO;2-X

↩

- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, *3*(1), 1–14. https://doi.org/10.1037/h0069608

↩

- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016

↩

- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49–63. https://doi.org/10.1016/S0163-6383(84)80022-3

↩

- Witzel, C., & Gegenfurtner, K. R. (2015). Categorical facilitation with equally discriminable colors. *Journal of Vision*, *15*(8), 22. https://doi.org/10.1167/15.8.22

↩