Cognitive and Neural Contributions to Understanding the Conceptual System

Lawrence W. Barsalou

Current Directions in Psychological Science 2008 17: 91
DOI: 10.1111/j.1467-8721.2008.00555.x

Cognitive and Neural Contributions to Understanding the Conceptual System

Lawrence W. Barsalou

Emory University

ABSTRACT—*The conceptual system contains categorical knowledge about experience that supports the spectrum of cognitive processes. Cognitive science theories assume that categorical knowledge resides in a modular and amodal semantic memory, whereas neuroscience theories assume that categorical knowledge is grounded in the brain's modal systems for perception, action, and affect. Neuroscience has influenced theories of the conceptual system by stressing principles of neural processing in neural networks and by motivating grounded theories of cognition, which propose that simulations of experience represent knowledge. Cognitive science has influenced theories of the conceptual system by documenting conceptual phenomena and symbolic operations that must be grounded in the brain. Significant progress in understanding the conceptual system is most likely to occur if cognitive and neural approaches achieve successful integration.*

KEYWORDS—*categories; conceptual system; knowledge; semantic memory*

Most generally, the conceptual system is an extensive system, distributed throughout the brain, that represents knowledge about all aspects of experience—including settings, events, objects, agents, actions, affective states, and mental states. Much research demonstrates that this knowledge is organized categorically. The brain is not like a camera or video recorder that captures holistic images of experience in which the components of experience are undifferentiated. Instead, the brain contains a powerful attentional system that focuses on individual components of experience, establishing categorical knowledge about them. For example, knowledge about the category *chairs* develops from focusing attention on chairs across experiences, extracting information about them, and integrating it. Categorical knowledge about other aspects of experience develops similarly. As people focus attention on the action of sitting, or on the affective state of happiness, they develop knowledge about these categories as well. By focusing attention on complex arrays of information in experience, people also develop more complex concepts—such as concepts for spatial relations (e.g., *above*), physical events (e.g., *carry*), and social events (e.g., *convince*).

Once the conceptual system is in place, it supports the broad spectrum of cognitive activities (e.g., Murphy, 2002). There is probably no such thing as a knowledge-free cognitive process. As people interact with the environment during online processing, knowledge in the conceptual system plays extensive roles. It supports perception, providing knowledge that completes perceptions and that generates anticipatory inferences; it makes the categorization of settings, events, objects, agents, actions, and mental states possible; and it provides rich inferences about categorized entities that go beyond the information perceived to support goal pursuit. The conceptual system also plays a central role when people cognize about situations not present, during offline processing. It supports the cuing and reconstruction of past events from memory; it contributes extensively to the meanings of words and sentences during language use; and it provides the representations on which thought operates during decision making, problem solving, and reasoning. The conceptual system is central to learning and development too. On encountering novel entities in a new domain, existing categorical knowledge is used to interpret them. As expertise in the domain develops, new categories that interpret the domain with greater sophistication develop, thereby expanding the conceptual system. Finally, the conceptual system is central to social cognition, playing central roles in categorizing social entities and events, in drawing social inferences, and in planning and remembering social interactions.

Address correspondence to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, GA 30322; e-mail: barsalou@emory.edu.

THE DOMINANT THEORY IN COGNITIVE SCIENCE

Across diverse areas of psychology, computer science, linguistics, and philosophy, the dominant account of the conceptual system is the theory of *semantic memory* (e.g., Smith, 1978). According to this theory, the conceptual system is a modular memory store that contains amodal knowledge about categories. Semantic memory is viewed as modular because it is assumed to be separate from the brain's episodic-memory system and also from the brain's modal systems for perception, action, and affect. Because semantic memory lies outside modal systems, its representations are viewed as different from theirs, providing a higher, amodal level of representation.

The *transduction principle* underlies the view that amodal representations develop for categories in a modular conceptual system. According to this principle, modal representations in the brain's systems for perception, action, and affect become active during experience with the world. In turn, the brain transduces these modal representations into amodal representations that represent category knowledge in a modular semantic memory. When encountering dogs, for example, modal representations arise as dogs are seen, heard, and touched. In turn, these modal representations are transduced into amodal symbols that stand for these experiences. For example, symbols become established in categorical knowledge for *dogs* that represent *fur*, *barking*, and *pat*. Later, when categorical knowledge represents dogs in offline processing, amodal symbols are retrieved to represent the category. Notably, none of the original modal representations are assumed to become active or to play any role in the representational process.

The idea that knowledge is represented outside the brain's modal systems underlies many theories of knowledge across the cognitive sciences. Although theories differ in whether they assume that rules, prototypes, schemata, or exemplars represent categories, they typically assume implicitly that amodal symbols represent these knowledge structures. Interestingly, this view was adopted largely for theoretical reasons associated with the cognitive revolution rather than because of extensive empirical evidence for the presence of amodal symbols in the brain (Barsalou, 1999).

THE DOMINANT THEORY IN COGNITIVE NEUROSCIENCE

A very different view of the conceptual system has arisen in cognitive neuroscience. According to this view, categorical knowledge is grounded in the brain's modal systems rather than being represented amodally in a modular semantic memory (e.g., Martin, 2001). For example, knowledge about *dogs* is represented in visual representations of how dogs look, in auditory representations of how dogs sound, and in motor representations of how to interact with dogs. Because the representational systems that underlie perception, action, and affect are also used to represent categorical knowledge, the conceptual system is neither modular nor amodal. Instead, perception and conception share overlapping systems.

Empirical evidence has been the driving force behind this view. In neuropsychology, lesions to the brain's modal systems produce deficits in category knowledge, suggesting that modal systems represent this knowledge, at least to some extent (e.g., Martin & Caramazza, 2003). Lesions to visual areas, for example, produce deficits in categories that rely heavily on visual processing (e.g., *animals*), whereas lesions to motor areas produce deficits in categories that rely heavily on actions (e.g., *tools*). In neuroimaging, processing particular categories activates associated modal areas (e.g., Martin, 2007). For example, processing *animals* activates brain areas for visual form and animate motion, whereas processing *tools* activates brain areas for action and inanimate motion.

Based on this evidence, neuroscience researchers increasingly view the representation of a category as a neural circuit distributed across the relevant modalities (e.g., Cree & McRae, 2003). Depending on the modalities relevant for interacting with a category, a corresponding circuit becomes active across modalities to represent it during conceptual processing.

NEUROSCIENCE CONTRIBUTIONS TO UNDERSTANDING THE CONCEPTUAL SYSTEM

Across different communities of researchers, the impact of neuroscience on understanding the conceptual system ranges from no impact to profound impact. As a result, tension exists between areas, along with considerable potential for rapid change in theory and research.

Impact on Cognitive Science Theories

Empirical findings and theories from neuroscience have had relatively little, if any, impact on theories of the conceptual system in mainstream cognitive science. Most behavioral researchers continue to implicitly adopt the semantic memory view in designing experiments and explaining empirical findings. Many researchers continue to believe that the brain does not provide useful constraints on cognitive theory and that cognitive theory should abstract over low-level implementation details such as neural mechanisms. Consequently, these researchers continue to assume that amodal knowledge in a modular semantic memory represents categories.

As theories of cognition become increasingly grounded in the brain, however, it will become increasingly necessary to explain how semantic memory theories of knowledge can be reconciled with the brain's anatomical and physiological properties. For example, where do amodal representations reside outside the brain's modal areas? What neural evidence supports the presence of amodal symbols? What neural principles explain how these symbols are processed as people perform categorization, draw categorical inferences, and combine simple concepts to form complex concepts? It is unlikely that the field will allow cognitive theories of knowledge to remain divorced from the brain indefinitely. Integration must occur eventually, and a key question is whether current theories will integrate successfully. At a minimum, neuroscience is likely to influence these theories by forcing them to become integrated with the brain. To the extent that problems arise during integration, constraints from neuroscience are likely to reshape these theories significantly.

Another major source of constraint is evidence from neuropsychology and neuroimaging. Ultimately, theories of the conceptual system must be compatible with empirical results from these areas. To the extent that existing theories cannot explain these findings, modification of these theories is likely.

Impact on Neural Network Theories

Not all researchers have been comfortable with the idea that theories of knowledge can ignore the brain. One source of dissatisfaction is the assumption that discrete amodal symbols underlie knowledge and that classic logical and computational operations manipulate these discrete symbols in the brain (e.g., Bechtel & Abrahamsen, 2002). In reaction to these views, large communities of researchers have developed computational accounts of cognition that rest upon principles of neural processing (e.g., O'Reilly & Munakata, 2000). Although these theories typically use idealized neural mechanisms that differ from actual neurons, they nevertheless assume that large populations of neuron-like units represent knowledge in a distributed manner, using simple activation and learning mechanisms that are neurally plausible. Such theories are used widely to motivate experiments and explain empirical results (e.g., Rogers & McClelland, 2004).

Although neural-network theories have not been developed to explain all cognitive processes, they have been applied to many cognitive processes successfully and with insight. Many researchers assume that it is simply a matter of time (and computational hardware) before neural network theories exist for most cognitive processes. In this sense, basic assumptions of neuroscience increasingly pervade computational theories. This is not to say that non-neural theories are completely incorrect or capture nothing important about cognition. Indeed, a likely scenario is that mechanisms in non-neural theories are likely to remain present to a significant extent in neural theories but will be implemented with neural mechanisms.

Impact on Grounded Theories

Another source of dissatisfaction with semantic memory theories is their modular status in the brain. On the one hand, it has not been clear how amodal representations in these theories interface with perception, action, and affect. On the other hand, neuroscience evidence strongly suggests that modal systems play central roles in representing categorical knowledge, such that it is neither modular nor amodal.

Grounded theories of the conceptual system address these concerns (e.g., Barsalou, 1999; Damasio, 1989). In grounded theories, modal states that arise during interactions with categories are captured and integrated in memory. Later when these categories are represented conceptually, previously captured modal states are partially reactivated to represent them. In other words, the brain attempts to simulate the states it was in while interacting with category members. To represent *dogs*, for example, the brain simulates visual, auditory, tactile, motor, and affective states experienced previously with the category.

Much accumulating evidence supports this view (e.g., Barsalou, 2008). As described earlier, findings in neuropsychology and neuroimaging results offer support. Much further support across cognitive, developmental, and social psychology exists as well (e.g., Barsalou, Niedenthal, Barbey, & Ruppert, 2003). Grounded theories offer another example of how neuroscience has produced considerable impact on theories of knowledge. Mechanisms in these theories are grounded both in brain architecture and in principles of neural processing. Empirical research that assesses these theories often makes predictions about neural activation or makes assumptions about the role of neural systems in producing behavior.

COGNITIVE SCIENCE CONTRIBUTIONS TO UNDERSTANDING THE CONCEPTUAL SYSTEM

In general, cognitive science has had considerable impact in shaping cognitive neuroscience. Often the mechanisms that neuroscientists attempt to identify in the brain were first identified in behavioral research (e.g., selective attention, working memory, repetition priming). The study of knowledge and conceptual processing is no exception. It is safe to say that behavioral research from cognitive science will continue to shape neuroscience research on the conceptual system for decades to come.

Empirical Phenomena

The behavioral literature on the conceptual system documents many phenomena that must be grounded in the brain (e.g., Murphy, 2002). These phenomena include category learning, typicality, the basic level, inductive inference, predication, conceptual combination, and many others. Rich literatures exist that describe these phenomena, along with the variables that affect them, in detail. Although the neural bases of these phenomena have generally not been established, they appear central to human cognition, given their many demonstrations under diverse conditions. Furthermore, these phenomena often arise in other research areas such as psycholinguistics, development, social interaction, education, clinical psychology, and so on. Thus, these phenomena appear central to the conceptual system.

Cognitive neuroscientists have barely scratched the surface in addressing these phenomena. Nearly all neuroscience work to date has simply addressed the question of where category representations reside in the brain. Specifically, most neuroscience research has sought to identify where particular types of properties associated with categories are stored (e.g., Martin, 2007) and how simple categories are learned (e.g., Ashby & Maddox, 2005). These are indeed basic phenomena. Nevertheless, many other important phenomena documented in the behavioral literature must be grounded in the brain as well. Both cognitive scientists and neuroscientists have their work cut out for them in achieving this integration. Much is likely to be learned about these phenomena by understanding how they are realized in the brain, contrary to the widespread assumption in cognitive science that the brain is irrelevant.

Symbolic Operations

Although skepticism that discrete amodal symbols underlie conceptual processing in the brain continues to increase, there is little doubt that the brain is a symbolic system. Unlike cameras and video recorders, the brain uses categorical knowledge to interpret regions of experience that contain agents, objects, actions, mental states, and so forth. The brain does not achieve its powerful forms of intelligence by processing holistic images.

Cognitive science theories have identified a common set of symbolic operations that occur ubiquitously across cognitive processes. These operations include the binding of types (categories) to tokens (individuals) during categorization and language processing, the extension of inferences from category knowledge to individuals, the combination of concepts during language and thought, the use of conceptual relations to integrate concepts, and the hierarchical embedding of conceptual structures to produce propositions. Not only have cognitive scientists observed, documented, and manipulated these processes throughout higher cognition, linguists and knowledge engineers rely on them to construct computational accounts of higher cognitive processes. Without these operations, it is impossible to model or implement intelligence in its full power and complexity.

Almost exclusively, cognitive scientists have assumed that discrete amodal symbols underlie this core set of symbolic operations. As the neural implausibility of such symbols increases, however, it becomes increasingly important to identify alternative ways that the brain could implement symbolic operations. We will not understand higher cognition until we understand how symbolic operations arise in the brain to interpret experience. Grounded theories offer one account of how the brain could implement symbolic operations (e.g., Barsalou, 1999, 2005).

CONCLUSION

Significant progress in understanding the conceptual system is most likely to occur if cognitive and neural approaches achieve serious integration. Cognitive science approaches are likely to continue down a path of neural implausibility if they do not incorporate neural principles extensively. Neuroscience approaches can only grow in sophistication if they ground well-established empirical phenomena and core symbolic operations in neural mechanisms.

Recommended Reading

Ashby, F.G., Maddox, W.T. (2005). (See References). Reviews empirical findings and theories associated with category learning.

Barsalou, L.W. (2008). (See References). Reviews empirical findings and theories associated with the grounding of categorical knowledge in the brain's modal systems for perception, action, and introspection.

Martin, A. (2007). (See References). Reviews empirical findings and theories associated with the grounding of object categories in neural mechanisms.

Murphy, G.L. (2002). (See References). Provides an extensive review across many literatures of behavioral findings and theories associated with concepts and categories.

O'Reilly, R.C., & Munakata, Y. (2000). (See References). Provides thorough treatment of neural networks used to model category learning and processing.

Acknowledgments—This work was supported by National Science Foundation Grant BCS-0212134 and by Defense Advanced Research Projects Agency contract BICA FA8650-05-C-7256 to Lawrence Barsalou.

REFERENCES

Ashby, F.G., & Maddox, W.T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149–178.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Barsalou, L.W. (2005). Abstraction as dynamic interpretation in perceptual symbol systems. In L. Gershkoff-Stowe & D. Rakison (Eds.), *Building object categories in developmental time* (pp. 389–431). Mahwah, NJ: Erlbaum.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Barsalou, L.W., Niedenthal, P.M., Barbey, A., & Ruppert, J. (2003). Social embodiment. In B. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 43, pp. 43–92). San Diego: Academic Press.

Bechtel, W., & Abrahamsen, A. (2002). *Connectionism and the mind: Parallel processing, dynamics, and evolution in networks* (2nd Ed.). Oxford, UK: Basil Blackwell.

Cree, G.S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.

Damasio, A.R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25–62.

Martin, A. (2001). Functional neuroimaging of semantic memory. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional* *neuroimaging of cognition* (pp. 153–186). Cambridge, MA: MIT Press.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.

Martin, A., & Caramazza, A. (2003). Special issue: Neuropsychological and neuroimaging perspectives on conceptual knowledge. *Cognitive Neuropsychology*, *20*, 195–587.

Murphy, G.L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

O'Reilly, R.C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

Rogers, T.T., & McClelland, J.L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.

Smith, E.E. (1978). Theories of semantic memory. In W.K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6). Hillsdale, NJ: Erlbaum.