Available online at www.sciencedirect.com

ScienceDirect

Neuroscience and Biobehavioral Reviews 32 (2008) 193–196

NEUROSCIENCE AND BIOBEHAVIORAL REVIEWS

www.elsevier.com/locate/neubiorev

Editorial

The cognitive neuroscience of category learning

The study of category learning has been a central paradigm within cognitive psychology for over 25 years. Cognitive neuroscientists have been drawn to this paradigm for several reasons: first, there is a large body of pre-existing empirical and theoretical analyses of category learning. Neuropsychological studies of clinical populations and neuroimaging of healthy subjects provide insights into the cognitive neuroscience of category learning. Second, category learning has aspects of both elementary associative learning as well as higher-order cognition. On one hand, category learning can be viewed as a ‘‘cognitive skill’’ that shares behavioral properties, and possibly some neural substrates, with motor-skill learning and conditioning. On the other hand, categorization underlies many higher-order cognitive abilities. When a connoisseur distinguishes a cabernet from a merlot, when a doctor recognizes that a patient’s symptoms are due to a particular disease, or when a weather forecaster uses today’s barometric pressure to predict tomorrow’s weather, they are all doing complex categorization. It is this dual nature—part elementary skill, part higher cognition—which helps make category learning a valuable paradigm for studying the fundamental aspects of human learning at both the behavioral and neural levels of analysis.

In a typical category learning experiment, subjects are presented with stimuli that vary along several different dimensions. Subjects are usually required to classify each item into one of a number of contrasting categories. For example, based on a list of symptoms (features) a subject may be asked to make a disease diagnosis (categorization). Although subjects may initially have to guess, they can eventually learn which features and patterns are more likely to fall into one category versus another based on trial-by-trial feedback. Building on an extensive base of empirical data from psychological studies of category learning over the last 30 years, mathematical modelers have developed rigorous quantitative models of category learning that account for a wide range of behavioral phenomena and predict novel psychological phenomena. Building on these behavioral models, cognitive neuroscientists have begun to explore and exploit the richness of category learning as a tool for increasing our understanding of the brain mechanisms of cognition.

Work on the cognitive neuroscience of category learning has already yielded many informative results. For example, neuropsychology studies have shown that many forms of category learning are impaired in patients with dysfunctional basal ganglia or frontal lobes. This pattern of data, mimicking what is often found with motor skills, provides some support for viewing some category learning as a cognitive skill. Other recent and ongoing neuropsychological studies are examining how category learning is affected by schizophrenia, obsessive–compulsive disorder (OCD), attention deficit disorder (ADD), and a host of other mental health disorders. Parallel to these neuropsychological studies, human brain imaging studies of category learning have identified several key brain regions critical to category learning, including the medial temporal lobes, basal ganglia, and prefrontal cortex.

Mathematical models provide a third line of research. Some researchers have used these models to analyze the nature of the underlying deficits in category learning to develop detailed hypotheses about the functional role of the basal ganglia and medial temporal lobes. Other researchers have used neural network models of these brain structures to begin to understand their contribution to category learning.

The cognitive neuroscience of category learning is a new and exciting interdisciplinary field with the potential to meld theoretical and behavioral analyses from cognitive psychology with techniques from cognitive neuroscience.

The Cognitive Neuroscience of Category Learning Collaborative Consortium

In the spring of 2002, the J.S. McDonnell Foundation awarded Rutgers University—Newark a three-year grant to create a multidisciplinary collaborative consortium of researchers working in the Cognitive Neuroscience of Category Learning. The consortium was directed by an Executive Committee, which included Robert Nosofsky (Indiana), Paul Reber (Northwestern), Russell Poldrack (UCLA), and Mark Gluck (Rutgers-Newark), the principal investigator. The consortium hosted three annual two-day workshops on the *Cognitive Neuroscience of Category Learning*, each of which was attended by approximately 50–70 people. Proceedings from the meetings, including

doi:10.1016/j.neubiorev.2007.11.002

copies of all talks and posters, were distributed each year to workshop attendees and to approximately 150 additional researchers and students in the field. In addition to the annual workshops, the consortium supported interdisciplinary summer rotations of students between laboratories, pilot collaborative projects, shared neuropsychological patient access, and collaborative travel expenses. The conference and other programs were expanded through supplemental matching support from the *National Science Foundation* and the *National Institute of Mental Health.*

With the conclusion of this three-year consortium, this special issue of Neuroscience and Biobehavioral Reviews seeks to present the broader scientific community with a collection of its highlights, emphasizing new multidisciplinary collaborations and research which emerged from this three-year program.

Our annual workshops provided a unique venue for bringing together researchers with a broad range of approaches, including animal and human researchers, clinical and experimental researchers, and psychologists and cognitive neuroscientists. What made these meetings unique was not just the broad interdisciplinary backgrounds of the participants, but the narrow and clear focus of the behavioral paradigms which they are studying—this combination of divergent methods and convergent paradigms—made these meetings especially valuable in moving forward research on the cognitive neuroscience of learning and memory.

Overview of this special issue

The nine articles included in this issue were selected to represent a broad swath of multidisciplinary approaches including both animal and human research, theoretical and empirical studies, drawing on data from behavioral, clinical, and brain imaging studies. Given the size and page limits of the journal, it is not possible to include a comprehensive survey of the state-of-field nor to include everyone who has made important contributions to the field over the last years. For this reason, we have to use the issue to highlight "selected topics" in the Cognitive Neuroscience of Category Learning, particularly projects involving new collaborations that emerged during the three-year consortium, and which are cross-laboratory and interdisciplinary, especially those using multiple methodologies and approaches.

In the first article, "Category learning and the memory systems debate," Russell Poldrack and Karin Foerde describe a substantial and growing body of evidence from cognitive neuroscience which supports the concept of multiple-memory systems. However, some theorists have questioned the existence of multiple systems and, instead, propose that dissociations can be accounted for within a single-memory system. Poldrack and Foerde present evidence from neuroimaging and neuropsychological studies of category learning in favor of the existence of multiple-memory systems for category learning and declarative knowledge, whereas single-system theorists have argued that their approach is more parsimonious because it only postulates a single form of memory representation. They argue that the multiple-memory systems approach is superior in its ability to account for a broad range of data from psychology and neuroscience.

Szabolcs Kéri, in his article on "Interactive memory systems and category learning in schizophrenia," reviews evidence suggesting that information is processed and represented by multiple-interacting-memory systems in the brain, including prefrontal cortex, basal ganglia, and medial temporal lobe, all of which are implicated in the pathophysiology of schizophrenia. Executive and declarative memory dysfunctions are well known in schizophrenia. However, there are also habit-learning deficits in schizophrenia which, Keri argues, suggest that dopaminergic and other neurochemical processes in the basal ganglia may play a crucial role in the pathophysiology and pharmacology of schizophrenia. In this article, Keri shows how the investigation of different classification learning functions, including reward- and feedback-guided learning and acquired equivalence learning, may shed light on the neuropsychology, pathophysiology, pharmacology, and behavioral genetics of schizophrenia.

In "Basal ganglia and dopamine contributions to probabilistic category learning," Daphna Shohamy, Catherine Myers, Jeevan Kalanithi, and Mark Gluck review behavioral, neuropsychological, functional neuroimaging, and computational studies of basal ganglia and dopamine contributions to learning in humans. The authors argue that, collectively, these studies implicate the basal ganglia in incremental, feedback-based learning that involves integrating information across multiple experiences. The medial temporal lobes, by contrast, contribute to rapid encoding of relations between stimuli and support flexible generalization of learning to novel contexts and stimuli. Shohamy and colleagues argue that by breaking down our understanding of the cognitive and neural mechanisms contributing to different aspects of learning, new studies are providing insight into how, and when, these different processes support learning, how they may interact with each other, and the consequence of different forms of learning for the representation of knowledge.

The fourth article, by Martijn Meeter, G. Radics, Catherine Myers, Mark Gluck, and Ramona Hopkins, is entitled, "Probabilistic categorization: How do normal participants and amnesic patients do it?" The authors compare and review the evidence for two alternative conceptualizations of learning in probabilistic categorization: as rule-based learning, or as incremental learning. Each conceptualization forms a different basis for analyzing performance: strategy analysis assumes rule-based learning, while rolling regression analysis implies incremental learning. In this article, Meeter and colleagues contrast the ability of each to predict performance of normal categorizers. Based on their analyses, they argue that the medial temporal lobe helps set up representations, which are then used by other regions to assign patterns to categories.

Ed Smith and Murray Grossman, in their article on "Multiple systems of category learning," review neuropsychological and neuroimaging evidence for the existence of three qualitatively different categorization systems. These categorization systems are themselves based on three distinct memory systems: working memory (WM), explicit long-term memory (explicit LTM), and implicit long-term memory (implicit LTM). They first contrast categorization based on WM with that based on explicit LTM. Neuroimaging studies show differences between brain activity in normal participants as a function of whether they are instructed to categorize novel test items by rule or by similarity to known category members. Studies with neurological patients in the same paradigms provide converging evidence. Their second contrast is between categorization based on explicit LTM with that based on implicit LTM. Neuropsychological studies with patients with medial temporal lobe damage show that these patients are impaired on tasks requiring explicit LTM, but perform relatively normal on an implicit categorization task. Neuroimaging studies provide converging evidence: explicit categorization is mediated by activation in numerous frontal and parietal areas, whereas implicit categorization is mediated by a deactivation in posterior cortex.

In "How do the basal ganglia contribute to categorization? Their role in generalization, response selection, and learning via feedback" Carol Seger examines how independent cortico-striatal loops linking basal ganglia with cerebral cortex contribute to visual categorization. The first aspect of categorization discussed is the role of the visual cortico-striatal loop, which connects the visual cortex to the body and tail of the caudate, in mapping visual stimuli to categories, including evaluating the degree to which this loop may generalize across individual category members. The second aspect of categorization discussed is the selection of appropriate actions or behaviors on the basis of category membership, and the role of the visual cortico-striatal loop output and the motor cortico-striatal loop, which connects motor planning areas with the putamen, in action selection. The third aspect of categorization discussed is how categories are learned with the aid of feedback linked dopaminergic projections to the basal ganglia.

In our seventh article, Emi Nomura and Paul Reber provide "A review of medial temporal lobe and caudate contributions to visual category learning" in which they summarize recent functional neuroimaging, neuropsychological and behavioral studies examining the role of the medial temporal lobe (MTL) and the caudate in learning visual categories either by verbalizable rules or without awareness. They show how the MTL and caudate are found to play dissociable roles in different types of category learning with successful rule-based categorization depending selectively on the MTL and nonverbalizeable information-integration category learning depending on the posterior caudate. These studies utilize a combination of experimental cognitive psychology, mathematical modeling, and cognitive computational modeling to examine the roles of the MTL and caudate in visual category learning.

"Learning strategies in amnesia," by Maarten Speekenbrink, Shelley Channon, and David R. Shanks, presents a new study addressing the question of whether or not there are separate implicit and explicit memory systems. Using a modeling-based approach, they argue that learning is indistinguishable between an amnesic and control group in their task. However, in contrast to earlier findings, they found that explicit knowledge of the task structure was good in both the amnesic and the control group. This, they argue, is inconsistent with a crucial prediction from the multiple-systems account. The results can be better explained, they claim, from a single-system account where previously found differences in later categorization performance can be accounted for by a difference in learning rate.

In our ninth, and last article, David Freedman and Earl Miller present "Neural Mechanisms of Visual Categorization: Insights from Neurophysiology." How, they ask, does the brain recognize the meaning of sensory stimuli? Although much is known about how the brain processes and encodes basic visual features (e.g. color, orientation, and motion direction), much less is known about how the brain learns and represents the behavioral relevance, or category, of stimuli. Their article reviews a number of recent experiments which suggest that neuronal activity in primate prefrontal, temporal, and parietal cortical areas likely plays significant, though complementary, roles in visual categorization and category learning.

Future directions and key open issues

What next? One facet of the maturation of this interdisciplinary area of inquiry in the past few years has been the emergence of a number of focused questions, which drive collective and collaborative research agendas. As a modest contribution towards encouraging future research, we point to the following key open questions as opportunities for pushing the field even further. Perhaps some amongst our readers, or their students, may find these appealing as future research topics:

1. How can behavioral analyses and mathematical theories from cognitive psychology further inform neuroimaging and neuropsychological studies of category learning?
2. What are the relative strengths or weaknesses of neuroimaging and neuropsychological studies of clinical populations as tools for understanding brain bases of category learning? How can these two methodologies work together to illuminate the brain bases of category learning?
3. How can pharmacological approaches, in healthy and brain-damaged populations, be a valuable tool to deeply understand human learning and decision making?

4. How do the basal ganglia, medial temporal lobe, and prefrontal cortex interact during category learning? What other brain regions are essential to consider and in what situations?
5. How can we better understand the functional role of neurotransmitters and their modulation in human learning?
6. How do task demands, stimulus design, surface features, and training methods in category learning experiments influence the brain regions involved?
7. How do probabilistic category structures influence category learning differently than deterministic or rule-based structures?
8. If there are multiple strategies by which subjects can solve category-learning tasks, what factors promote the use of one strategy (and associated brain regions) versus others?
9. How can animal research enhance our understanding of the cognitive neuroscience of category learning? In particular, do animal lesion studies provide converging insights for studies of neurologically impaired clinical human populations, and do animal electrophysiology studies provide converging insights into human brain imaging studies?
10. How does the higher-level structure of a knowledge domain affect the cognitive and neural processes involved in category learning?
11. To what extent do categorization and recognition share common or distinct cognitive and brain mechanisms?
12. How does awareness affect category learning?
13. How can basic research on the cognitive neuroscience of category learning be transitioned and leveraged to further contribute to our understanding and treatment of neurological and psychiatric disorders?

Acknowledgments

For their support of the three-year collaborative consortium, and the production of this special issue, we are indebted to Susan Fitzpatrick and the James S. McDonnell Foundation. Additional support for these activities was provided by conference grants from the *National Science Foundation* and the *National Institute of Mental Health.*

Guest Editors
Mark A. Gluck
Rutgers University – Newark, USA
E-mail address: gluck@pavlov.rutgers.edu

Russell A. Poldrack
UCLA, USA

Szabolcs Kéri
Semmelweis University, Hungary