NeuroImage 71 (2013) 284–297

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Brain activity across the development of automatic categorization: A comparison of categorization tasks using multi-voxel pattern analysis

Fabian A. Soto [a,b,*], Jennifer G. Waldschmidt [b], Sebastien Helie [c], F. Gregory Ashby [b]

[a] Sage Center for the Study of the Mind, University of California, Santa Barbara, USA
[b] Department of Psychological & Brain Sciences, University of California, Santa Barbara, USA
[c] Department of Psychological Sciences, Purdue University, USA

ARTICLE INFO

Article history:
Accepted 8 January 2013
Available online 17 January 2013

Keywords:
fMRI
Multi-voxel pattern analysis
Categorization
Automaticity
Multiple memory systems

ABSTRACT

Previous evidence suggests that relatively separate neural networks underlie initial learning of rule-based and information-integration categorization tasks. With the development of automaticity, categorization behavior in both tasks becomes increasingly similar and exclusively related to activity in cortical regions. The present study uses multi-voxel pattern analysis to directly compare the development of automaticity in different categorization tasks. Each of the three groups of participants received extensive training in a different categorization task: either an information-integration task, or one of two rule-based tasks. Four training sessions were performed inside an MRI scanner. Three different analyses were performed on the imaging data from a number of regions of interest (ROIs). The *common patterns* analysis had the goal of revealing ROIs with similar patterns of activation across tasks. The *unique patterns* analysis had the goal of revealing ROIs with dissimilar patterns of activation across tasks. The *representational similarity* analysis aimed at exploring (1) the similarity of category representations across ROIs and (2) how those patterns of similarities compared across tasks. The results showed that common patterns of activation were present in motor areas and basal ganglia early in training, but only in the former later on. Unique patterns were found in a variety of cortical and subcortical areas early in training, but they were dramatically reduced with training. Finally, patterns of representational similarity between brain regions became increasingly similar across tasks with the development of automaticity.

Introduction

The ability to group objects and other stimuli into classes, despite their perceptual dissimilarity, is extremely helpful for organizing the environment and adaptively responding to its demands. For this reason, categorization has attracted much attention as a subject of behavioral and neurobiological research. Research in the neuroscience of human category-learning has shown that a variety of areas are recruited during learning and performance of categorization tasks, including visual, prefrontal, parietal, medial temporal and motor cortices, as well as the basal ganglia (for reviews, see Ashby and Maddox, 2005; Seger and Miller, 2010).

Multiple systems of category learning

A body of behavioral and neurobiological evidence suggests that the brain areas associated with categorization are organized in relatively separate category-learning systems and that different categorization tasks engage the systems differently (Ashby and Maddox, 2005; Nomura and Reber, 2008; Poldrack and Foerde, 2008). Information-integration (II) tasks, which require the integration of information from two or more stimulus components at a pre-decisional stage, recruit a procedural-learning system that relies on feedback-based learning of associations between stimuli and responses. An example is shown in the top-right panel of Fig. 1, where information about the orientation and width of stripes must be integrated to categorize the stimuli correctly. Rule-based (RB) tasks, in which the optimal strategy is easy to verbalize and can be learned through a logical reasoning process, recruit a declarative-learning system that is based on explicit reasoning and hypothesis testing. An example is shown in the bottom-left panel of Fig. 1, where the simple verbal rule “respond A if the stripes are narrow and B if the stripes are wide” can solve the task.

The COVIS model of category learning (Ashby et al., 1998; for recent versions of the model, see Ashby et al., 2011; Ashby and Valentin, 2005) is a formal description of these two learning systems and the brain regions subserving each of them. Learning of sensory-motor associations in the COVIS procedural system is implemented in the synapses from visual sensory neurons onto medium spiny neurons in the striatum, the input structure of the basal ganglia. The output of the basal ganglia

* Corresponding author at: Department of Psychological & Brain Sciences, University of California Santa Barbara, Santa Barbara, CA 93106, USA. Fax: +1 805 893 4303.
E-mail address: fabian.soto@psych.ucsb.edu (F.A. Soto).

1053-8119/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.neuroimage.2013.01.008

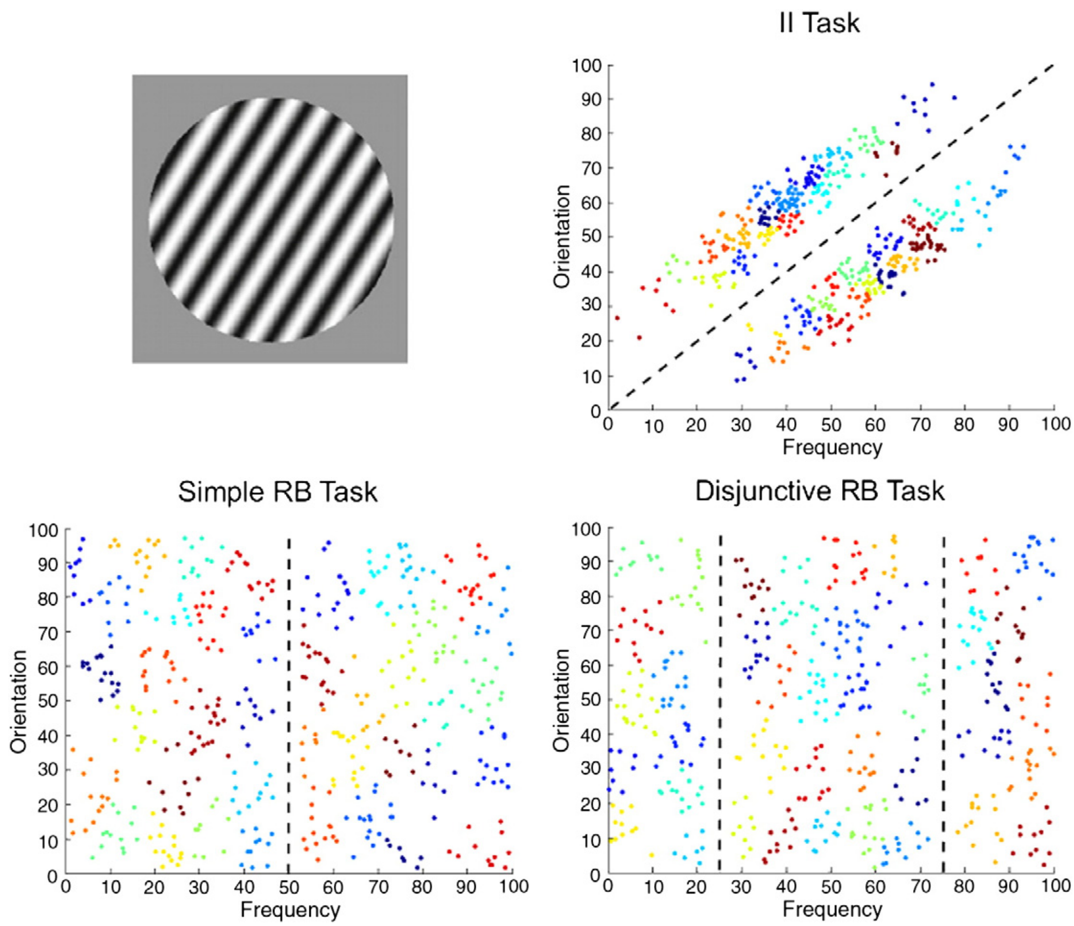

Fig. 1. Information about the tasks and stimuli used in the present study. The top-left panel shows an example stimulus. The other three panels show the category structures in each of the tasks. Dashed lines represent optimal bounds separating the two categories and different colors represent different clusters of stimuli revealed by a *k*-means cluster analysis (see Neuroimaging analysis section).

controls motor responses through its influence on premotor areas, via the ventral lateral and ventral anterior thalamic nuclei. Learning through explicit reasoning and hypothesis testing in the COVIS declarative system is implemented in a network of frontal, medial temporal and basal ganglia areas. Candidate rules are maintained in working memory representations in the lateral PFC, via a series of reverberating loops through the medial dorsal nucleus of the thalamus. These rules are selected from among all available rules via a network that includes the anterior cingulate cortex. If evidence is accumulated that a particular rule does not lead to accurate performance, the rule is switched by reducing attention to it via frontal input to the head of the caudate nucleus, which ultimately inhibits the thalamic-PFC loops in charge of working memory maintenance.

Many behavioral studies have found dissociable effects of experimental manipulations in RB and II tasks. For example, switching the locations of response keys after categorization learning interferes with performance in II tasks, but not with performance in RB tasks (Ashby et al., 2003; Maddox et al., 2004), suggesting that category learning is tied to specific motor responses only in the former. Learning in II tasks is also disrupted if feedback is absent (Ashby et al., 1999), presented before the stimulus (Ashby et al., 2002) or delayed by a few seconds (Maddox et al., 2003). The same manipulations have smaller or no effects in RB tasks. On the other hand, asking participants to perform a simultaneous task during category learning, which demands working memory and attention, interferes more with RB tasks than with II tasks (Waldron and Ashby, 2001). Similarly, dual-task interference has been found for declarative, but not implicit knowledge about a probabilistic categorization task (Foerde et al., 2007).

Neurobiological studies also suggest that the brain areas involved in category learning differ for II and RB tasks. For example, in the only fMRI study that has directly contrasted task-related activity in II and RB tasks, Nomura et al. (2007) found that activity in the hippocampus, anterior cingulate cortex, middle frontal gyrus and body of the caudate all correlated with successful performance in the RB task, whereas only activity in the body and tail of the caudate correlated with successful performance in the II task. Direct comparison of task-related activity between the tasks in several regions of interest (ROIs) revealed higher activity for the RB than the II task in the hippocampus, and higher activity for the II than the RB task in the caudate, suggesting a dissociation between a hippocampal-based declarative system and a basal ganglia-based procedural system in category learning.

Several other studies have found results that are generally in agreement with COVIS. During the early stages of learning of RB categorization tasks, accuracy is found to be correlated with activation in the hippocampus, head of the caudate, dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, and posterior parietal cortex (Filoteo et al., 2005; Helie et al., 2010a; Seger and Cincotta, 2006). On the other hand, activity during learning of II tasks increases in the body and tail of the caudate, and in the putamen (Cincotta and Seger, 2007; Waldschmidt and Ashby, 2011).

Other studies have used a “weather prediction” categorization task, in which feedback about category membership is usually probabilistic. Participants can use a variety of strategies to achieve good performance in this task (Gluck et al., 2002) and neuroimaging studies suggest that dissociable learning systems might underlie such strategies (for a review, see Poldrack and Foerde, 2008). For example,