Brain and Cognition 74 (2010) 88–96

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Refining the visual-cortical hypothesis in category learning

Mariana V.C. Coutinho [a,*], Justin J. Couchman [a], Joshua S. Redford [a], J. David Smith [b]

[a] Department of Psychology, University at Buffalo, The State University of New York, United States
[b] Department of Psychology and Center for Cognitive Science, University at Buffalo, The State University of New York, United States

ARTICLE INFO

Article history:
Accepted 4 July 2010
Available online 1 August 2010

Keywords:
Categorization
Prototypes
Typicality
Perceptual representation system
Visual cortex

ABSTRACT

Participants produce steep typicality gradients and large prototype-enhancement effects in dot-distortion category tasks, showing that in these tasks to-be-categorized items are compared to a prototypical representation that is the central tendency of the participant's exemplar experience. These prototype-abstraction processes have been ascribed to low-level mechanisms in primary visual cortex. Here we asked whether higher-level mechanisms in visual cortex can also sometimes support prototype abstraction. To do so, we compared dot-distortion performance when the stimuli were size constant (allowing some low-level repetition-familiarity to develop for similar shapes) or size variable (defeating repetition-familiarity effects). If prototype formation is only mediated by low-level mechanisms, stimulus-size variability should lessen prototype effects and flatten typicality gradients. Yet prototype effects and typicality gradients were the same under both conditions, whether participants learned the categories explicitly or implicitly and whether they received trial-by-trial reinforcement during transfer tests. These results broaden out the visual-cortical hypothesis because low-level visual areas, featuring retinotopic perceptual representations, would not support robust category learning or prototype-enhancement effects in an environment of pronounced variability in stimulus size. Therefore, higher-level cortical mechanisms evidently can also support prototype formation during categorization.

1. Introduction

Categorization—the formation and use of psychological equivalence classes—is a basic ability that is critical to all domains of cognition and to survival. For this reason, categorization is a sharp research focus (e.g., Ashby & Maddox, 1992; Homa, Sterling, & Trepel, 1981; Kruschke, 1992; Medin, 1975; Murphy, 2003; Rosch & Mervis, 1975; Smith & Minda, 1998).

There is strong evidence that humans sometimes categorize objects using prototypes. That is, category learners average or blend the category members they experience to form a prototype, compare new to-be-categorized items to it, and accept these items as category members if they are similar enough to the prototype (e.g., Ashby & Maddox, 2005; Cook & Smith, 2006; Smith, 2002; Smith, Chapman, & Redford, 2010; Smith & Minda, 1998, 2001, 2002).[1]

Humans' capacity for prototype abstraction can be illustrated using the influential dot-distortion category task (e.g., Blair & Homa, 2001; Homa et al., 1981; Knowlton & Squire, 1993; Posner, Goldsmith, & Welton, 1967; Smith, Redford, & Haas, 2008; Smith et al., 2010). In this task, participants are trained on a family of shapes that are all distortions of an originating prototype, and they are then asked to endorse (or not) previously unseen probe items as belonging in the trained category. These probe items are copies of the prototype, low- and high-level distortions of it, and random items outside the trained category (Fig. 1, rows 1–4, respectively).

Humans in this task show strong endorsement of the prototype relative to other category members, as they would if they had ab-

* Corresponding author. Address: Department of Psychology, Park Hall, University at Buffalo, The State University of New York, Buffalo, NY 14260, United States.
E-mail address: mvc5@buffalo.edu (M.V.C. Coutinho).

[1] That humans have prototype formation as one component of their category-learning capacity does not imply that humans learn only through prototypes or that their categorization is unitarily prototype based. It is not. A complete description of the categorization system of humans and possibly nonhumans would include, in addition to prototypes, rules, decision bounds, exemplar processes, and so forth. Indeed, a multiple-systems approach is an important part of the human categorization literature (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Ell, 2001; Erickson & Kruschke, 1998; Homa et al., 1981; Minda & Smith, 2001; Rosseel, 2002; Smith & Minda, 1998) and the comparative categorization literature (Cook & Smith, 2006; Herbranson, Fremouw, & Shimp, 1999; Smith, Beran, Crossley, Boomer, & Ashby, 2010; Wasserman, Kiedinger, & Bhatt, 1988). Prototype formation is one important component of humans' overall categorization system that we emphasize in the present article simply because it is the article's empirical and theoretical focus.

0278-2626/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2010.07.001

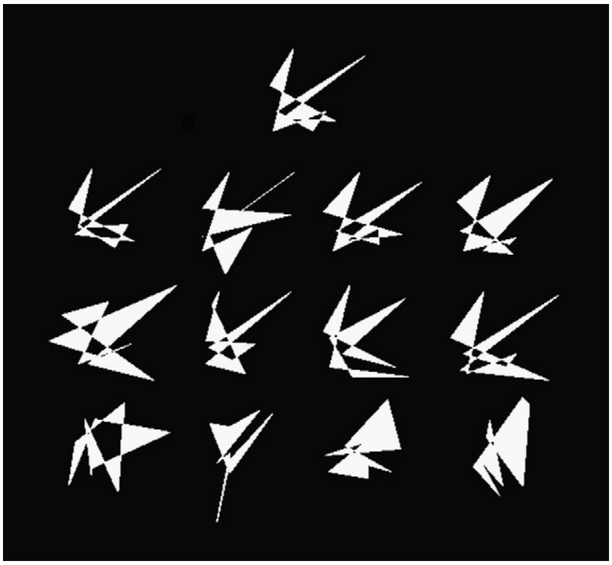

Fig. 1. A dot-distortion category, with the originating prototype, low-level distortions, high-level distortions, and random-unrelated shapes in rows 1–4, respectively.

stracted with privilege that central category representation (e.g., Knowlton & Squire, 1993). They also produce a steep typicality gradient—that is, a large change in category endorsement level from prototypes to low-level distortions to high-level distortions (Fig. 2). This steepness also points to an underlying prototype representation because the prototype is the unitary, central point in psychological space, and as one creates higher-level distortions, one always moves directly away from it, producing more dissimilar stimuli and weaker category members (Smith & Minda, 2001, 2002). These results are fit distinctively well by a prototype model—in Fig. 2, the model's predictions missed the actual observations by only 1.5% per data point.[2] Thus, these results confirm that prototype abstraction is one component of humans' multifaceted categorization capacity (Ashby & Maddox, 2005).

It is possible now to grant humans' prototype-formation capacity greater phylogenetic depth. Smith et al. (2008) found that monkeys (*Macaca mulatta*) also showed steep typicality gradients that were distinctively well fit by a prototype model (Fig. 3). In this case, the model's predictions missed the actual observations by only .2% per data point. Clearly, humans and animals share a capacity for prototype formation in the dot-distortion category task, suggesting that there may be basic mechanisms common to human and monkey perception that underlie prototype-enhancement effects. The dot-distortion task is the focus of the present research because it currently provides the strongest evidence for prototype-abstraction processes and because it is the dominant task within the cognitive-neuroscience literature that is reviewed shortly. In fact, though, a wide range of prototype and typicality effects have now been demonstrated in the comparative-cognition literature (e.g., Aydin & Pearce, 1994; Cook & Smith, 2006; Huber & Lenz, 1993; Jitsumori, 1996; Smith et al., 2008, 2010; von Fersen & Lea, 1990; White, Alsop, & Williams, 1993).[3]

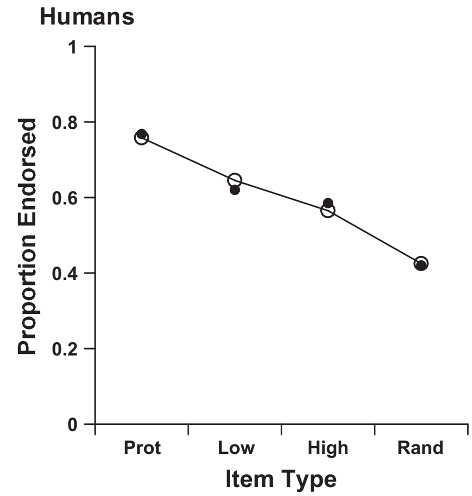

Fig. 2. The composite observed profile (filled circles) of category endorsements (i.e., affirmations that shapes belong in the training category) produced by humans in four dot-distortion category-learning studies (controls and amnesics in Knowlton & Squire, 1993; participants in Reber et al. (1998a, 1998b)). Also shown (open circles) is the average of the four best-fitting predicted profiles when a prototype model fit the individual data sets. Prot: prototype; low, high: low- and high-level distortions of the prototype; Rand: random shapes unrelated to the target category.

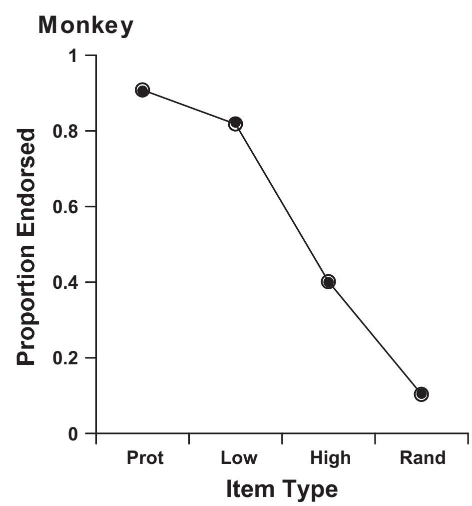

Fig. 3. The composite observed performance profile (filled circles) from 10 dot-distortion sessions by a rhesus macaque (*Macaca mulatta*) in Smith et al. (2008). Also shown (open circles) is the average of the 10 best-fitting predicted profiles when a prototype model fit the individual data sets. Prot: prototype; low, high: low- and high-level distortions of the prototype; Rand: random shapes unrelated to the target category.

Given the consensus about prototype abstraction as one aspect of category learning, theorists have begun to analyze the mecha-

[2] Detailed modeling procedures for the dot-distortion paradigm can be found in Smith (2002), Smith and Minda (2001, 2002), or Smith et al. (2008). Generally, the prototype models illustrated in Figs. 2 and 3 of the present article begin with the stimulus distance between a to-be-categorized stimulus and the prototype of the category that is the presumed reference standard. Distance is transformed into psychological similarity using an exponential-decay function that incorporates a freely varying sensitivity parameters. Similarity is transformed into a category-endorsement level using a choice rule. Standard hill-climbing algorithms are used to maximize the fit between observed and predicted performance profiles and thus to find the configuration of the model that best reproduces the observed profile. The prototype models illustrated in Figs. 2 and 3 are particularly constructive within the literature because they offer a direct comparison to exemplar models that take identical inputs, have identical parameters, and differ only in their underlying representational assumption.

[3] An apparent exception to findings of prototype formation by humans and animals was reported by Sigala, Gabbiani, and Logothetis (2002). However, the reasons for this exception are clear. Humans and monkeys in these tasks were trained using only five exemplars in each category. The small size of these exemplar sets is well known to encourage exemplar memorization as a task strategy and to discourage abstraction or prototype formation (Homa et al., 1981; Smith, Murray, & Minda, 1997; Smith & Minda, 1998, 2000). Moreover, Sigala et al. categorization task was not linearly separable along two of its four dimensions, definitionally ruling out the adaptive use of abstraction or averaging along those dimensions. Finally, the tasks in Sigala et al. were essentially solvable by a single-dimensional rule, discouraging multi-dimensional prototype formation for yet another reason.