Cognition 162 (2017) 87–102

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Alignability-based free categorization

John P. Clapper *

Department of Psychology, California State University San Bernardino, United States

ARTICLE INFO

Article history:
Received 2 December 2015
Revised 6 February 2017
Accepted 8 February 2017

Keywords:
Categorization
Concepts
Unsupervised learning
Category construction
Alignment

ABSTRACT

Much evidence suggests that real-world natural kinds are based on overall similarity or family resemblance, but people often appear surprisingly insensitive to family resemblance in laboratory studies of sorting or free categorization. In such experiments, all stimuli generally vary along the same discretely-varying dimensions and family resemblance is defined in terms of the proportion of matching or mismatching values along those dimensions. This article argues for an alternative conception of family resemblance based on structural alignability, i.e., whether objects have corresponding parts-in-relations that can provide the basis for a shared schema or conceptual model. Five experiments using two new free categorization tasks demonstrate that structural alignment, even without specific matching parts, is sufficient for people to perceive objects as essentially similar and group them into common family-level categories. Importantly, the experiments demonstrate that this categorization is based on abstract alignment rather than shared parts or features, because when the parts of the individual objects are randomly rearranged, eliminating their shared spatial structure, people no longer perceive them as belonging to a common category. These results suggest that people do construct perceptual categories on the basis of overall similarity, at least when similarity is defined in terms of spatial correspondence or alignability rather than individual shared parts or features.

1. Introduction

Philosophers and scientists have long wondered at the mind's ability to carve its experience into natural and appropriate categories, elegantly captured by Plato's famous metaphor of a butcher dividing a carcass at its natural joints (*Phaedrus*, trans. 1997). The research described in this article attempts to address some of the issues raised by Plato's metaphor within a simple but ecologically important task context, namely, the basic perceptual categorization of novel visual objects. The main issues can be framed as follows: How does a normal observer presented with a set or series of novel objects detect natural kinds or equivalence classes among those objects? In particular, how do they construct meaningful categories in the absence of external guidance or feedback, when they are free to partition the objects any way they wish? While far from providing a complete account of human categorization, answering these questions might at least help provide a basic understanding of some of the processes that underlie simple forms of learning in daily life.

It has often been assumed that perceptual categorization should be based on overall visual *similarity* (e.g., Hampton, 2001; Rosch, 1975; Wittgenstein, 1957), i.e., that people should put things that look alike into the same categories, and things that do not look alike into different categories. Supporting this intuition is the fact that most of the object categories we are familiar with from everyday life (e.g., types of plants, animals, and artifacts) do indeed seem to be based on overall similarity or *family resemblance* (e.g., Rosch & Mervis, 1975). Both prototype (Posner & Keele, 1968, 1970; Reed, 1972) and exemplar (Medin & Schaffer, 1978; Nosofsky, 1984) models of categorization assume that objects are assigned to categories based on their similarity to other objects within those categories. There are also technical arguments related to category *utility*, which assume that categories serve certain adaptive functions, such as feature inference (e.g., Anderson, 1990, 1991), that tend to be maximized when they have high internal and low external similarity (Corter & Gluck, 1992; Fisher & Langley, 1990; Gluck & Corter, 1985; Lassaline & Murphy, 1996; Rosch & Mervis, 1975). The category utility view, in particular, suggests that people should create categories based on overall similarity to maximize their predictive value in the real world.

Given these intuitions, demonstrating similarity-based free categorization in the laboratory using artificial stimuli and controlled tasks would seem to be a natural and important goal for research in

* Address: Department of Psychology, California State University San Bernardino, 5500 University Parkway, Sa Bernardino, CA 92407-2397, United States.

E-mail address: jclappe@csusb.edu

http://dx.doi.org/10.1016/j.cognition.2017.02.003
0010-0277/

this area. A number of studies have in fact attempted to do that, with surprisingly negative results. Below, I briefly review the literature on free categorization and show that this research has failed to provide convincing demonstrations of similarity-based categorization that scale robustly across different tasks and stimulus types. I then suggest a possible explanation for this failure, arguing that it relates to the way in which similarity and category structure have been defined in previous research, and suggest an alternative definition based on the structural correspondence or *alignability* (Gentner, 1983) among objects within a category. The experiments reported below aimed to demonstrate that people can acquire categories based on this definition of similarity, to eliminate alternative explanations for that result, and to provide some basic information about how this learning occurs.

1.1. Research on similarity and free categorization

Most previous research on free categorization has employed some type of sorting procedure. For example, several experiments (e.g., Handel & Imai, 1972; Imai & Garner, 1965; Pothos & Chater, 2002) used free sorting – in which people are simply asked to sort a set of stimuli into categories of their own choosing – to investigate how people categorize simple stimuli varying along one or two continuous dimensions (e.g., circles varying in area and radius orientation, lines varying in length and thickness, etc.). One early result was that stimuli constructed of *integral* dimensions (e.g., colors varying in brightness and saturation) were sorted on the basis of overall similarity while those constructed of *separable* dimensions (e.g., geometric figures varying in shape and color) were sorted on the basis of a single dimension, with no apparent role for overall similarity (Handel & Imai, 1972).

Some free-sorting studies do seem to provide evidence for similarity-based sorting of separable stimuli, at least under certain conditions (e.g., Pothos & Close, 2008; Pothos et al., 2011). For example, Pothos et al. (2011) presented participants with stimuli that varied along two continuous dimensions and that fell into two, three or more similarity-based clusters. It was clear that in at least some cases, e.g., when they constructed three or more categories for a given set, people were using more than one dimension. On the other hand, there was very little convergence in participants' overall sorting behavior in this study; even when the particular sort predicted by overall similarity was the most frequent one for a given set, it would typically be produced only by a minority of participants. So, while these studies suggest that people are sometimes capable of sorting simple stimuli on the basis of overall similarity, they provide little assurance that this is an easy or natural thing for them to do.

The issue of whether similarity-based sorting is an easy or natural form of categorization has been investigated using a variation of free sorting known as the *triad* task (e.g., Smith & Kemler Nelson, 1984; Ward, 1983; Wills, Inkster, & Milton, 2015). In this task, three stimuli are presented on each trial and the participant selects which two go best together. Two of the objects match exactly on one dimension, while another two match on neither dimension but are more similar overall. Early results (e.g., Smith & Kemler Nelson, 1984; Ward, 1983) appeared to show that people preferred the pair with higher overall similarity, as opposed to the pair matching on a single dimension. However, later research (Wills et al., 2015) has convincingly shown that participants' actual strategy is to simply pick one dimension and then always choose the pair that is most similar on that dimension, ignoring the other dimension completely. This tendency to engage in one-dimensional sorting increases under time pressure and cognitive load (Wills et al., 2015), reinforcing the conclusion that one-dimensions sorting is the easiest and most natural mode of categorization in this task.

The studies discussed so far all employed simple stimuli that varied continuously along only two dimensions. However, it might be argued that a fair test of the similarity hypothesis requires more complex stimulus sets, in which objects vary along multiple dimensions. Since variation along any one dimension necessarily makes up a smaller proportion of total variation in such a set, one might expect people to be less inclined to sort on the basis of one dimension and more inclined to focus on overall similarity.

A number of experiments have searched for evidence of similarity-based sorting using multi-dimensional, discretely-varying stimulus sets. A particularly well-known set of studies was reported by Medin, Wattenmaker, and Hampson (1987). Their stimuli were composed of four binary dimensions and were divided into categories based on overall similarity, as illustrated in Table 1. Each category had a set of characteristic values present in most examples of that category, but no single value was perfectly diagnostic of either category, consistent with standard notions of family resemblance. The results provided no evidence that people recognized this family resemblance structure, however; instead, they divided the stimuli along a single dimension, while ignoring all other dimensions. Subsequent research has strongly reinforced these results, showing an overwhelming preference for one-dimensional as opposed to family resemblance sorting across a wide variety of stimulus types and minor task variations (Ahn & Medin, 1992; Regehr & Brooks, 1995).

Regehr and Brooks (1995) proposed that the preference for one-dimensional sorting reported by Medin et al. (1987) and others could be an *array effect*, i.e., due to the fact that the stimulus objects were all presented simultaneously for participants to examine and compare. They argued that people find it difficult to track multiple dimensions across objects in this situation, and so focus on a single dimension as an easy way to perform the sorting task. To test this claim, they created a new task in which the prototypes from each category served as standards, and all the other objects had to be sorted into one of two categories defined by these standards. Here, participants made a series of pairwise comparisons between each stimulus and the two standards, and should thus select the standard that is most similar to the current stimulus each time. Consistent with this line of reasoning, Regehr and Brooks found that people were much more likely to categorize on the basis of overall similarity in this task than in the regular full-array task. Subsequent research (Milton, Longmore, & Wills, 2008; Milton & Wills, 2004; Wills, Milton, Longmore, Hester, & Robinson, 2013) has confirmed this result.

The match-to-standards results show that it is possible to induce family resemblance sorting if the category prototypes are specified in advance and the comparison process is constrained in specific ways (see also Murphy, Bosch, & Kim, 2016; Norenzayan, Smith, Kim, & Nisbett, 2002). However, they still provide no real support for the claim that family resemblance is a primitive or basic principle of human categorization (Regehr & Brooks, 1995). One problem is that a substantial amount of one-dimensional sorting is still observed in some match-to-standards experiments (e.g., Milton & Wills, 2004). Another problem, first noted by Regehr and Brooks (1995), is that participants who categorize on the basis of family resemblance often seem to do so slowly and with noticeable effort. Later research (e.g., Milton & Wills, 2004; Milton et al., 2008; Wills et al., 2013) has shown that family resemblance sorting in the match-to-standards task results from an effortful *dimensional summation* strategy, in which people count the number of matching dimensions between the target and the two standards and select the standard with more matches as winner. Consistent with this account, increasing cognitive load or time pressure increases one-dimensional sorting and decreases similarity-based sorting in the match-to-standards task (Wills et al., 2013). The obvious implication is that one-dimensional