The Ontogeny of Units in Object Categories

Philippe G. Schyns and Gregory L. Murphy

Dept. of Cognitive and Linguistic Sciences
Brown University
Providence, RI, 02912
schyns@clouseau.cog.brown.edu

Abstract

Theories of object recognition and categorization rely on a set of primitives to represent objects. The nature and the development of these primitives have been neglected in computational vision and in concept learning theories. We present a theory of part ontogeny in which not only perceptual, but also categorical constraints play a role. A two-phase experiment using categories of synthesized 3D objects (Martian rocks) was conducted to test the theory. The first phase tested the hypothesis that part identification is dependent on categorical context. The second phase tested whether the units extracted in the first phase played a conceptual role in learning a new category. In both phases, subjects interactively delineated the parts of the stimuli while learning the categories. The units subjects identified in the first phase were those that were predictive of the object's category. These units then influenced the perception of parts in the new categories of the second phase.

Introduction

Most accounts of object recognition and categorization assume that in order to identify an object, one must first identify its components. Of course, not all components of an object must be identified before an object can be categorized, but many of them are probably recognized during the categorization process. For example, many accounts of categorization (see Smith & Medin, 1981) would claim that a person who sees a cat might identify a global shape, a size, body parts such as the head, tail and whiskers, a color, a texture, and so on. After these perceptual components of the object have been identified, they are compared with memory representations in order to categorize the object. Theories that embody this general approach include Rosch and Mervis's (1975) family resemblance model, Marr and Nishihara's (1978) generalized cones, and Biederman's (1987) Recognition By Components.

The topic of this paper is the nature of the components themselves. How are these components derived? Are the components independent entities, or does category learning influence them? In particular, we will be interested in what might informally be called the *parts* of an object. Tversky and Hemenway (1984) showed that the most useful categories (i.e., Rosch et al.'s 1976 basic categories) can be distinguished by parts whereas less useful categories cannot (see also Murphy, in press). Furthermore, parts have a privileged place in a number of accounts of object identification (Biederman, 1988; Biederman & Ju, 1988; Binford, 1981; Marr, 1982; Marr & Nishihara, 1978; Pentland, 1989).

Although parts may well be important in categorization, the question arises as to how people identify an object's parts and encode them in memory. The most common assumption concerning parts and other components of an object is that they are the result of perceptual structures. That is, our perceptual apparatus identifies the colors, size and parts of an object which then form the basis for conceptual tasks. Certainly, perceptual structures provide important constraints on what can be a part of an object. But important research in computationnal vision and psychophysics has shown that part extraction without constraints is essentially an ill-posed problem (Bertero et al., 1987; Biederman, 1987; Braunstein et al., 1989; Hoffman & Richards, 1985; Koenderink & van Doorn, 1982; Poggio et al., 1985). However, an additional constraint, which has received little attention, is the effect of category learning itself. The process of concept formation may alter the interpretation of objects such that new parts and combinations of parts are encoded.

To understand how this might occur, consider this thought experiment. Suppose you were a Martian with no experience of Earth objects. The first object you see is a cup which can be segmented into the analytical units *handle* and *container*. However, your Martian conceptual system does not possess these primitives. Supposing your low-level visual system were roughly like ours, you could describe the cup by mentioning at least that something seems to break the regularity of its surface: something sticks out. Suppose the second object you experience is a glass. This object is much smoother; nothing sticks out. From now on, you could use the primitive *Xd@#* (you certainly would not call it a handle), because it categorizes the two objects: one has a *Xd@#* and one doesn't. Before encountering the glass, you might have represented the cup as a whole, as a single unit. Thus, the category learning situation may select some parts and suppress others in representing objects.

This phenomenon illustrates what we call *the Homogeneity Principle*. This principle says that if a component or group of components occurs frequently throughout a category, then it will be treated as a single unit. In other words, the objects of the category provide a sufficient context to change a component's status from a potential part to an actual part--or unit--that is used to represent the object. We will call these analytical parts "conceptually instantiated units," or simply "units."

Units are the parts that are input into concept representation. Examples of units are *head*, *tail*, *handle* and so forth. Conceptually instantiated parts play a *conceptual role* by segmenting the external world into categories of objects. Once a unit is instantiated in the context of one particular category, it can then play a conceptual role in other contexts of categorization. To illustrate, if the Martian had extracted what we call a handle, it could use it in categorizing new objects that have a handle. The Homogeneity Principle also assumes that exposure to different contexts may result in the instantiation of different units. Therefore, this theory provides a framework to study the ontogeny of part/whole relationships. The representation of an object could eventually be decomposed into two units if the parts composing the object appear to play a conceptual role in another context; otherwise, the whole may be treated as one instantiated unit.

The study reported here examines the Homogeneity Principle through a two-part experiment using categories of novel objects. The first phase tested the hypothesis that part instantiation into units is dependent on categorical context. It also investigated a claim of *sufficiency*, that the context of one category is sufficient to instantiate the parts needed to represent the objects of the category. Components that were homogeneous throughout the category were expected to be instantiated as units, and other components were expected to be rejected, over the course of learning. The second phase introduced a new category, which overlapped somewhat with the previous one. This allowed a test of the conceptual role played by the units learned in the first phase. The overlapping units were expected to be parsed first, and thereby to bias the extraction of other components in the new category. A group with no prior experience with those units served as a control. It was expected to identify a different part structure (see below).

Studying the ontogeny of units presents certain methodological challenges. Using familiar objects is obviously inappropriate, because the theory concerns the effect of category learning on part analysis. However, most artificial categories used in previous experiment of this sort could not be used here because it was necessary to avoid familiar *units* as well as familiar objects. So, it would not do to teach subjects a new kind of furniture if they could already identify familiar units such as legs, seats, cushions, drawers, etc. One would not expect to counteract subjects' extensive prior knowledge of these parts. Therefore, the experiment used 3-dimensional spherical objects with various mountain-like projections. Each projection or group of projections could serve as a potential part. The experiment examined the degree to which subjects' perception of parts could be influenced by the context of one category of objects, and how the units acquired in this context would influence the extraction of new units in another category.

Phase 1

Subjects were randomly assigned to one of two conditions in which they were exposed to exemplars of a particular category (A or B). Category A was a collection of objects on which a specific part (*a*) was always present. Category B was also a collection of objects, but they were characterized by another consistent part (*b*). We recorded the parts subjects considered important before, during and after category learning.

The Homogeneity Principle predicts that the parts subjects select after exposure to a category should be the ones that hold the category together. Therefore, subjects should pick part *a* or part *b* if they are in condition A or B, respectively. Also, the number of nontarget parts subjects pick should decrease as they gain more and more experience with the category. A corrolary of homogeneity is that subjects should not consider as relevant parts those that do not glue a category together. Thus, if subjects from group A were exposed to the target part of group B, they would not pick part *b* as a relevant part.

Method

Subjects. Twelve Brown University undergraduates participated in the first phase of the experiment. Subjects were assigned randomly to condition A or B with the constraint that the number of subjects be equal in each condition.

Stimuli. "Martian rocks" were synthetized three-dimensional gray-level shaded objects displayed on the screen of a graphic mini-supercomputer (Stellar GS 1000). Each stimulus was generated on a Lisp Machine using the S-Geometry package. A Martian rock was created from a sphere by specifying eight equidistant regions along its equator. Each region covered a surface extending from the equator to the two poles of the sphere, and deformations were applied to each region. A deformation consisted of pulling out a facet of a region for a certain distance along the surface normal at that point. The deformations were deterministic for the target parts, while the seven remaining regions were randomly distorted. The top pictures of Figure 1 show the deformations that created unit *a* and unit *b*, and the bottom pictures present complete exemplars of categories A and B. Random Martian rocks were also created by randomly distorting *all* the regions. We will call the objects in the category being learned the M1 rocks.

Procedure: Initial delinations. Before experiencing objects of their category, subjects were asked to delineate

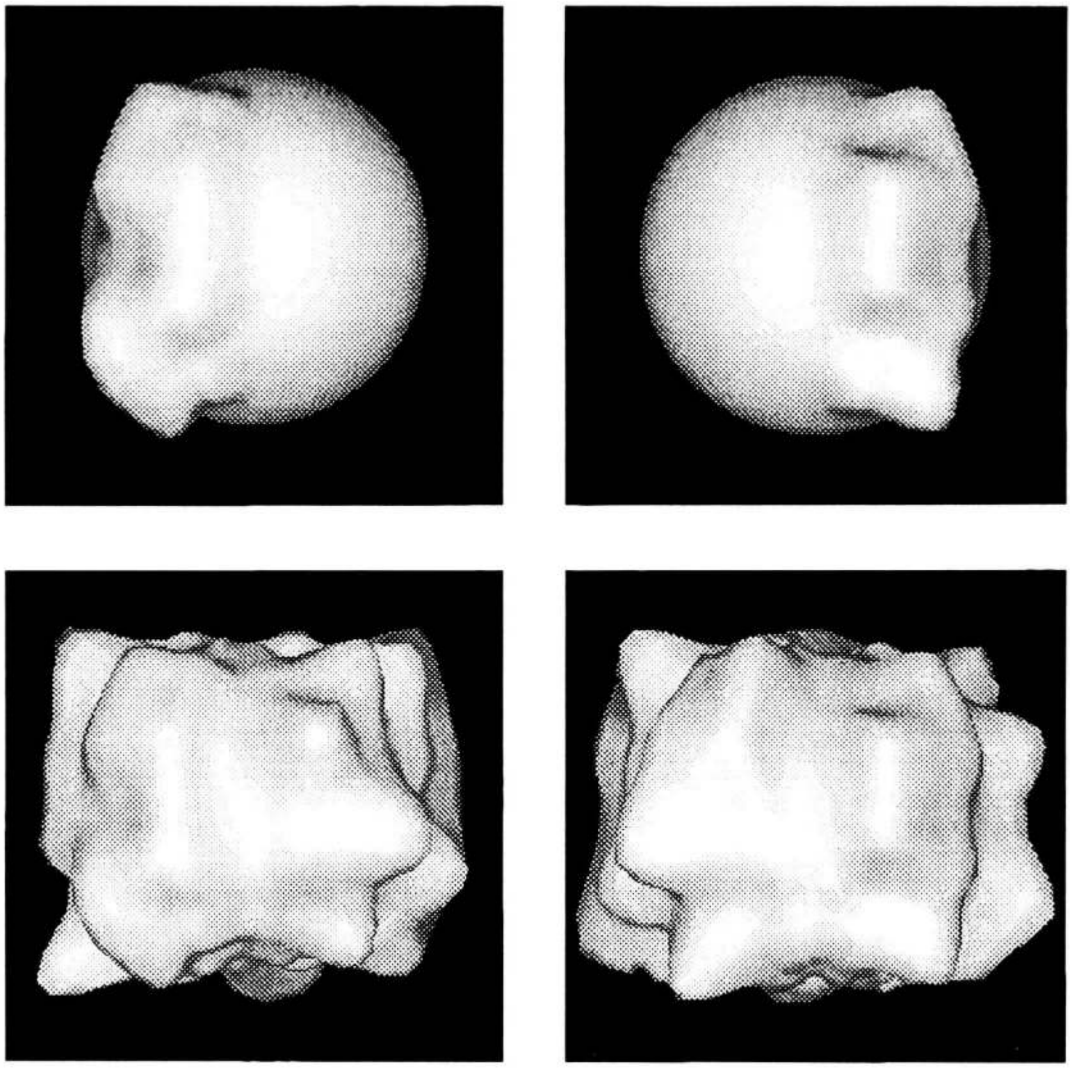

Figure 1: This figure illustrates the stimuli used in Phase 1. The top pictures show part a (left) and b (right) in isolation. The bottom pictures show examples of the complete objects.

the parts they thought were important on one M1 rock. Subjects were instructed to delineate at least one part per object, but no upper limit was specified. The purpose of this initial stage was to test for the salience of the target parts. Part delineation was done using a computer mouse. Using the left and right buttons of the mouse, subjects had to center the part they wanted to delineate with respect to a vertical bar that appeared on the screen. Then, subjects moved a cursor on the screen with the mouse in order to mark with a red line the part they had previously centered. That is, by clicking the mouse, subjects extended the line from its last vertex to the current location. They continued this delineation until the entire part was indicated. Once one part was delineated, subjects could either center another part of the object or quit part delination for that object.

Learning stage. The learning stage was divided into 6 trials. On each trial, subjects first saw an M1 rock. This rock would rotate once clockwise and once counterclokwise in front of them. Subjects were told to study this rock to learn what M1 rocks were like in general. Then this stimulus was replaced with another rock on which subjects were instructed to delineate parts as described earlier. For five out of six delineation rocks, this item was an M1, and the other delineation rock was from the category of the other condition (category B for subjects in group A, and vice versa).

Test stage. In order to ensure that subjects learned the category, a categorization stage with 6 test rocks was conducted. Two of the test rocks were M1 rocks, and the other four were random rocks. One test rock at a time would rotate--once clockwise, once counterclockwise--on the screen and subjects had to decide, by pressing the "yes" or the "no" key on the keyboard, whether the rock was an M1. If subjects were correct on all six items, the first phase of the experiment was over. Otherwise, they repeated as many learning blocks as needed to perform the test stage without mistake. No feedback was given during the testing phase.

Results

We constructed two grids of points in the exact configurations of part *a* and part *b*. The measure of subjects' accuracy in part identification was the number of points of the target parts contained within their delineations. Fifty percent or more of the target's surface had to be delineated in order to score the subject as having identified the target. To test the salience of the target parts, we collected the total number of part delineations in the initial delineation stage (before learning). A total of 44 parts were initially delineated; only one of them met the criterion for being a target part. Thus, the targets were not very salient units before learning.

To test the first aspect of the Homogeneity Principle, we considered the last learning session for all subjects (the one preceding a successful test stage). In this session, we computed the average number of points included in the delineation of the last target part (the M1 part), and compared it with the average number of points included in the delineation of the part from the other condition. Subjects delineated 93% of the target from the category they learned, but only 8% of the target from the other category on average. These means were significantly different $t(11) = 11.88, p < .001$.

The corrolary of homogeneity was tested by recording the number of nontarget delineations per object for all subjects. We expected the number of nontarget selections to decrease with learning. Thus, we divided each subject's learning stage into five blocks (since each subjects saw 5 M1 rocks for delineations per learning stage, the total number of M1 objects seen in learning was always a multiple of 5) and counted the number of nontargets delineated in each block. The number of nontargets decreased significantly from the first to fifth block ($F(4,44) = 18.63, p < .001$; means of 2.78, 2.09, 1.68, 0.70 and 0.21). A contrast showed that this decrease was linear ($F(1,44) = 73.59, p < .001$), and there was little remaining variance ($F(3,44) = .32$).

Discussion

The results confirmed the two predictions of the Homogeneity Principle. The conjunction of perceptual constraints and the context of several objects forming a category provide a sufficient means to conceptually instantiate potential parts. At the end of learning, subjects picked the part that was redundant throughout the category. It is worth emphasizing that prior to the experiment, subjects didn't know the parts composing the objects. The corollary of homogeneity is that as subjects gain more and more experience with the context of one category, they should progressively reject infrequent potential parts as plausible candidates for a conceptual unit status. This prediction was corroborated with a linear decrease of the nontarget delineations.

If the unit instantiated in one particular context has a conceptual role, it should be used to describe objects of a new category. Once instantiated, these units have a status similar to *hand, wheel, handle* and so forth. That is, they can be used to characterize new objects. The purpose of the second phase of the experiment was to test both implications of the conceptual role by presenting subjects with a new category of objects in which a number of units are distributed throughout the category, with either one or none of them being already instantiated.

Phase 2

Subjects from both groups of the first phase were exposed to a new group of objects, category C. In addition to groups A and B, a new group of subjects, group C, who did not participate in Phase 1, served in Phase 2. All objects of Phase 2 were characterized by part *a* and part *b* adjacent to one another.

Method

Stimuli. New stimuli were created in which part *a* and part *b* were adjacent to one another and only 6 regions were random. Figure 2 shows different views of a Martian rock from this phase.

Subjects. 18 Brown undergraduates participated in the second phase of the experiment. 12 of them were the subjects of the previous phase and 6 subjects were new to the experiment.

Procedure. Phase 2 took place immediately following Phase 1 for groups A and B. As in the previous phase, subjects went through three main stages: an initial delineation stage, a learning stage and a testing stage. As before, subjects went through six trials in the learning phase, and they had to delineate 5 rocks of category C (the remaining rock was a random rock). In the test phase, 6 rocks were presented (2 of them with part *a* only, two with part *b*, and two with part *a and* part *b*), and subjects had to categorize them correctly to end the experiment. No feedback was given during the testing phase.

Results

We recorded the number of times part *a* or part *b* was delineated in the only object tested before learning. The

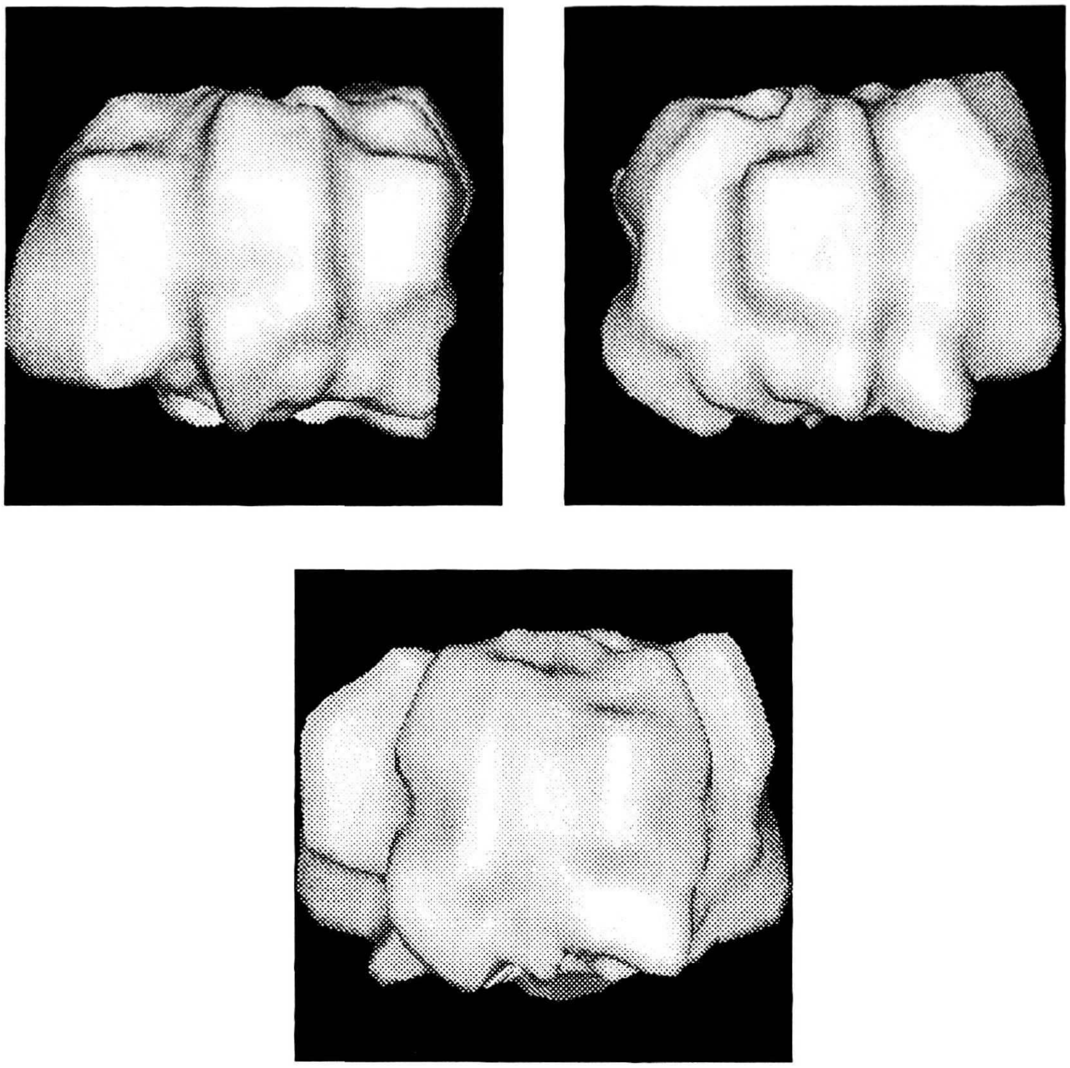

Figure 2: This figure illustrates the stimuli from Phase 2. Subjects in conditions A and B delineated the two parts presented in the top pictures. Subjects in condition C selected the part shown in the bottom picture as one unit. This part is the conjunction of the two parts shown above. During the experiment, subjects could rotate the object to center the part they wanted to delineate.

subjects of group A always delineated part *a*, and never *b*. The subjects of group B always delineated *b*, and never *a*. On the final learning trial, the subjects from groups A and B all delineated parts *a* and *b* as two distinct units, while all subjects from group C (the group with no prior experience with Martian rocks) delineated part *a and b* as a single unit. Typical delineations from groups A and B are presented on Figure 2 (the top pictures) and a typical delineation from group C is presented on Figure 2 (the bottom picture).

Discussion

This study investigated the role of previous knowledge in parsing new objects. We can see that the units instantiated in the first phase were applied to categorize objects of category C. This shows that they played a conceptual role in the sense explained earlier. The bias created by units on part extraction is also illustrated in the second phase. Specifically, subjects with prior knowledge ended up instantiating two units--the one they had learned in the previous phase, either *a* or *b*, and the other part adjacent to the one they knew already. In contrast, subjects from condition C didn't segment *a* and *b* into two instantiated units. Instead, they considered *a* and *b* to be one single unit *c*. This is interesting for two reasons. First, *a* and *b* adjacent to one another can clearly be segmented as two units--as shown by the other subjects. This means that the two parts are separable on a perceptual basis. In fact, the parts could be extracted with a scheme along the lines of Hoffman and Richard's (1985) minima rule. Subjects in group C apparently encoded *c* as a perceptual unit which was not composed of the units *a* and *b*, which were not known as such to the subjects of condition C. This leads to the second point. The context of categorization may influence what is considered a unit in an *object representation*. The context enables the selection of a perceptual interpretation that maximizes the homogeneity of a category. Since parts *a* and *b* are always present on the object of the category, there is no reason to segment them into two separate parts, unless they have already been distinguished. This last result has far-reaching consequences for understanding the development of a conceptual system. Potential parts become instantiated units if the right categorical context offers them a functional or conceptual role. This provides, among other things, a theoretical framework within which the development of part/whole relationships can be understood, as well as the circumstances under which hierarchies occur in conceptual representations.

The theory of unit ontogeny presented in this paper may be considered an extension of the principles of concept learning. The arbitrary situation of the Martian rocks could be comparable to the kind of learning that takes place in infants when they start forging a conceptual system (Baldwin, 1989; Clark, 1973; Schyns, 1990). The conjunction of perceptual constraints with a context of categorization seems to be a promising framework to ground the development of a simple conceptual system on perceptual foundations. Clearly that will not account for all concepts, but these simple principles may account for the genesis of object concepts from "scratch."

Acknowledgments

This work was supported in part by NIMH grant MH 41704. The authors would like to thank Heinrich Bülthoff and Shimon Edelman for useful help. Gregory Murphy is now at the University of Illinois.

References

Baldwin, D. A. 1989. Priorities in children's expectations about object label reference: Form over color. *Child Development* 60:1291-1306.

Biederman, I. 1987. Recognition-by-components. A theory of image understanding, *Psychological Review* 94:115-147.

Biederman, I., & Ju, G. 1988. Surface vs. edge-based determinants of visual recognition. *Cognitive Psychology* 20:38-64.

Binford, T. E. 1981. Inferring surfaces from images. *Artificial Intelligence* 17:205-244.

Bertero, M., Poggio, T., & Torre, V. 1987. Regularization of ill-posed problems. AI memo. 924. AI Lab, MIT.

Braunstein, M. L., Hoffman, D. D., & Saidpour, A. 1989. Parts of visual objects: An experimental test of the minima rule. *Perception* 18:817-826.

Clark, E. V. 1973. What's in a word? On the child's acquisition of semantics in his first language. In T.E. Moore (ed.), *Cognitive development and the acquisition of language*. New York: Academic Press.

Hoffman, D. D., & Richards, W. 1985. Parts in recognition. *Cognition* 18:65-96.

Koenderink, J., & van Doorn, A. 1982. The shape of smooth objects and the way contours end. *Perception* 11:129-137.

Marr, D. 1982. *Vision*. San Francisco: Freeman.

Marr, D. & Nishihara, K. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society London B* 200:269-294.

Murphy, G. L. in press. Parts in object categories. *Memory & Cognition*.

Pentland, A. 1989. Part segmentation in object recognition. *Neural Computation* 1:82-91.

Poggio, T., Torre, V., & Koch, C. 1985. Computational vision and regularization theory. *Nature* 317:314-319.

Rosch, E., & Mervis, C. B. 1975. Family resemblances: Studies in the internal structure of cateogries. *Cognitive Psychology* 7:573-605.

Schyns, P. G. in press. Concept acquisition in a modular neural network architecture. *Cognitive Science*.

Smith, E. E., & Medin, D. L. 1981. *Categories and concepts*. Cambridge, MA: Harvard University Press.

Tversky, B., & Hemenway, K. 1984. Objects, parts, and categories. *Journal of Experimental Psychology: General* 113:169-193.