Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Domain-Specificity in Shape Categorization and Perception

Permalink

https://escholarship.org/uc/item/0274j4kk

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN

1069-7977

Authors

Balas, Benjamin J.
Tenebaum, Joshua B.

Publication Date

2004

Peer reviewed

Domain-Specificity in Shape Categorization and Perception

Benjamin J. Balas (bjbalas@mit.edu)
Joshua B. Tenenbaum (jbt@mit.edu)
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139-4307 USA

Abstract

We examine the influence of domain-specific knowledge on the process of learning and representing simple visual categories. Depending on whether subjects' construe simple objects as living kinds or machines, they show differential sensitivity to the importance of basic shape features. In particular, subjects who treated the objects as machines placed less importance on coarse shape differences unrelated to the described function of the objects in both categorization and productive (drawing) tasks. These findings suggest that domain-specific functional understanding of objects may influence the formation of shape categories, and perhaps the perception of these shapes.

Introduction

It is typical in the study of pattern recognition to discuss the classification of objects as though they are a relatively homogeneous class of stimuli. Many sets of 2D and 3D shape primitives have been proposed, all with the goal of describing how a wide variety of forms can be catalogued by a small set of simple building blocks (Hoffman & Richards, 1986; Biederman, 1987). Special-purpose mechanisms are seen as add-ons that are only used to process particularly interesting and ecologically significant stimuli. The nature of these specialized processes has also been explored, and there is much debate over what defines a "special" class of stimuli (Yin, 1969; Diamond & Carey, 1986; Gauthier & Tarr, 1997).

The suggestion that object recognition only employs unique processes for a small minority of object classes stands in stark contrast to work concerned with the formation and application of domain-specific theories (Carey, 1985). It has been suggested that knowledge of particular domains (like biology, physics, or psychology) may substantially affect the reasoning employed in a range of tasks. Given that theory-based reasoning may guide performance in complex scenarios, it may also be possible that human observers possess theories concerning the visual properties of objects in various domains that affect the way they recognize and represent object categories. Should this prove to be the case, it may suggest that general accounts of object recognition are too coarse, in that they fail to consider the richness of subjects' visual knowledge of a particular object category.

It is important to establish exactly what we mean when we suggest that subjects may possess theories about visual properties of objects that affect perception. We envision a simple hierarchy of object knowledge (Figure 1) ranging from high-level theories to low-level perception of shapes. At the top is amodal knowledge of particular domains. At this level (L1), abstract facts about objects in a broad domain (e.g., "artifacts") are stored propositionally (as in, "Machines are often built in factories to precise specifications"). At the second level (L2), object categories are defined in terms of the visual properties shared by members of common groups. This knowledge could be symbolic ("Tigers have stripes") or encoded in terms of visual measurements ("Tigers have lots of contrast energy at a particular spatial frequency"). Finally, at the bottom level (L3) lies the representation constructed in perceiving an object as a category instance, expressed in terms of its visual features such as size, shape, color, etc. Our work set out to explore the possibility that abstract knowledge at the highest level of this hierarchy could affect representations underlying both shape categories (L3 → L2) and shape perception (L3 → L1).

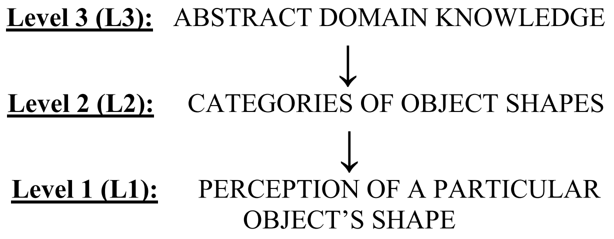

Figure 1: A schematic view of the relationships between domain theories and perception.

Previous experiments by Goldstone (1994) have demonstrated the influence of categorization on perception (L2 → L1). After forming categories of simple objects based on perceptual attributes (like luminance), subjects' ability to discriminate between luminance levels was altered to support the newly learned perceptual groups. No abstract domain knowledge (L3) was implicated in these studies. Kelemen and Bloom (1994) demonstrated an influence of domain knowledge on the representation of shape categories (L3 → L2), but did not look at influences down to L1. Their stimuli were uniform circles that could vary both in size and color. Subjects who were told the circles were microscopic animals preferred to categorizethem according to color, while those who construed them as machines preferred to categorize based on size.

Visual categorization relies on the ability to learn object attributes that vary more across categories than within them, and to accurately measure those attributes in new images.

Knowing that an object is a member of a particular domain (say “living things”) may bias the observer to expect particular patterns of variability. The Kelemen and Bloom study, as well as Keil’s work (1998), have both shown that color is expected to be useful for categorizing living things. We suggest that different aspects of shape may also be differentially weighted to categorize objects in different domains. The abstract (Level 3) knowledge observers possess about animals’ growth and movement as well as knowledge about artifact construction may lead to influences on representations of shape categories (Level 2), and the ability to make fine-grained perceptual distinctions at Level 1.

Another possibility relevant to the categorization of living and non-living things concerns the influence of functional information on shape perception. It has been proposed that function helps set the core meaning of artifact concepts, and perceived shape may be relevant to artifact categorization primarily to the extent that it supports a functional interpretation (Bloom, 2000). A theory of non-living things may not induce any *a priori* preferences for particular visual features, but rather, flexibly bias resources towards functionally relevant information. Function may also influence which shape features are perceived to be important for biological categories, but in different ways. Landau et al. (1998) have shown that people will be more tolerant of certain non-rigid shape variation when classifying objects construed as animals (relative to those construed as artifacts).

In sum, our studies here ask two main questions not addressed in previous work. First, how far down in the hierarchy of Figure 1 do domain-specific conceptual influences extend? In particular, do they extend down to perceptual representations of individual objects? Second, what is the range of conceptual influences? Specifically, is there a role for functional understanding in forming representations of shape categories and individual shapes? In the spirit of Kelemen and Bloom, we have created novel 2-D shapes that we have named “Marcons” and “Draxels.” These two populations of objects both consist of ellipses whose perimeters have been modified to contain sinusoidal bumps (Figure 2). The result is a set of stimuli that can be distinguished both by coarse shape information (ellipse eccentricity) and fine details (the frequency and/or amplitude of the bumps). This allows us to explore subtler aspects of object appearance than previous studies have done, and provides us with the ability to easily attribute function to shape properties.

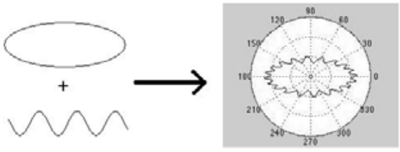

Figure 2: Constructing a Draxel.

In the experiments we present, subjects were asked to learn the distinction between Draxels and Marcons. While both groups are given identical information about the behavior of these objects and the function of particular features, some subjects are told that the stimuli are nanomachines and others are told they are amoeba-like animals. In Experiment 1, we use productive data (subjects’ drawings) to determine whether domain knowledge influences perceptual representations. Drawings are a particularly useful tool in that they require subjects to make their perceptual representations explicit. In Experiment 2, we look for the effects of domain knowledge on category representations through a more controlled categorization task.

Experiment 1

We begin by presenting the results of a learning task in which subjects learn to discriminate between Marcons and Draxels. We assess the nature of their post-learning representations of the two kinds of object by collecting drawings from all subjects, and examining the extent to which differences in elongation and ‘bumpiness’ are expressed.

Methods

Stimuli 32 Marcons and 32 Draxels were hand-drawn on 2 ⅛” x 2 ¾” cards for this task using Crayola™ magic markers. A stencil was used to enforce a major/minor axis ratio of 2.9:1 for Draxels compared to a 2:1 ratio for Marcons. Bumps were applied to the perimeter of each ellipse such that Draxels contained 2.9 bumps/cm and Marcons contained 2.0 bumps/cm. The length of the major axis could take on one of 8 values for both Draxels and Marcons, and a dual-color contour was applied to the perimeter of the finished figures. The figures were also depicted at four different orientations (0, 90, +20, -20 degrees from vertical). No conjunction of color, size, and orientation was diagnostic of object identity, leaving only ellipse eccentricity and bump frequency as useful criterion for discrimination.

Subjects 24 naïve subjects (9 men, 15 women) were recruited from the MIT community to participate in this task. Subject age ranged from 18-40 years of age.

Procedure Subjects were initially presented with a brief introductory paragraph explaining that Marcons and Draxels were either unicellular organisms (Animals condition) or nanomachines (Machines condition). In both cases, the two kinds of object were said to participate in “agricultural revitalization” by grabbing onto various chemical compounds with their bumps, and redistributing them across depleted soil. Subjects were told that Marcons and Draxels were quite similar, but that experts could identify them very accurately despite the range of individual sizes, shapes, and colors in which the objects appeared.

After reading this paragraph, subjects were shown one example each of a Marcon and Draxel on an 8 ½" x 11" placemat. These two items were the same dimension as the largest exemplars in the stimulus set mentioned previously, but were depicted with a novel dual-color contour. Subjects were presented with the shuffled deck of 64 cards, and asked to sort the pile into two stacks such that Draxels were on one side and Marcons on the other. Subjects were permitted to take as much time as they liked to sort the cards.

After each round of sorting, the experimenter determined what false classifications were made, and presented these items to the subject for further study (grouped underneath their proper place on the sorting mat) before they were shuffled back into the deck for the next round. Subjects sorted the cards until they made fewer than 8 errors, or until they had sorted through the entire deck 4 times.

When the card-sorting task had been completed, subjects were then presented with new instructions asking them to produce drawings of Draxels and Marcons. 8 examples of each object were requested, with the additional instruction that their drawings should depict what they believed "typical" members of each category looked like, and that their set of 8 drawings should attempt to cover the range of variations that existed within each category. After completing their drawings, subjects were given a brief questionnaire asking them to rate on a 1-10 scale how important were various visual features in their concepts of these two classes, and to enumerate in free-response style the differences they perceived between the two kinds of objects.

Results

Learning Rates To determine if either group showed a particular affinity for learning to discriminate between Marcons and Draxels, we examine the number of errors made by each group after their first and second rounds of sorting. These two rounds are of particular importance in that they indicate to what extent the task is difficult with only one example of each object type and how much improvement each group undergoes by viewing a population of labeled examples.

A two-way ANOVA, with sorting round and domain as factors, revealed only a main effect of sorting round ($p < 0.05$). Subjects improve from across rounds, but neither group was particularly better at performing the discrimination between Marcons and Draxels, nor benefited more than their counterparts from receiving multiple labeled examples after their first round of sorting. We note that of the 24 participants who performed this learning task, 6 subjects (3 Animals, 3 Machines) were unable to reach our performance criterion of fewer than 8 mistakes after 4 rounds of sorting. We take this to mean that the difficulty of this initial task was intermediate, and unrelated to the domain of the objects.

Post-Learning Questionnaire Subjects' responses to the post-experiment questionnaire were analyzed to determine if there were differences between the Animal and Machine groups' explicit feature preferences. A two-factor ANOVA was run on subjects' ratings of the importance of shape differences between the two object categories, with feature type as one factor (elongation v. bumps) and subject group (animal v. machine) as the other factor. No main effects were found in the analysis, but a significant interaction ($p < 0.05$) was found between feature type and subject group. Subjects in the Animal group rated eccentricity as a more important feature, compared to subjects in the Machine group who preferred to use the bumps. (Table 1)

Table 1:
Mean ± SD ratings of feature importance (1-10 scale)

	Animals	Machines
Elongation	8.3 ± 4.1	6.5 ± 2.6
Bumps	6.8 ± 2.0	8.8 ± 3.0

Subjects' Drawings Using these responses as a guide, we turn next to the drawings of Marcons and Draxels produced by subjects after learning. (Figure 3) The eccentricity and number of bumps/cm for each figure was measured by first inscribing the largest ellipse possible inside the bumpy contour. If a particular drawing was sufficiently irregular that this proved impossible, that figure was excluded from the analysis. Only 4 out of 384 drawings (all from different subjects) were excluded in this fashion. Eccentricity was determined for each figure by measuring the major and minor axes of the inscribed ellipse, and the number of bumps/cm was determined by counting the number of bumps on the drawing and dividing by its perimeter. We use the YNOT approximation of the perimeter of an ellipse (Maertens & Rousseau, 2000) here, which was also used in the creation of the stimuli.

For each subject, we then compute the mean values of both eccentricity and bumps/cm for Marcons and Draxels across all eight drawings. The difference between these means is taken for each feature type, and divided by the maximum standard deviation of that feature within an object population. In this way, we express the differences in Marcon and Draxel shapes as a function of the separation and spread of the populations produced by each individual subject(Table 2).

Table 2:
Normalized differences between Marcon and Draxel features (Mean ± SD)

	Animals	Machines
Elongation	1.93 ± 0.98	0.79 ± 1.04
Bumps	2.42 ± 1.55	2.31 ± 1.27

A two-way ANOVA was performed on these measurements, using feature type and subject group as factors. A main effect of feature type was found, (bumps > eccentricity, $p < 0.05$) with a marginally significant effect of subject group (Animals > Machines, $p < 0.08$). No significant interaction was found. However, to tease apart the contributions of each feature type to the weak effect of perceived domain, we conducted a further analysis for simple main effects across the Animal and Machine groups. We find in this analysis that subjects in the Animal condition expressed differences in eccentricity significantly more than subjects in the Machine condition ($p < 0.05$), while no such difference exists for the expression of bump density ($p > 0.8$).

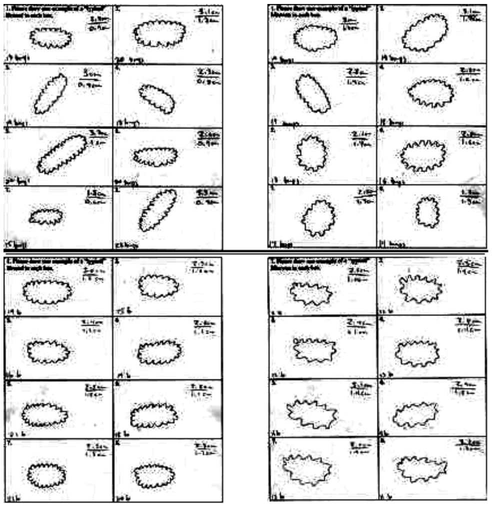

Figure 3: Examples of Draxels (left) and Marcons (right) created by subjects in the "Animals" condition (top) and the "Machines" condition (below). Note the lack of elongation differences in the lower drawings.

Discussion

Subjects' ratings of feature importance indicate an intriguing interaction between domain and the visual features perceived to be important. The drawings produced after completing the learning task show significantly more exaggeration of bumps across both domains, as well as significantly more expression of elongation differences in the Animal condition compared to the Machine condition. Bumpiness was not significantly more exaggerated in the Machine group compared to the Animal group, as one might expect from the interaction in subjects' ratings, yet there are some interesting qualitative differences in the way differences in bumpiness are expressed. Subjects in the Machine condition often pointed out subtle differences between the two categories that were not expressed or described by subjects in the Animal condition. Subjects in the Machine condition often pointed out aspects of bump symmetries and asymmetries that they felt were important to the task, and often included these details in their drawings (Figure 4). Though these features are not directly related to the difference in bump frequency, and therefore did not contribute to our quantitative analysis, it is interesting to see how these features appear in the drawings of Machine subjects while remaining almost wholly absent from the drawings of Animal subjects. It may be that more aspects of bump shape need to be explicitly considered when constructing stimuli and examining subjects' drawings.

We note that both of these effects may be a consequence of the introductory scenario given to subjects at the beginning of the task. By indicating that the bumps had a particular functional importance, the representation of bump differences may have been weighted more heavily in both groups. It is interesting to note that in the Machines condition, this seems to have resulted in an overall tendency to ignore additional shape differences. The significance of elongation differences in this group may have been compromised by the direction of functional information away from these features.

Experiment 1 provides us with direct evidence that domain knowledge can influence shape categorization (L3 → L2, in Figure 1). Evidence for effects of L3 on L1 (perceptual shape representations) is only indirect, insofar as subjects are thought to produce drawings by translating a Level 2 representation into perceptual primitives. This is an important distinction, since Level 2 representations may be more symbolic ("Draxels are skinnier than Marcons") rather than truly visual ("Draxel elongation is about 3:1"). A valuable next step would be to measure perceptual abilities directly using psychophysical methods (e.g., Goldstone, 1994), thereby establishing a more direct L3 → L1 link.

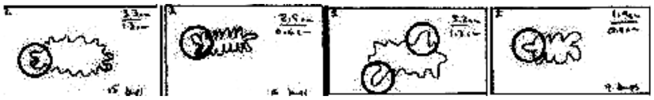

Figure 4: Over-expression of symmetry and asymmetry in "Machines" subjects. Draxels are at left, Marcons right.

Experiment 2

In our second task, we pursue possible domain-specific differences in shape processing using a more controlled task. Rather than relying on subjects' drawings, we use a triad task in which subjects are asked to classify a new object given labeled examples of our two categories. This second task allows us to more closely align our findings with the color/size asymmetry noted by Kelemen and Bloom, and allows us to balance the freedom given to subjects in Experiment 1 with a more constrained environment.

Subjects in this task are asked to make classifications given first a single example of each class, and then multiple examples of each class. Use of a single example allows for a relatively pure measure of subjects' prior beliefs concerning what visual features are important to the task, while multiple examples gives us some sense of what (if anything) changes when subjects are given evidence of what features vary across a population of objects. We look for evidence of the interaction between domain and perceived relevance of features suggested by our previous experiment, while also investigating whether or not the domain-general preference for expressing differences in bumpiness persists in a task in which the two feature types are pitted directly against one another. In terms of our proposed hierarchy, we are looking for influences of Level 3 on Level 2, without examination of Level 1 representations in any form.

Methods

Stimuli For this task, a subset of the original stimuli were used. 4 Draxel/Marcon pairs of different color schemes were selected, with 2 pairs taken from the largest items in the original set, and the remaining 2 pairs being of intermediate size. All of these stimuli appeared at horizontal orientation to ensure that subjects would only consider shape information when making their decisions.

Additionally, two novel stimuli were created for each pair to serve as 'unclassified' items. Each novel item matched both their parents' color and size, but would match their Draxel parent for one feature type (say, elongation), and their Marcon parent for the other (bump frequency, in this example). The two stimuli in each "hybrid" pair were complementary, such that each "Draxel elongation/Marcon bumps" item had a partner with the opposite pattern of feature inheritance.

Subjects 64 subjects participated in this task, drawn from the MIT community.

Procedure Each subject read the same short description of Draxels and Marcons presented in our first task. Subjects were then told that they were being asked to help classify a new object that was either a Draxel or a Marcon, but currently unlabeled. Subjects were told that their initial answer would be based on the observation of only one example each of a Draxel and a Marcon, and that after their first response they would be given multiple examples to look at regardless of their first answer. To ensure that subjects did not feel undue pressure to change their answer given new information, all stimuli to be presented to the subject were laid face-down on the table before any responses were solicited. In this way, we minimize the possibility that subjects' might consider the new stimuli as additional information selected by the experimenter to guide them to a particular answer.

Subjects were first shown one example each of a Draxel and a Marcon, (Figure 5) matched for all attributes except eccentricity and bump frequency. A new item was then shown, drawn from one of the two hybrid stimuli created for that initial pair of stimuli. This third item matched the color and overall scale of both original examples. Subjects were then asked to classify the new item as a Draxel or a Marcon.

After their response was recorded, they were shown the three remaining examples of Draxels and Marcons. Subjects were asked a second time for a response, which was then recorded prior to rewarding the subject with M&M's for volunteering. The initial pair of examples displayed, as well as the particular probe used as the third stimulus were balanced across subjects, as was the left-right arrangement of Draxels and Marcons.

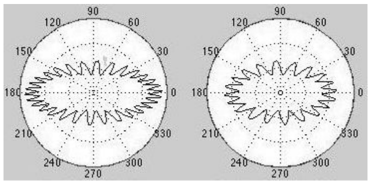

Figure 5: Digital versions of Draxels (left) and Marcons (right). Actual stimuli were hand-drawn, but we present these more regular, schematic images for clarity.

Results

On subjects' first judgments, 21 subjects in the Animals condition choose to categorize the novel stimulus by referring to the bump frequency of the examples, with 11 favoring eccentricity. In the Machines condition, 25 subjects used the bumps as a diagnostic feature, with 7 participants using the elongation of the ellipse. A chi-squared test reveals that these two distributions do not differ from each other. However, an additional goodness-of-fit test shows that the distribution of responses in the Machines condition differs from chance ($p < 0.05$), while the responses of subjects in the Animals condition do not.

Moving on to consider the responses given by subjects after viewing multiple examples, we see that 19 subjects in the Animals condition favor bump frequency for classification, compared to 13 individuals who prefer to use eccentricity. In the Machines condition, the corresponding numbers are 27 and 5 subjects respectively. Unlike the single-example data, these two distributions do differ from one another by a chi-squared test ($p<0.05$). Also, the responses of the Machines subjects differ from chance while the other responses do not.

In each group of subjects, only two subjects changed their mind when presented with multiple examples of the object classes. This indicates that our efforts to minimize undue pressure on the subjects to change answers suceeded, and that subjects were confident in their classifications.

Discussion

In our triad task, we continue to see domain-specific differences in the preference for different aspects of object shape. In correspondence with our results from Experiment 1, we see that subjects in the Animals condition do not significantly prefer one feature type to another. Moreover, as a whole, subjects who perceive the objects as machines have a greater tendency to ignore differences in elongation in favor of differences in bump frequency.

Additionally, in both conditions we see the same overall bias towards using bumps that characterized subjects' drawings in Experiment 1. The distribution of responses in the Animals condition is never different from chance in this study, making it difficult to draw a firm conclusion on this matter. However, the qualitative agreement between the pattern of results obtained here and those obtained in Experiment 1 suggests that the effects we observe genuinely reflect people's representations of these categories, rather than task artifacts.

General Discussion

Two studies demonstrated that perceptual representations of simple visual shape categories differ depending on whether the stimuli are conceived of as living or non-living things. Our results show a difference in how various shape features are weighted across domains, but they do not yet speak to the origins of those differences. The difference could be simply one of spatial scale, or it could reflect deeper differences in how function or shape variation is conceptualized in different domains. We conjecture that the functional description of the objects plays a driving role in the effects we observed. In a pilot study, we have presented subjects with the same stimuli and task of Experiment 2, eliminating only the instruction sheet's explicit description of bump function. Preliminary results indicate that this manipulation eliminates the domain effects we see here, in that both "animal" and "machine" subjects appear equally to favor the bumps for categorization, and at a level equal to "machine" subjects in Experiment 2. Hence, for these shape stimuli, functional information appears to exert a greater influence when the objects are construed as living kinds, and its primary role appears in *decreasing* the distinctiveness of the bumps for classifying these two kinds of microorganisms. The bumps may be seen as pseudopods, non-rigid appendages of the microorganisms that are used to grab compounds but are not essential shape features that are stable across time in an individual, let alone as a distinctive feature for classifying objects into kinds. This interpretation is consistent with our data, although subjects' post-experiment surveys did not mention any explicit reasoning of this sort. Further research is needed to determine precisely how functional knowledge and shape representations interact here, but it is intriguing to speculate that intuitive domain theories are guiding implicit inferences concerning the possible dynamic aspects of simple shapes.

The potential effects of exposing subjects to multiple examples of an object class are also worth exploring further. In this task, we see very little change in subjects' behavior from one response to the other, but the number of observations they are allowed to make is still quite small. If presented with an extremely large population of Marcons and Draxels, subjects might undergo a more profound evolution of shape processing strategies. We note that in the Experiment 1, subjects gained far more experience with Marcons and Draxels than those that participated in our triad task. Subjects in the Animals condition also expressed an explicit preference for elongation rather than bumps in that task, which we did not see in Experiment 2. The difference between these two patterns of response may be related to the size of the observed population of each object class. The possible interaction of statistical reasoning given a population of novel objects and prior beliefs about feature relevance may prove to be a rich area for further research.

Conclusions

Taken together, these studies demonstrate that the categorization and perception of simple visual stimuli may be affected by the domain in which these objects are construed, functional reasoning about object properties, and by the amount of experience one has with a particular stimulus set. Understanding all of these influences, both as separate mechanisms and as a coherent whole, may lead to a richer understanding of human object categorization and the perception-cognition interface.

Acknowledgments

The authors would like to thank the members of the Sinha lab, as well as Erin Conwell, who put up with talking about fictitious unicellular organisms for far too long. An anonymous reviewer also provided many helpful comments. BJB was supported by an NIH pre-doctoral training grant. JBT was supported by the Paul E. Newton chair.

References

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

Bloom, P. (2000). *How Children Learn the Meanings of Words.* Cambridge, MA: MIT Press

Carey, S. (1985). *Conceptual Change in Childhood.* Cambridge, MA: MIT Press.

Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology, 115(2),* 107-117.

Gauthier, I. & Tarr, M.J. (1997). Becoming a 'greeble' expert: Exploring the face recognition mechanism. *Vision Research, 37(12),* 1673-82.

Goldstone, R.L. (1994) Influence of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123,* 178-200.

Hoffman, D. & Richards, W. (1986). Parts of Recognition. In A.P. Pentland (Ed.), *From Pixels to Predicates,* Norwood, N.J: Ablex Publishing Corporation.

Keil, F.C. (1998). Cognitive science and the origins of theoretical knowledge. In W. Damon & R. Lerner (Eds.), *Theoretical Models of Human Development.* New York, NY: John Wiley & Sons.

Kelemen, D. & Bloom, P. (1994). Domain-specific knowledge in simple categorization tasks. *Psychonomic Bulletin & Review, 1(3),* 390-395.

Landau, B., Smith, L, & Jones, S. (1998). Object perception and object naming in early development. *Trends in Cognitive Science, 2(1),* 19-24.

Maertens, R. & Rousseau, R. (2000). Een nieuwe benaderde formule voor de omtrek van een ellips, *Wiskunde & Onderwijs, 26,* 249-258.

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141-145.