Normative Representation of Objects: Evidence for an Ecological Bias in Object Perception and Memory

Talia Konkle (tkonkle@mit.edu)
Aude Oliva (oliva@mit.edu)
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

Objects in the world can be viewed at almost any distance, and so can subtend a variety of visual angles. An apple for instance can be held in the hand if you are eating it, or be perceived in a fruit bowl from a few meters away. When interacting physically or perceptually with a given object, our brain is exposed to a particular distribution of distances and their corresponding visual angles. To which extent is our visual representation sensitive to these statistical regularities? In this paper, we report results demonstrating that both perception and memory of an object's visual size are influenced by a "normative size." The normative size corresponds to the optimal size for viewing the object based on observer's reports. We show that the perception of real world objects is implicitly sensitive to a normative visual size and that this norm is strongly correlated with the actual size of objects in the world (Experiment 1). Using a size-memory task and a change detection task, we show that long-term and short-term memory errors for an object's size are systematically biased towards the normative size (Experiment 2 and 3). Altogether the results support the claims that perception of objects is sensitive to a normative size and that object memory is biased toward this perceptual norm.

Introduction

"For each object, as for each picture in an art gallery, there is an optimum distance from which it requires to be seen, a direction viewed from which it vouchsafes most of itself: at a shorter or greater distance we have merely a perception blurred through excess or deficiency."
- Merleau-Ponty

While watching out for incoming cars from a distance or getting your keys while approaching your car, we see each object in the world from a variety of viewing distances. During natural vision, the physical size of an object in the world and your viewing distance to that object determine the angle it subtends in your visual field. Objects can appear at almost any size in the visual field depending on how close or far you stand in relation to them. However, objects typically are interacted with at a range of specific distances, which correspondingly leads to a distribution of visual angles at which they are most often viewed (Figure 1a). Does the mode of visual experience with an object lead to a perceptual expectation of when the object is the "right size" in one's visual field?

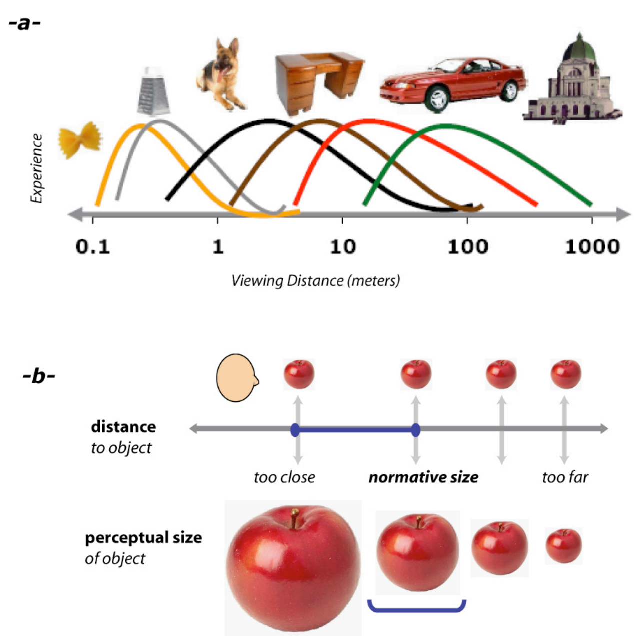

Figure 1: (a) Hypothetical distribution of visual experience with objects, as a function of object sizes in the world and viewing distances. (b) The concept of a "normative size" is the perceptual size at which an object, here the apple, is seen as neither "too close" nor "too far."

In the current study, we explore the notion that each object has a perceptually privileged size, which we term the "normative size" (figure 1b). Further, we explore the possibility that, if perception is sensitive to a normative size, then memory for that object might also be biased toward a norm. In Experiment 1, observers were presented with a picture of an isolated object and adjusted the visual angle of the object until the object was the "right size" on the screen. Despite the subjectivity of the task, the selected size for each object was remarkably consistent across observers. In experiment 2, observers were presented with pictures of objects which subtended a range of visual angles. Afterwards, participants adjusted the visual angle of the object so as to match the size viewed at learning. The remembered objects showed a systematic bias towards the normative size found in Experiment 1. Experiment 3 measured the sensitivity to detect when an object changed in size. The object could either change to be slightly larger in visual angle or slightly smaller in visual angle. Results show that when an object changed size in a direction towards the normative size, this change was more difficult to detect then when the object changed size away from the normative size. Experiment 2 and 3 both reveal predictable biases in memory for an object's presented size, biased toward the normative size for that object.

Experiment 1: Existence of a statistically constrained perceptual norm

The first experiment examines if there is a perceptually privileged size to view an object. The goal of this experiment was to answer the two following questions: (1) how consistent are people when they choose an intuitively "best size" to view a real-world object on a monitor? (2) what is the relationship, if any, between the selected size of each object and the actual physical size of that object in the real world? For clarity, the term "visual-size" will refer to the visual angle of the object on the screen, and the term "real-world size" will refer to the physical size of the object in the world (e.g. airplanes have a large real-world size, paperclips have a small real-world size, but both these objects can be any visual-size, depending on observer's distance to the object).

One prediction is that observers will choose the same visual angle for each object, either by filling the fovea or parafovea, reflecting a visual acuity constraint, or by filling the monitor to the edges, reflecting a framing/bounding box constraint. An alternate prediction is that observers will select a size that is correlated with the real-world size of the object.

Method

Object norming experiment A: Seven observers with normal or corrected vision (age range: 18-35 years old) were presented with 40 norming trials, each consisting of one object shown on a blank background. The objects were all color photographs of real world objects (see examples in Figure 2) selected from a commercial database (Hemera Photo-Objects, Vol. I & II) to correspond to a variety of object sizes, poses, and surface characteristics. Observers could freely increase and decrease the visual angle of the object by pressing the up and down arrow keys. They were seated approximately 57 cm from a 20 inch monitor, and the range of visual angles each object could subtend was 0 to 30 degrees. Observers were given the following instructions: "For each object, select the best size to see it. Intuitively, when the object is at the smallest extreme, this is too small. When the object is at the largest extreme, this is too big. Use the keys to adjust the object's size and then click when the size of the object is not too big or too small, but just right. There is no right answer, so select the size that is best for you." The order of objects was randomized across observers.

Object norming experiment B: Additionally, we repeated the norming task with 100 objects using a sliding-mouse method of adjustment instead of the key press described above. Six adult observers with normal or corrected vision (age range: 18-35 years old) participated. 100 objects were used in this experiment which included the 40 objects from object norming experiment 1. To adjust the visual angle of the object, participants moved the mouse up and down, and clicked to select the intuitively right size. All other procedures and instructions were the same.

Size sorting experiment: Six observers gave ground truth rankings on the real-world size of 100 objects. This was accomplished using a hierarchical sorting method (Oliva and Torralba, 2001). Thumbnails of the 100 objects were put on a 30" monitor and participants iteratively divided the images into two groups by dragging and dropping the thumbnails, until there were 8 groups of objects, ranked by real-world size.

Results

In object norming Experiment A, the selected visual angle of the height dimension for each object ranged on average from 4 degrees (e.g. peanut, thumbtack) to 14 degrees (e.g. Arc de Triumph, crane). We refer to the average selected size for each object as the "normative size." Interestingly, despite the subjectivity of the task, inter-rater reliability for each object across observers was remarkably consistent, ($R=.9$, $p<0.01$). This correlation indicates consistency in which objects observers set the smallest and which objects they set the largest. However, two observers appeared to use more of the total range, making several very large or very small settings. For the purposes of experiment 2 and 3, we set the normative size of each object equal to the average size selected, excluding these two of the seven subjects. This was done solely to select a best normative size for each object that would be representative of most observers in the absolute value of the normative sizes. In fact, without the two observers included the inter-reliability measure was slightly lower ($R = .88$, $p<0.05$).

In the object norming Experiment B with 100 objects, we found similar results. Inter-rater reliability was again very high ($R=.7$, $p<0.05$). For the 40 objects that were tested in both experiments, the average selected sizes were within 3 degrees for all items, with no significant bias to be either smaller or larger.

Next, we examined if there was a relationship between the visual angle of the normative size and the real-world size. Six observers arranged the objects into 8 categories ranked by real-world size. An objects size-rank was taken to be the mode rank across observers. A rank of one is the smallest real-world size and a rank of eight is the largest (see figure 2). The number of objects per rank was not constrained to be equal, and across all ranks there was an minimum of 7 and a maximum of 23.

Using the selected visual angles from the norming experiment with all 100 objects, we averaged the normative size for the objects in each rank group, and plotted it against the rank size. This allows us to examine the correlation between the normative size and the real-world size. As shown in Figure 3, the visual-angle of the object was highly correlated with the real-world size rank of the object (R^2=0.96).

Figure 2: 100 objects were sorted into 8 groups by their real-world size. Example objects in each size rank are shown.

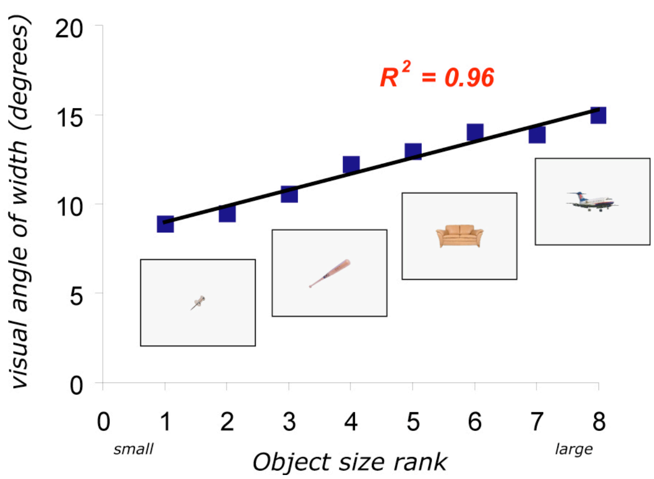

Figure 3. Correlation of visual angle of the normative size and real-world size ranking. As the real-world size of the object increases, the visual angle of the normative size also increases.

Discussion

The results show striking reliability for normative sizes across observers and across reporting methods, despite the subjectivity of the task to select the "intuitively right size" to see the object. This provides initial evidence for the existence of a perceptually privileged, or "normative size" for each object. Interestingly, the normative size was relatively small on the screen (4 – 14 visual degrees); thus participants were not simply filling the monitor with the object, which is suggestive that some acuity factors of the visual field may be also playing a role in the normative size. However, the normative size was very significantly correlated with real-world size (i.e. the larger the real-world size of the object, the larger the selected visual angle), indicating that the normative size is also influenced by knowledge or experience of these objects in the real world.

Experiment 1 gives a perceptual norm for each object, corresponding to the average size which observers selected as the best size to see it. If this size is indeed privileged in perception, then it might also be privileged in memory. Experiment 2 tests this hypothesis directly, and predicts that long-term memory for objects will be biased towards the normative size.

Experiment 2: Bias of Long-Term Memory towards the Normative Size

Previous literature on the boundary extension effect has shown that people tend to remember a picture of an object or scene as farther than it was originally viewed (Intraub & Richardson, 1989). A normative theory predicts that memory for an object is biased toward the norm, and that boundary extension is only one possible result of this memory bias. When an object is presented closer than the norm, the object should be remembered as farther away, consistent with the classic boundary extension effect. However, if the object is presented too far, a normative theory predicts that the object should be remembered as closer. This opposite effect has not been observed in studies on object memory using boundary extension paradigms.

Experiment 2 used a classical boundary extension paradigm in which a stream of objects were presented and observers were told that that they were going to be tested on their memory for these objects, without explicit instruction about the kind of memory test. Afterwards, subjects had to report the visual size of the object that was presented. As in boundary extension paradigms, the images can either be "close-up" or "wide-angle". In the current experiment, we selected object visual sizes as "too big" and "too small" relative to each object's normative size from experiment one.

Method

Twenty categorically unique objects from the set of 40 used in Experiment 1 were selected, uniformly across the eight real-world size groups. For each of the 24 participants, ten objects were randomly selected for the too-big condition and ten were randomly selected for the too-small condition. The experiment consisted of two phases. During the learning phase, each object was presented on the screen for 5 seconds with a 1 second inter-stimulus-interval. Participants were informed that after a learning phase of 20 objects, they would be "tested on their memory of the objects." As with the classic boundary extension paradigm, observers were *not* informed that the memory test was specifically going to be for the object's size on the screen. Following the learning phase, participants were presented with each of the 20 objects, one at a time in randomized order, and used the up and down arrows on the keyboard to resize the object to match the size they saw during the learning phase.

Object 'step-sizes' were linear steps in visual angle of the height dimension corresponding to approximately 1 visual degree in height increase per step. In the learning phase, each object was presented 5 steps larger or 5 steps smaller than its normative size. Key presses advanced the object size one step. For each object there were 40 possible step sizes. During the memory test, the object was initially presented jittered around the middle step position.

Results

Memory performance was quantified by calculating the number of steps between the object size selected during the testing phase and the object size presented during the learning phase. *Negative* numbers indicate that the object was reported as *smaller* than at learning (object contraction); *positive* numbers indicate that the object was reported as *larger* than at learning (object expansion).

Participants showed significant contraction (remembering a smaller object) for the objects presented larger than the normative size (too-big condition: $t(23)=2.45$, $p<0.05$). This is consistent with the known boundary-extension effect. Critically, participants also showed significant expansion (remembering the object bigger) for objects presented smaller than the normative size (too-small condition: $t(23)=2.80$, $p<0.05$, figure 4). None of the 24 subjects showed memory errors with the opposite trend. In an item analysis, 15 of the 20 objects presented "too big" during learning showed significant compression ($p<.05$) and 18/20 objects presented "too small" showed significant expansion ($p<.001$).

Discussion

Long-term memory for objects showed a systematic bias toward the normative size, both for objects presented too big and objects presented too small. When an object is seen larger than its norm, it is remembered as smaller (closer to the norm), whereas when an object is seen smaller than its norm, it is remembered as bigger (again, closer to the norm). Here, observers were not explicitly informed that they would be tested on memory for the objects visual-size. Further, observers were required to remember 20 objects before being tested. This shows that long-term memory for object size is biased toward the normative size, in the absence of explicit encoding of the size of the object. In other words, when you have to remember a number of objects and subsequently report their size, your memory errors are biased in a predictable direction toward the object's normative size.

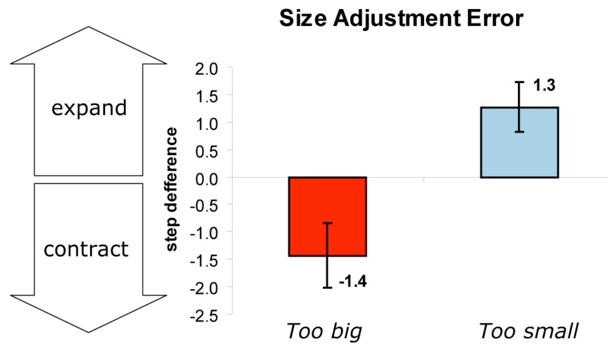

Figure 4. Deviation between presented visual size and selected visual size. Negative numbers indicate that the object was reported as smaller than at learning. Positive numbers indicate that the object was reported as larger than at learning.

Experiment 3: Bias of Short-Term Memory towards the Normative Size

The aim of Experiment 3 was to generalize the effects of the normative size to a situation where memory is tested immediately after the presentation, and where observers explicitly know they are being test on memory for size. A change detection paradigm is suitable to evaluate short-term memory of a single event (Luck & Vogel, 1997). Suppose an object is presented at a larger visual angle than the normative size. We predicted that memory for this object will be shifted slightly smaller, toward the direction of the normative size for that object. Thus, if the object is presented again at a slightly smaller visual angle than the first, this change should be more difficult to detect than if the object is presented again at a slightly larger visual angle. Similarly, suppose your first view of an object is smaller than it's norm; again, the normative theory implies that your memory of that object will be larger. Thus, if the object reappears slightly larger, this change should be more difficult to detect than if the object appears slightly smaller.

Put succinctly, we hypothesized that a change in size *toward the normative size* should be more difficult to detect than a change of visual-size *away from the normative size*. If there is no systematic bias in the memory of the first stimulus, then there should be no difference in detection if the object changes toward or away from the normative size.

Method

Twelve observers participated in the change detection task. On each trial, an object was presented for one second, masked for 200ms, followed by a blank screen for one second, and then re-appeared at the same or a different size. The object remained on the screen until participants pressed a key indicating whether the size of the object was the same or different. Forty objects were used in the experiment. Each object was repeated in 12 trials: on six trials the object was presented as too-big relative to its norm, and on six trials the object was presented as too-small. In two of the six trials, the object changed toward the norm, in two of the six trials, the object changed away from the norm, and in the remaining two trials, the object was presented at the same size. The first image of the object was presented at 5 steps smaller or larger than its normative size from experiment one. (See step description from Experiment 2.) The second image could change by 3 steps toward or away from the norm. Figure 5 shows an example object and the size of the changes.

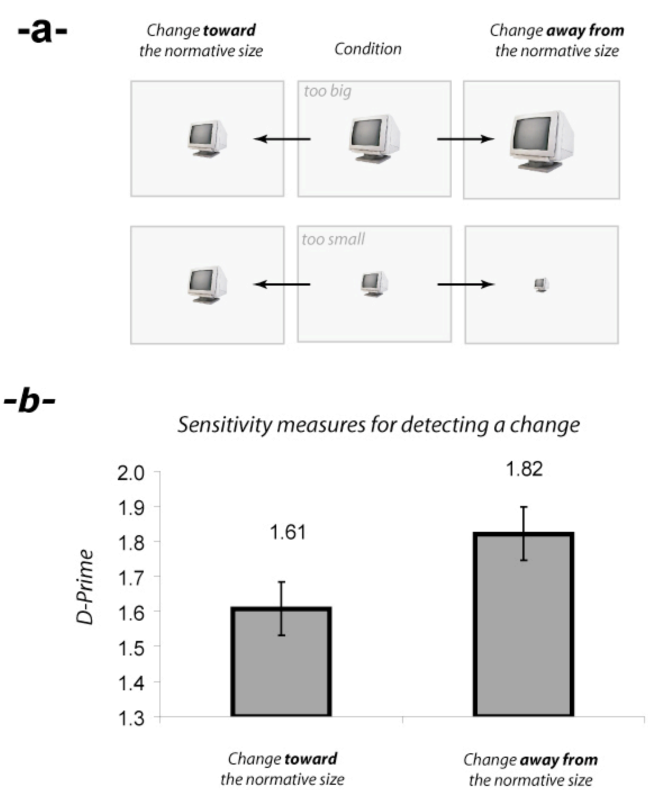

Figure 5. (a) Visual-angle changes are illustrated here for a sample object. The first stimulus is presented 5 steps larger or smaller than its normative size (middle column). The second stimulus changed towards the normative size (left column) or away from the normative size (right column). (b) Sensitivity to detecting a change. D-prime measures are plotted for the change-toward condition and the change-away condition. Changes toward the normative size were significantly more difficult to detect than changes away from the normative size.

Results

A measure of sensitivity (d-prime) was calculated for each type of change (change toward the norm, change away from the norm), by taking the z-score of the percentage of hits minus the z-score of the percentage of false alarms. The results are shown in figure 7. Participants were significantly less sensitive in detecting a change toward the norm compared to detecting a change away from the norm ($t(11)=2.78$, $p<0.01$). When the first stimulus was too-close, paired t-tests show that detecting a change toward the norm was significantly more difficult then when the change was away from the norm ($t(11)=2.20$, $p<0.05$). However, when the first stimulus was too-far, this trend did not reach significance ($t(11)=1.35$, $p=.2$).

Discussion

The results of Experiment 3 indicate that a change toward the normative size was significantly harder to detect than a change away from the normative size. This suggests that short-term memory is also biased toward the normative size, even in a situation in which the visual-size of the object is explicitly the dimension of interest for the task.

The d-prime measures of sensitivity to change were very high (all d-primes > 1), indicating that the change-detection task was easy to perform. This makes it even more surprising that any significant differences in the change-toward vs. change-away conditions were observed. While the one of the two individual paired t-tests did not reach significance, this is likely due to factors like the magnitude of the change in size and variation in individual subjects normative size.

General Discussion

Experiment 1 showed that for each object, there is a particular visual-angle, termed the normative size, which is *perceptually* privileged across observers. Experiments 2 and 3 demonstrated that *memory* for object size is biased in a systematic direction that is predicted by the normative size. Long-term incidental memory for object size is biased toward the normative size: objects that were presented larger than their normative size tended to be remembered as smaller, and objects that were presented smaller than their normative size tended to be remembered as larger. Short-term explicit memory for object size is also biased toward the norm: a change in object size toward the normative size was harder to detect than a change away from the norm. Taken together these experiments provide support for the following claims:

1. Perception of objects is implicitly sensitive to a normative size.
2. The normative size is related to the real-world size of objects.
3. Memory for object size is biased towards the normative size.

While this work demonstrates a normative concept for an object's size, this could be extended to other properties of real world objects, such as viewpoint, elevation, color, identity, or state. Indeed, research on canonical viewpoints

(Tarr et al, 1998) and color (Tanaka & Presnell, 1999), suggests that there are privileged perceptual views along other spatial and featural dimensions. Further, systematic memory errors towards a norm have implications for the nature of object representation, e.g. objects could be represented as a sum of their parts (Biederman, 1987), or as implied here, as deviations from their statistical mode. Some theories of efficient coding (Barlow, 2001) and prototypes (Rosch, 1981), are suggestive and consistent with this theory of object representation.

One striking result was the correlation between the visual angle of the normative size and the real-world size. This suggests that the process driving observers to select the "intuitively right size" is influenced by the natural statistics of the world (Gibson, 1979). One broad hypothesis is that the default representation of an object is the statistical mode of visual experience with that object along *any* relevant dimension. Whether an object's norm corresponds to the mode of visual experience with that object or object category is an open question.

The normative hypothesis poses an alternate explanation for the current interpretation of the boundary extension phenomenon. In the boundary extension effect, close-up scenes are remembered as farther away than they were actually perceived (Intraub & Richardson, 1989, Intraub & Bodamer 1993, Gottesman & Intraub 2002, 2003). While classically this effect is thought to be a phenomenon about the visual information at the edges of the scene, more recent evidence suggests this effect is driven by the central object (Bertamini et al, 2005). Our results suggest that boundary extension reflects a memory bias towards the normative size of the central object.

To date, these experiments have only been run on monitors that have clearly defined edges. How much do the present results, and those of boundary extension, depend on the presence of visual edges? Ongoing work is examining how this normative concept operates in an embodied context, where people walk towards and maneuver around real-world objects in a natural setting.

Acknowledgments

The authors wish to thank George Alvarez, Michelle Greene, and Barbara Hidalgo-Sotelo for comments on the manuscript. This research is supported by a NDSEG fellowship awarded to T.K. and a NSF Career Award grant to A.O.

References

Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12, 241-253.

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

Bertamini, M., Jones, L.A., Spooner, A., & Hecht, H. (2005). Boundary extension: the role of magnification, object size, context and binocular information. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1288-1307.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

Gottesman, C.V., & Intraub, H. (2002). Surface construal and the mental representation of scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 589-599.

Gottesman, C.V. & Intraub, H. (2003). Constraints on spatial extrapolation in the mental representation of scenes. View-boundaries versus object boundaries. *Visual Cognition*, 10, 875-893.

Intraub, H., and Richardson, M. (1989). Wide-angle memories of close-up scenes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 179-187.

Intraub H., & Bodamer, J.L. (1993). Boundary extension: Fundamental aspect of pictorial representation or encoding artifact? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 1387-1397.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279:281.

Oliva, A., & Torralba, A. (2001). Modeling the Shape of the Scene: a Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*, 42, 145-175.

Rensink, R. (2002). Change detection. *Annual review of Psychology*, 53, 245-77.

Rosch, E. (1981). Principles of Categorization. *Readings in Cognitive Science, a Perspective from Psychology and Artificial Intelligence.* In A. Collins & E. E. Smith, Morgan Kaufmann Publishers (pp. 312-322).

Tanaka, J.W., & Presnell, L.M. (1999). Color diagnosticity in object recognition. Perception & Psychophysics, 61, 1140-1153.

Tarr., M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, 1(4), 275-277.