Title

Completion in the Wild: Perception of 3D forms from cross-sections

Permalink

https://escholarship.org/uc/item/7c94q19x

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Gagnier, Kristin
Shipley, Thomas

Publication Date

2013

Peer reviewed

Completion in the Wild: Perception of 3D forms from cross-sections

Kristin M. Gagnier (kristin.gagnier@temple.edu)
Spatial Intelligence and Learning Center, Department of Psychology, 1701 N. 13th Street
Philadelphia, PA 19122 USA

Thomas F. Shipley (Shipley@temple.edu)
Spatial Intelligence and Learning Center, Department of Psychology, 1701 N. 13th Street
Philadelphia, PA 19122 USA

Abstract

Under conditions where an object is inside another object and only a single face is visible, is there a bias to assume smooth continuation of the surface straight back into the object? To examine the ability to estimate how features progress into a volume, participants viewed 16 pictures of everyday objects (rocks, food, wood) presented with only a single face visible (see Figure 1). Participants reported whether a highlighted region of a picture was present on the surface or extended into the object. If they perceived the region as extending in, they positioned a rod to indicate the angle. Surface responses were rare and instead participants' readily perceived 3D forms from 2D views. Inspection of frequency histograms revealed a systematic bias to estimate the angle of extension in the 80-110° range. This type of completion process suggests constraints on models of visual completion and has implications for STEM education, in particular, how students deal with ambiguity.

Keywords: perception of 3d volumes, amodal continuation, penetrative thinking.

Introduction

How does the visual system estimate 3-dimensional (3D) forms of objects embedded in other objects? While this question has received very little empirical attention (Chariker, Naaz, & Pani, 2011; Hegarty, Keehner, Cohen, Montello, & Lippa, 2007), we believe it is central to our understanding of amodal completion. Take for example the image shown in Figure 1. There is little doubt that the dark brown region (cinnamon swirl) is 3D and extends into the object, however from this 2-dimensional (2D) view or cross-section, the 3D shape is unknowable (note even if you knew the true 3D shape the answer would still be ambiguous within a mirror reflection because the image could come from either side of a cut). In order to infer the 3D shape of the region, one would have to have a view of the region from another side. Inferring shape from partial information is particularly important for disciplines that rely on 3D visualization (e.g., astronomy, neuroscience, geosciences). However, a common sentiment echoed by geoscientists and noted by Kali & Orion (1996) for geological stimuli, is that students neglect the ambiguity inherent in a single 2D view and instead are biased to assume that surface boundaries extend perpendicularly into solids.

The goal of this paper was twofold: 1) to examine whether people recognize that the 3D form of an object is unknowable from a single 2D view and 2) to examine whether participants do indeed exhibit a bias to assume smooth continuation of a surface straight back into the solid as suggested by the anecdotal reports of geoscience educators.

To answer this question we showed undergraduate psychology majors pictures of everyday objects such as food, rocks and wood. For each picture, a region was indicated with a colored line, as shown in Figure 1 top. Students indicated if the highlighted region was visible only on the surface or whether it extended into the object. Note that the answer to this question is in fact unknowable. While some of the objects are familiar to viewers (kiwi, bread, etc) and thus the overall shape of the object can be inferred, one cannot know how the cut was made or for unfamiliar objects, whether the indicated region continues into the object or is present only on the surface. To infer the 3D shape one would have to see more than one 2D cross-section of the object. If students saw the indicated region as extending in, they used a rod attached to an inclinometer to indicate the angle at which the region continued into the object (the correct answer for the bread is shown in Figure 1 bottom). We predicted that participants would have a strong sense that the indicated regions were 3D and would exhibit a bias to see the regions as extending back at a 90° angle relative to the ground surface.

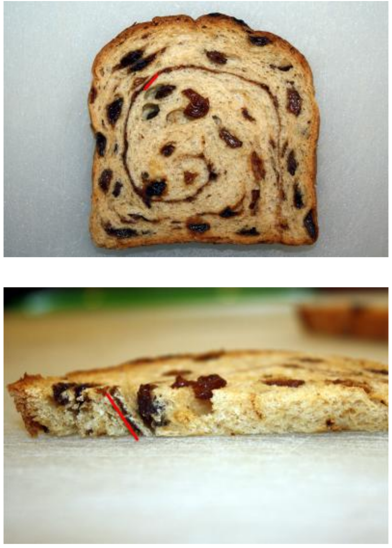


Figure 1: The bread stimulus. Top: Participants indicated if the region highlighted by the red line, was present on the surface or extend into the solid. Bottom: Red line shows the angle at which the swirl extends into the bread. Note participants never saw this view.

Method

Participants. Participants were 30 Temple University Undergraduates.

Stimuli. The stimuli consisted of 17 color photographs of common objects such as food, wood and rocks. There was one practice image and 16 experimental images. For each picture we selected a specific region of the picture to ask participants about. As shown in Figure 1, this region was indicated with a red line. In each image the area represented a region where a plane might intersect the visible surface. For example in Figure 1, the plane defined by the cinnamon layer between two regions of dough is indicated. Images when presented on the screen were approximately 25 x 18 cm.

These categories of images were chosen with two constraints: 1) that we were physically able to slice each object and measure the angle at which each highlighted region extended into the object and 2) that we sample a range of objects that might be familiar to participants.

Our stimuli fell into six broad categories defined by their internal structure: 1) rocks (granite slab), 2) wood (tree ring and a knot), 3) fruits (pineapple, papaya, kiwi), 4) vegetable (onion), 5) animals (fish and beef) and 6) food products that were originally liquid and are now solid (blue cheese, chocolate with almonds and cinnamon bread). These categories were selected because the internal structure ranged from highly structured and constrained by the environment (e.g., wood grain) to relatively unconstrained (e.g., minerals in rock) and thus the orientation is either knowable within a certain range or completely unknowable.

For example, the internal structure of wood is constrained by the environment. As tree structures are generally concentric cylinders, the extension into a slice is a function of the angle of the cut relative to the cylinders. For the fruit stimuli we selected fruits with radial symmetry and thus the internal structure is also structured. An onion although somewhat irregular in shape, has an organized internal structure. For the fish and beef stimuli we asked about how regions of fat extend in. Fat deposits are structured in complex ways by the surrounding muscles and thus organized but not to the same degree as wood, or fruit. The internal composition of rocks can be structured, but the orientation of a mineral's surface relative to the cutting plane is essentially arbitrary.

Apparatus. Stimuli were presented on a 20-inch Dell monitor. As shown in Figure 2, the monitor was positioned parallel to the ground.

Figure 2: The display used in the experiment. Participants used the black rod to indicate the orientation at which the highlighted region extended into the object.

Procedure. Participants were tested individually. They viewed each picture while standing with their nose over the center of the monitor. Participants were told that we were interested in their opinions of how regions of images continue in 3-dimensions. They were told that sometimes they would see pictures where they might have a strong sense that a region continued and sometimes they might have a sense that something was present only on the surface. To illustrate these cases, students were shown a picture of a Swiss roll and crayon marks on paper. All students reported seeing the layers of the Swiss roll as extending into the object while the crayon marks were present on the surface.

Participants were shown 16 pictures. For each, their task was to indicate whether the region indicated with the red line was present only on the surface or extended into the object. If they thought it extended into the object, they used a stainless steel rod with an inclinometer (angle measure) attached to indicate the orientation of continuation. Participants placed the edge of the rod on the red line and then moved the rod up and down to indicate the angle. The 0° was defined relative to the ground plane (i.e. if positioned the rod to indicate straight down, as shown in Figure 2, the angle measurement was 90°). After they estimated the angle, they reported their confidence in their response on a 5-point scale. Prior to viewing the 16 pictures participants practiced using the angle measurement device on the image of the Swiss roll. To be sure that there were not differences in the estimates based on the orientation of the picture, after viewing all the pictures and making their responses participants were shown the 16 pictures again but this time the images were rotated 180 degrees. This allowed us to calculate any bias due to their body position relative to the image. Finally, participants were shown the pictures a third time and asked to identify each picture. For any response with a confidence rating of 0 or 1, we further probed their uncertainty. Participants were asked to select which of the following reasons best described why they were uncertain in their response: 1) they have no idea what it could be, 2) the answer is unknowable 3) there could have been a range of possible angles. Additionally, for the pictures that they selected surface for, we told them that the region did extend in and asked them if they could make a guess about its orientation using the rod (note these estimates were not included in any analyses). After this, participants completed the Geologic Block Slicing Test (a measure of inferring internal spatial structures from views of multiple sides; Ormand et al, 2011), Spatial Orientation Test (Kozhevnikov & Hegarty, 2001) and the Mental Rotations Test (Shepard & Metzler, 1971). Data for these spatial tests are not presented in this paper.

Results

Although the single 2D view is insufficient to define a 3D shape, participants reported that the answer was unknowable on only 1% of the trials (12 times out of 960 trials), suggesting that participants do not recognize the need for multiple views to solve the intersection of constraints problem. Consistent with this, participants were confident in their angle estimates. The mean confidence was 3.2 (SD = 1.1) on a 5-point scale. Participants did perceive some of the highlighted regions as being only on the surface, but this was the case on only 26% of the trials, suggesting that participants tended to perceive the highlighted regions as extending into the object in three dimensions.

In order to calculate the participant's unbiased estimate for each picture, the two estimates were combined by calculating the average of the first estimate and 180° minus the second estimate. By presenting each picture in two orientations, we could remove any bias that a participant had to orient the rod towards (or away from) their body. For example, consider a case where the estimate for the first view and the second view (when the picture was rotated 180°) of a picture was 80°. If the participant were truly responding with an unbiased estimate, this would mean that if the estimate for the first view was 80°, the estimate for the second view should have been 100°. Thus, by subtracting the second estimate from 180, we can avoid any systematic bias (overall participants exhibited an ~4 degree bias towards their body).

Figure 3 shows the mean angle estimate without bias for each picture along with the 95% confidence interval for that picture. Inspection of the figure reveals that mean estimates tended to be biased towards 90°. Fifteen of the 16 pictures have mean estimates that are not significantly different from 90° (the red line denotes 90°). The only picture that has a mean estimate significantly greater than 90° was the "onion" picture.

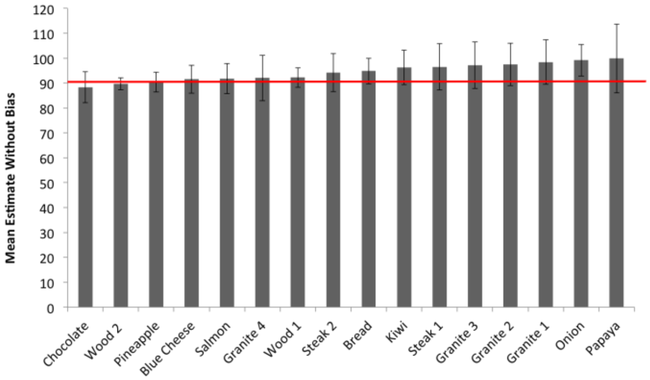


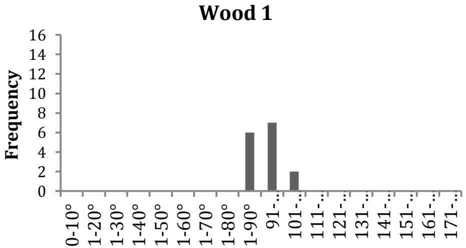
Figure 3: Mean estimate for each picture. Error bars show 95% confidence interval for each mean.

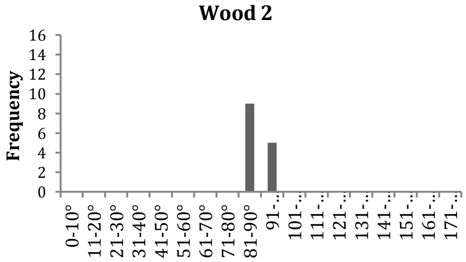
Next we examine whether the distribution of responses for each picture was random or whether there was a preferred direction for the distribution. Surface responses were excluded from this analysis. Thus the N for each picture varies based on the number of estimates for that picture. To examine the structure of the distributions we conducted the *Rayleigh test* (Zar, 1999) using the EZ Rose program (Baas, 2000). The null hypothesis is that the distribution of responses are randomly distributed around a semicircle. If the null is rejected then the distribution of responses has a preferred direction. To examine whether the null is accepted or rejected one compares R (mean vector length) for each picture to the critical value of the test statistic $R_{0.05}$ (see Baas, 2000 equation 10). If R is greater than $R_{0.05}$ then the distribution of responses has a preferred direction (i.e. is not random). As can be seen in Table 1 the Rayleigh test was rejected for 14 of the 16 pictures. For the "Tree Knot" and Papaya pictures, participants tended to perceive the region as on the surface (44% and 58% were judged to be "on surface," respectively). Thus the number of estimates was less than the 15 recommend for this test (Baas, 2000). However, an inspection of the frequency distributions (see below) reveals that when estimates were made, they centered around 90°.

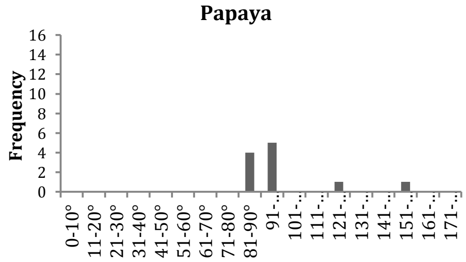
Table 1: Results of the Rayleigh Test.

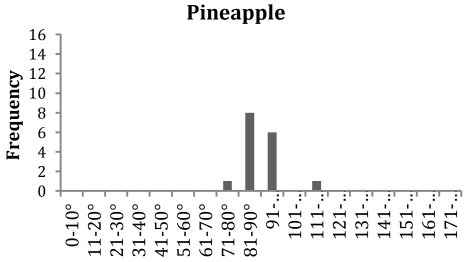
Picture	R	$R_{0.05}$	H_o
Chocolate	0.91	0.38	rejected
Tree Knot	-	-	-
Pineapple	0.97	0.43	rejected
Blue Cheese	0.94	0.43	rejected
Salmon	0.87	0.32	rejected
Granite 4	0.84	0.38	rejected
Wood 1	0.97	0.45	rejected
Steak 2	0.82	0.33	rejected
Bread	0.92	0.33	rejected
Kiwi	0.85	0.33	rejected
Steak 1	0.80	0.36	rejected
Granite 3	0.79	0.37	rejected
Granite 2	0.85	0.39	rejected
Granite 1	0.85	0.43	rejected
Onion	0.89	0.37	rejected
Papaya	-	-	-

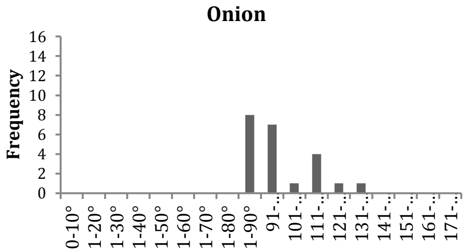
Finally, we examined whether the mean estimates centered around 90° because some participants estimated the angle at 10° and others at 170° and this averaged out to 90° or whether there was consistency among estimates for all participants. Figure 4 shows the frequency distribution for response for each of the 16 pictures grouped by category. As can be seen in the figure, the distributions are fairly uniform. There are certain pictures, for example the papaya, where participants on averaged perceived the indicated region to be on the surface (thus number of estimates is smaller). However, there were other pictures, like the kiwi and bread where participants agreed the region extended in. Also evident in the distributions is the limited spread. Mean estimates did not encompass the entire 0-180° spectrum; instead on average they were concentrated around 90.

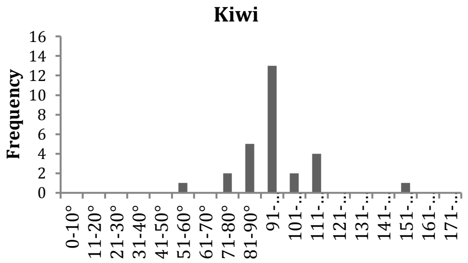


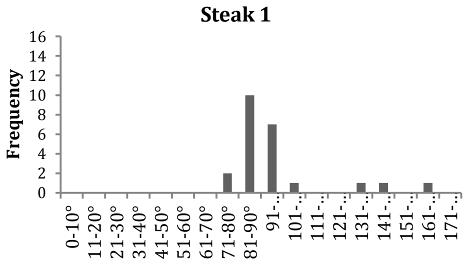


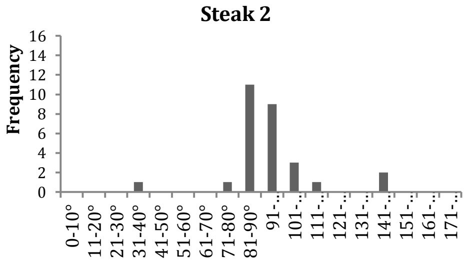


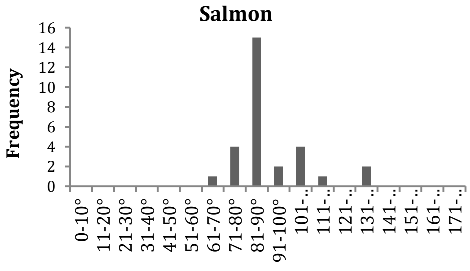












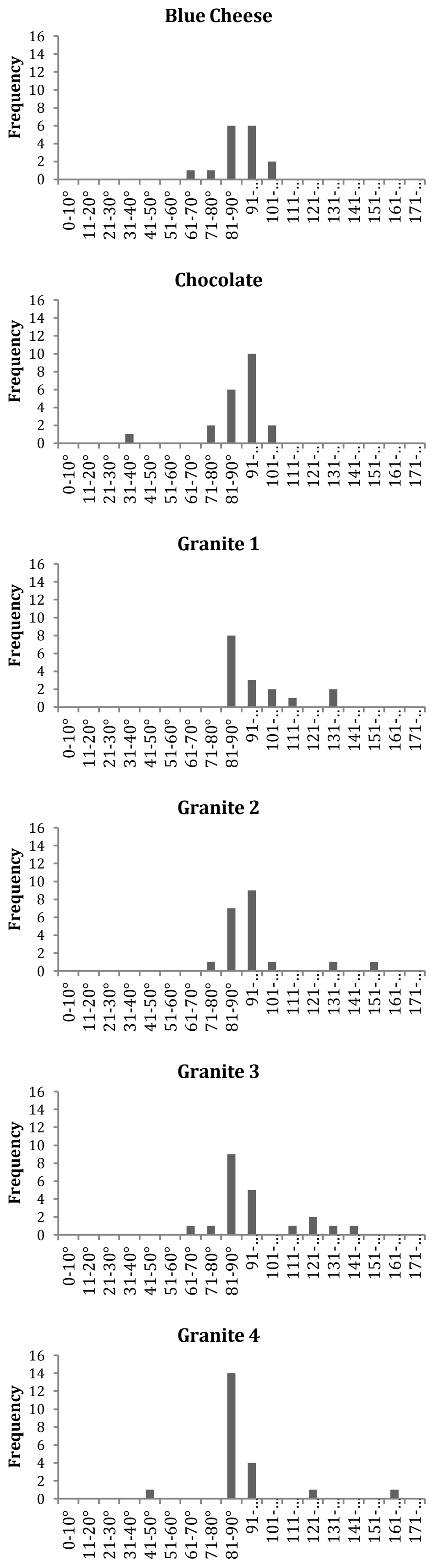


Figure 4: The frequency of surface response and estimates that fall within a 10° bin from 0-180° for each picture.

Discussion

Two conclusions can be drawn from this work. The first is that participants failed to recognize that one cannot know the orientation of a 3D structure from a single 2D view. Instead of recognizing that this situation is ambiguous, participants tended to have consistent intuitions about how regions of pictures extend in 3D into the object. The second conclusion that can be drawn from these findings is that participant's estimates tended to be clustered around 90°, suggesting that estimates were biased to assume that the surface continues straight back into the object. These findings 1) suggest possible constraints on models of amodal completion processes and 2) have implications for STEM education, in particular, challenges that arise when students do not recognize ambiguity. We consider each of these in turn.

In developing our understanding of how the visual system estimates the 3D form of objects, there are important completion phenomena that should constrain computational models. Here we describe a completion phenomenon that to our knowledge has not been recognized -- participants readily perceived 3D forms from 2D views (of both familiar and unfamiliar objects) and the perception is that surface boundaries extend perpendicularly into a solid.

This bias may be informative about the filling *out* process that occur under conditions where the 3D completion of surfaces can not occur because the edges of an object are not aligned in the 2D projection on the retina (see Tse, 1999). The bias evident in this study may be a product of the way the visual system handles a more common instance of having partial information about an object – when viewed head-on. Normally, if one can only see a single side of an object, it is a result of your current viewpoint (the line-of-sight is perpendicular to the front face). So, the visual system may represent the portions of the occluded object using past experience (i.e. knowledge about cinnamon swirl bread) or some properties of the front face. This process becomes evident when the object (the cinnamon swirl) is surrounded by an opaque region (more bread) and thus other sides are not visible to the observer. The visual system is not flummoxed by this situation but instead rapidly extrapolates from the available single surface to represent an extended 3D structure. Under these conditions, the completion process reflects the assumption that edges on the surface project straight into the object.

Our observations suggest the existence of a visual process that uses available visual information to extend form representations into regions where the form is not visible. Models of visual completion argue that completion processes reflect the system's attempt to construct a representation of the most likely 3D form. An ongoing debate in the literature has examined whether completion processes occur as a result of extrapolation (filling out) or interpolation (filling in). Here, where only one face of a 3D form is visible, the completion processes must be based on extrapolation - filling out from available visual information (Shipley & Kellman, 2003) – rather than interpolation between defined regions. Extrapolation occurs in both amodal (Kanizsa, 1979) and modal displays (Shipley & Kellman, 2003), but previous demonstrations have been extrapolation of planes or edges, not 3D volumes.

In addition to informing our understanding of visual completion processes, these results have important implications for STEM education. Consider a field geology student examining a rock face and trying to make inferences about its 3D structure, or an anatomy student learning where best to make an incision during dissection or surgery. Both of these tasks require inferences about a 3D form from a 2D view. What is critical in both these cases is that an accurate estimate of the orientation of the 3D form requires more information – either in the form of looking at another angle of the rock to see how features penetrate in, or knowing something about the true 3D shape and using that to constrain the estimate of orientation of how the region extends into the volume. There are aspects of the world that might place constraints on the probable internal structure (i.e. grains in wood has a cylindrical structure), however students must recognize the need to seek out additional information in order to make inferences about the 3D structure.

An extensive body of research has examined decision-making regarding uncertain events (Kahnamen, & Tversky 1982), however to our knowledge work in this area has not examined uncertain perceptual situations and their relationships to confusion in the classroom. We believe this is an interesting area to pursue in future research. How best to convey to students that information in an image may appear to determinant, but is in fact ambiguous.

Our aim in this paper is threefold: first, examine 3D completion from 2D views and to bring this type of process to the attention of the research community. By making researchers aware of completion processes in the “wild”, we hope to begin a dialogue that may move completion research forward into new domains. Second, to expand the phenomena considered by any model of visual completion. Third, we wish to illustrate the importance of scaffolding student’s ability to recognize ambiguity and the necessity to seek out additional information for solving a problem.

References

Baas, J. (2000) EZ-ROSE a computer program for equal area circular histograms and statistical analysis of two dimensional vectorial data. Comput. Geosci., 26, 153–166.

Chariker, J. H., Naaz, F., & Pani, J. R. (2011). Computer-based learning of neuroanatomy: A longitudinal study of learning, transfer, and retention. Journal of educational psychology, 103(1), 19.

Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In G. Allen (Ed.) Applied Spatial Cognition. Mahwah, NJ: Lawrence Erlbaum Associates.

Kanizsa, G. (1979). Organization in Vision. New York: Praeger.

Kali, Y. & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. Journal of Research in Science Teaching, 33, 369–391.

Kozhevnikov, M. & Hegarty, M. (2001). A dissociation between object-manipulation and perspective-taking spatial abilities. Memory & Cognition, 29, 745-756.

Ormand, C., Shipley, T. F, Manduca, C, and Tikoff, B (2011). Improving Geoscience Students' Spatial Thinking Skills: Applying Cognitive Science Research in the Classroom. American Geophysical Union Fall meeting (San Francisco, CA).

Shipley, T. F., & Kellman, P. J. (2003). Boundary completion in illusory contours: Interpolation or extrapolation? Perception, 32, 985-999.

Shepard, R and Metzler. J. "Mental rotation of three dimensional objects." Science 1971. 171(972):701-3.

Tse, P. U. (1999). Volume completion. Cognitive Psychology, 39, 37-68.

Zar, J. H. (1998). Biostatistical analysis (4th ed.). Upper Saddle River, NJ: Prentice Hall.