CrossMark

BRIEF REPORT

Feature-based attention resolves depth ambiguity

D. Yu [1] · B. Levinthal [1] · S. L. Franconeri [1]

Abstract Perceiving the world around us requires that we resolve ambiguity. This process is often studied in the lab using ambiguous figures whose structures can be interpreted in multiple ways. One class of figures contains ambiguity in its depth relations, such that either of two surfaces could be seen as being the "front" of an object. Previous research suggests that selectively attending to a given *location* on such objects can bias the perception of that region as the front. This study asks whether selectively attending to a distributed *feature* can also bias that region toward the front. Participants viewed a structure-from-motion display of a rotating cylinder that could be perceived as rotating clockwise or counterclockwise (as imagined viewing from the top), depending on whether a set of red or green moving dots were seen as being in the front. A secondary task encouraged observers to globally attend to either red or green. Results from both Experiment 1 and 2 showed that the dots on the cylinder that shared the attended feature, and its corresponding surface, were more likely to be seen as being in the front, as measured by participants' clockwise versus counterclockwise percept reports. Feature-based attention, like location-based attention, is capable of biasing competition among potential interpretations of figures with ambiguous structure in depth.

Keywords Visual selective attention · Visual perception

✉ D. Yu
dianyu2017@u.northwestern.edu

[1] Department of Psychology, Northwestern University, 2029 Sheridan Rd, Evanston, IL 60208, USA

Sensory input is fit to internal models, refined over evolution and experience, that represent best guesses about what the world is like. When inputs can be fit to multiple potential models, a competitive process within our perceptual system declares a single winner at a time (Beck & Kastner, 2009; Desimone & Duncan, 1995). This process is often studied in the lab using ambiguous figures that can be interpreted in multiple ways, like the identity of the ambiguous duck/rabbit (Attneave, 1971). The competitive process can be biased in several ways to make a given percept more likely, such as by first showing a viewer an unambiguous rabbit (Goolkasian, 1991).

Some stimuli are ambiguous in how their surfaces are ordered in depth, such as the Necker Cube (see Figure 1a). Here, the perceptual system can also bias competition among possible depth orderings, where one of multiple surfaces could be seen as being the "front" of an object. This competitive process of surface depth assignment can be biased to make a given percept more likely, such as by first showing an unambiguous version of a cube (Long, Toppino, & Mondin, 1992). But another way to bias the viewer's interpretation of the cube is to fixate on or selectively attend to the location of either the lower-left or upper-right vertex of the cube (Kawabata, 1986). Similarly, even when fixating the center of an ambiguous Necker Cube, reports of seeing one surface as being in the front are strongly associated with shifts of covert visual attention to the location of that surface, as measured by an electrophysiological correlate (Xu & Franconeri, 2012).

This location-based biasing strategy may not be the only way for attentional selection to bias completion for depth ordering. In the 3-D molecular diagram in Figure 1b, occlusion cues allow an unambiguous interpretation of depth relations. But in the dash-wedge notation diagram next to it, a chemistry student must see the atomic bindings marked by black wedges as being in front (closer) of the carbon atoms, and the bindings marked by dashed wedges as being behind (farther away) than
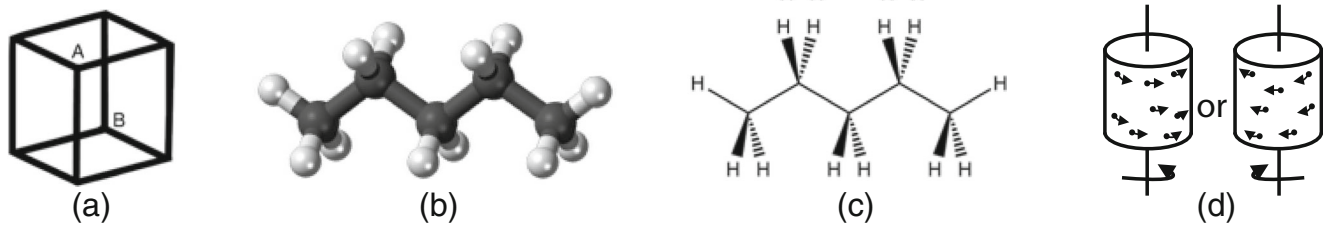
Fig. 1 a The classic Necker Cube ambiguous figure, showing multiple depth interpretations: either corner A or B could be the front. Fixation or location-based attention to either corner can bias that surface to be seen as the front. **b** 3-D model of a pentane molecule, containing unambiguous information about surface depths through occlusion cues. *Note.* From Pentane 3D ball, In Wikimedia Commons. Retrieved February 2, 2016, from https://commons.wikimedia.org/wiki/File:Pentane_3D_ball.png. Copyright 2014 by Creative Commons CC0 1.0 Universal Public Domain Dedication. Reprinted with permission. **c** Dash-wedge notation example of the same molecule, where a student must imagine the black wedges as being atoms in a closer depth plane, and the dashed atoms being in a more distant depth plane. Although this percept might be biased by multifocal location-based attention to five locations, the less restricted capacity limits of feature-based attention render it a more likely mechanism. **d** Two percepts of the experiment's rotating cylinder stimulus

the carbon atoms. How does her perceptual system construct such 3-D mental images, marking the set of black wedges as being in the front, even when she has never encountered a 3-D model of this particular object? A location-based biasing strategy would require her to noncontiguously select five objects within a spatially dense region, which should be extremely difficult (Scimeca & Franconeri, 2015), especially for even more complex molecules.

Instead, we argue that feature-based selection—in this case to the five black wedges—could contribute to her ability to see those parts as being in the front. Features such as color, orientation, or motion directions can be broadly simultaneously selected (Liu, Larsson, & Carrasco, 2007; Saenz, Buracas, & Boynton, 2002; but see Leonard, Balestreri, & Luck, 2015, for evidence of constraints), regardless of the number of objects selected. This should allow a viewer to bias a large set of object parts toward a "front" depth plane.

To test whether feature-based selection can bias depth assignments in the absence of location-based selection, we used a depth-ambiguous stimulus where location-based attention could not distinguish between surfaces, because they fully overlap. The stimulus was a rotating cylinder constructed of two arrays of moving dots: red and green, moving in opposite directions. When red moves rightwards and green leftwards, if the red dots are perceived as being in front, the cylinder is seen as rotating counterclockwise (as seen from the top; see Figure 1c for illustration), and clockwise when the green dots are seen in front. Asking the viewer to judge rotation direction—instead of relative surface depth percepts—minimized potential experimenter demand characteristics for reporting associations between the cued color and that color being at the front. To cue observers to select either the red or green dots, a secondary task was embedded in peripheral field of red and green dots. Trials unpredictably required observers to report on either the state of the peripheral field of the instructed color, or the perceived depth assignment within the ambiguous cylinder. To preview the results, feature-based attention biased depth

assignments—attending to red/green in the peripheral task biased viewers to see the cylinder rotating in the direction such that the surface in the attended color would appear in the front.

Experiment 1

Method

Participants

Eighteen participants (8 female; age range, 18–30 years) participated in Experiment 1. The number of participants was not based on a power analysis, but was specified a priori as a large but standard sample across past psychophysical studies. All participants had normal or corrected-to-normal vision, were paid for participation or granted course credit, and gave written consent.

Apparatus

Stimuli were generated using MATLAB with the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) driven by an Apple Mac Mini running OS 10.6 and were displayed on 17-inch monitors at a resolution of 1028 × 768 pixels and refresh rate of 75 Hz. Head movement was unrestrained, but the average viewing distance was approximately 47 cm; at this distance, the screen area subtended 39.8° × 30.3° of visual angle (24.7 px/degree).

Stimuli

The display consisted of a simulated parallel projection of a transparent rotating cylinder at screen center. The cylinder subtended 4° vertically and horizontally, and was constructed of randomly distributed red and green dots (0.12° × 0.12°; 15 dots per square visual degree) moving horizontally against a black background. The speed of each dot followed a sine wave

function, creating the appearance of a rotating cylinder with half of its surface being red and half being green, or occasionally, two concave or two convex half-pipe surfaces. The angular velocity of the dots within the cylinder was 45.8°/s. All red dots moved from left to right and all green dots right to left. Red was set at a RGB value of [180, 0, 0]. Green was initially set as [0, 120, 0], but was adjusted for some participants (see the Procedure section).

The cylinder was surrounded by two arrays of red and green dots moving in opposite directions. These dots shared the same center as the cylinder and covered an area of 12.0° × 12.0°, except for the space occupied by the cylinder. There was a 0.4° gap between the edge of the cylinder and the inner edge of the surrounding arrays. Dots on the surrounding arrays were of the same size and color as those that belonged to the cylinder, but of lower density (one dot per squared visual degree). The two arrays of dots surrounding the cylinder traveled in opposite directions (up–down) at 1.6°/s. In a trial where either red or green dots were flashing, the RGB values of the flashing set alternated between 60 % and 140 % of the original RGB values at a frequency of 25 Hz.

Procedure

Each participant initially practiced perceiving the cylinder as rotating clockwise versus counterclockwise (the surrounding dots were not present in these displays). If a participant reported that it was consistently easier to see the green surface as being in the front (or back), the green value was increased (or decreased) until it was equally easy to see both perceptions based on participants' report. Once the color adjustment was completed, participants entered a training phase to ensure that they were able to see and confidently report both percepts. They were asked to alternate their percept of the cylinder (clockwise vs. counterclockwise) 10 times. If a participant failed to achieve 10 flips within 4 minutes, the experiment ended.

The testing session consisted of eight blocks plus an initial practice block, with 16 trials per block. Each trial started with a yellow fixation dot center screen. Participants were instructed to fixate on this dot throughout the trial, and press a key to start each trial. A rotating cylinder and surrounding dot array as described above appeared upon the key press. Either the red or green surrounding dot array would be flickering. Participants were required to attend to the flickering set of dots and monitor for an increase in its speed. In an unpredictable half of the trials, the attended set of dots doubled its speed within 3 to 6 seconds and persisted with increased speed for 2 seconds. Participants were instructed to press the space bar as soon as the speed change occurred. In the other half of trials, the fixation dot would turn blue within 4 to 6 seconds of the trial's start, signaling participants to report their current perception of the cylinder. Keys were labeled with icons

showing either clockwise or counterclockwise rotation of a cylinder, and a third key was reserved for when participants did not see a clear percept of a rotating cylinder (e.g., two convex half-pipes). The process of a typical trial is illustrated in Figure 2.

Participants monitored for speed changes in one set of peripheral dots by selectively attending to either red or green, and potentially the upward or downward motion that was associated with each color. Note that even if our participants chose to attend primarily to motion directions, the associated color should become selected as well (Boynton, Ciaramitaro, & Arman, 2006; Lustig & Beck, 2012).

Results

Four of the 18 participants could not see 10 alternations of the direction of the cylinder in the training phase and were excluded from the remainder of the experiment. Among the remaining 14 participants, six used an increased green luminance [0, 150, 0] and one used a decreased [0, 110, 0]. For the speed-increase detection task, we coded trials with response times longer than 2,000 ms as misses. Participants had an average hit rate of 90.4 % (SE = 2.2 %). Among the hit trials, average response time was 860 ms (SE = 29 ms). "Unclear" percepts of the cylinder were rare (M = 2.2 %, SE = 1.1 %), for both attend-red (M = 2.9 %, SE = 1.6 %) and attend-green (M = 1.6 %, SE = 0.9 %) trials.

Feature-based attention biased depth assignment. The surface sharing the attended feature was more likely to be seen as the front of the rotating cylinder, compared to 50 % chance levels (M = +7 %, SE = 2.6 %), t(13) = 2.70, p = 0.02, Cohen's d = 0.72. Figure 3 depicts bias rates for individual participants, and shows several subjects could bias their percepts with high rates, some over 70 %. This effect tended to be stronger for
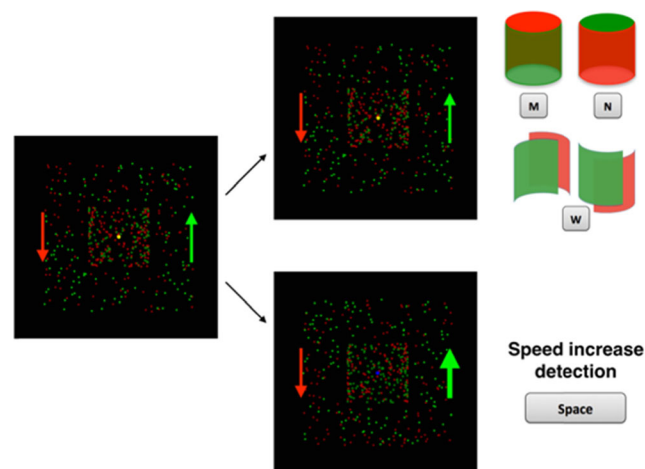

Fig. 2 Participants attended to a given color in the peripheral task, which should bias that color's surface to the front of the cylinder. Half the trials required a response to the secondary task (below), and half a judgment of the cylinder's shape
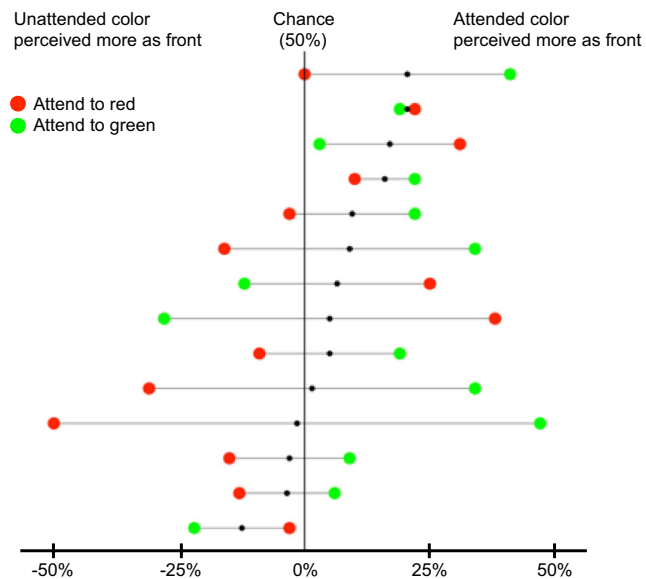
Fig. 3 Results from Experiment 1, showing that feature-based attention can bias people to see the attended color as being in the front

attend-green ($M = +13.9$ %, $SE = 6.1$ %) than for attend-red ($M = -1$ %, $SE = 6.6$ %), though this trend did not approach significance, $t(13) = 1.29$, $p = .22$, Cohen's $d = .33$.

Experiment 2

One might argue that the flickering dots in Experiment 1 led attention to be drawn to that color exogenously, instead of endogenously and voluntarily. Experiment 2 therefore cued the to-be-attended group of dots symbolically, with an up or down arrow before each trial. We also added a balanced design for the combinations of colors and motion directions within the cylinder, adding trials where the green dots moved leftwards and red dots moved rightwards. We also more formally measured baseline preferences for seeing each color as being in the front, for a clearer view of the additional effect of feature-based attention on each color compared to this baseline.

Method

Participants

Twenty-four participants (17 females; age range, 18–24 years) participated in Experiment 2.

The number of participants was increased because we doubled the number of conditions, and each condition received fewer trials.

Apparatus

The apparatus was the same as in Experiment 1.

Stimuli

Stimuli were identical to those in Experiment 1, except for (1) in half of the trials, the cylinder consisted of red dots moving rightwards and green dots moving leftwards, and vice versa in the other half of the trials, and (2) dots in the surrounding array no longer flickered.

Procedure

Each participant first completed a color-adjustment session using an ambiguous cylinder composed of red and green moving dots. The RGB values for color red were [180, 0, 0], except for a single participant who required a lower red value of [160, 0, 0] to achieve roughly equal salience with green. In each trial, initial RGB values for color green were either [0, 255, 0] or [0, 80, 0]. Participants first indicated their initial perception of how the cylinder rotated (all participants reported seeing the brighter color as front initially in every trial). Then they would adjust the brightness of the green color until the rotation direction of the cylinder flipped, and indicate the "flipping point" by pressing the space bar. There were 16 trials in total, half in which green dots moved rightwards and red dots leftwards, and the other half vice versa. The RGB values of the green color at each flipping point were recorded and averaged separately for the two color–motion combinations and were used in the remainder of the experiment.

We then measured the baseline percept of the cylinder's rotation direction using these roughly balanced colors. Stimuli in this session were identical to the experimental displays in Experiment 1. The fixation dot would turn blue within 4 to 6 seconds of the trial's start. Participants then indicated whether they saw the cylinder rotate clockwise, counterclockwise, or did not see a rotating cylinder, at the moment when the fixation dot turned blue. There were 32 trials in this session, with 16 trials for each color–motion–direction combination.

Participants then began the test session, which was identical to Experiment 1 except for three changes. First, to-be-attended dots were cued by an upward- or downward-pointing arrow at the beginning of each trial. Second, to prevent participants from just attending to motion speed-ups in general, as opposed to more selectively to motion speed-ups in the attended collection, the unattended collection also sped up within the same time window, with independent timing. Third, participants received warnings and a 7-second time-out if they responded to the speed increase of the unattended group of dots. This test session consisted of seven blocks, each with 32 trials.

Results

One participant was excluded due to extremely low accuracy in the speed detection task (53 %), and another for a reported

inability to hold the percept of a 3-D cylinder. Baseline percepts of cylinder with individually adjusted colors did not differ between before and after the testing session and were combined in the analysis. Participants reported seeing a clear perception of the cylinder rotating for the majority of the time ($M = 96.4$ %, $SE = 1.3$ %), and did not show a preference for seeing a certain color as being in the front (Green as front: $M = 53.4$ %, $SE = 3.4$ %).

For the speed-increase detection task in the testing session, we coded trials with response times longer than 1,500 ms as misses. Participants had an average hit rate of 90.5 % ($SE = 1.1$ %) and average miss rate of 6.2 % ($SE = 0.8$ %). Among the hit trials, the average response time was 650 ms ($SE = 27$ ms). The surface sharing the attended feature was more likely to be seen as the front ($M = +8.2$ %, $SE = 2.2$ %), $t(21) = 3.67$, $p = .001$, Cohen's $d = .78$, and was not influenced by cylinder color–motion combination (Green-left: $M = +7.7$ %, $SE = 2.7$ %; Green-right: $M = +8.6$ %, $SE = 2.0$ %). Comparing this attention effect to each participant's baseline for red versus green, we again see a numerically stronger effect for green ($M = +12.2$ %, $SE = 3.3$ %) than for red ($M = +4.1$ %, $SE = 4.0$ %), although not statistically significant, $t(21) = 1.38$, $p = .18$, Cohen's $d = .30$. Figure 4 depicts bias rates for individual participants. "Unclear" percepts of the cylinder were rare ($M = 1.2$ %, $SE = 0.8$ %), for both attend-red ($M = 1.0$ %, $SE = 0.7$ %) and attend-green ($M = 1.5$ %, $SE = 1.0$ %) trials.

General Discussion

Feature-based selection can bias competition within interpretations of ambiguous depth, such that the attended feature is
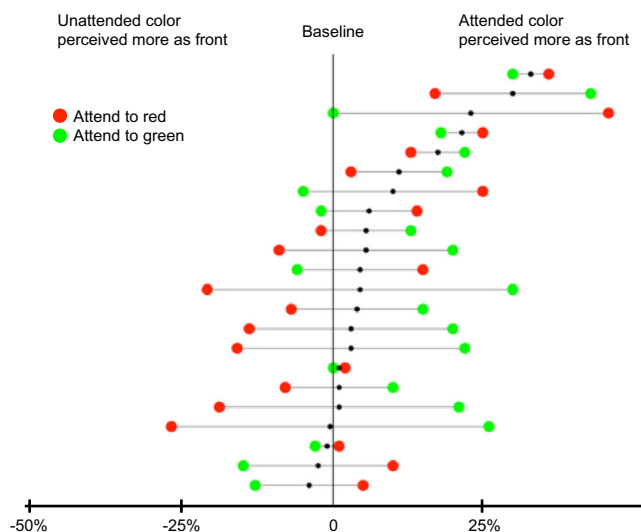

Fig. 4 Results from Experiment 2, showing how much attending to red/green increases chances of seeing the attended color as the front surface for each participant, compared to their default percentage in seeing such surface as the front in the baseline condition

more likely to be perceived as the closer surface. We manipulated attention to features (Saenz et al., 2002) with a peripheral secondary task and found that the surface of an ambiguous figure with the corresponding color was more likely to be seen as the front of that figure.

These results are consistent with informal reports from at least two past studies. In one, participants tracked the identity of a Gabor patch that directly overlapped another, each continuously changing in orientation, size, and spatial frequency. Although peripheral to the purpose of the study, participants reported that the tracked Gabor—which should benefit from feature-based attention along one or more dimensions—was more likely to be seen as the "front" object (Blaser, Pylyshyn, & Holcombe, 2000). Another study constructed a dot sphere that rotated in an ambiguous direction, similar to our ambiguously rotating cylinder. A patch of increased dot density (or no dots) (which is likely to attract attention; Nothdurft, 1993) was placed randomly on the surface of the sphere and rotated at the same speed as the rest of the dots constituting the sphere. Participants were more likely to perceive the sphere rotating in a direction consistent with this salient patch being on the sphere's front hemisphere (Brouwer, & van Ee, 2006). The current finding is also related to the resolution of another type of depth ambiguity—figure-ground assignment, where spatial attention also appears to bias competition (Wagemans et al., 2012). When observers view an image segregated by a jagged vertical line, either the left or right side can be seen as the closer figural surface, and selective attention to (or fixating on) the location of one surface can bias competition toward that side being perceived as the figure (Peterson & Gibson, 1994; Vecera, Flevaris, & Filapek, 2004).

Why does attending to a surface—either by location or by feature—bias it to the front? One possibility is that there are existing associations between attending to a surface and that surface being in the front, because we are more likely to attend to an object's front surfaces, as opposed to the back surfaces, which are typically occluded (Xu & Franconeri, 2012; see Goldreich & Peterson, 2012, for a similar argument for figure-ground assignment). Another possibility is that the perceptual consequences of attention, such as increased resolution or contrast (Carrasco, Ling, & Read, 2004; Ling, Liu, & Carrasco, 2009; Treue & Trujillo, 1999), are also commonly associated with properties of closer surfaces (Egusa, 1983; O'Shea, Blackburn, & Ono, 1994).

Feature-based attention, like location-based attention, will not be the only, or even the dominant, factor in determining surface depth assignment. Other factors that will carry even more power, including occlusion cues (Grossberg, 1997); geometric characteristics such as convexity and symmetry (Peterson & Gibson, 1994; Peterson & Salvagio, 2008); image blur and contrast (Mather, 1996; O'Shea, et al., 1994); and prior probabilities for different interpretations (Dobbins & Grossmann, 2010; Huang & Pashler, 2009).

However, attention may be a useful cue for depth assignment when other cues are lacking, as in the dash-wedge molecular diagram in Figure 1. If feature-based attention does support the percept of a 3-D structure for such examples, then it carries predictions for how to design these diagrams to be easier to perceive. For example, the "closer" atoms in the dash-wedge diagram are cued by their color, but the orientation of those wedges varies, with some tilting left versus right. This featural difference might make it more difficult to select all of the wedges at once because the orientation dimension might be involuntarily selected (Lustig & Beck, 2012). Although this characteristic may be unavoidable or even desirable for other reasons in this particular example, it could be part of a design decision in other examples. As a second example, the greater perceptual salience of black wedges, perhaps driven from being the strongest instance of that particular color within the diagram, seems to make them easier to be "pulled" to the front. In contrast, the dashed wedges seem to be more difficult to be seen as being in front. If feature-based attention plays a role in such diagrams, then designers must use features that are easily selected and consider the relative salience and selection power of each feature value, which will determine the surfaces that are brought to the front easier, or as a default percept.

References

Attneave, F. (1971). Multistability in perception. *Scientific American, 225,* 63–71.

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49*(10), 1154–1165.

Boynton, G. M., Ciaramitaro, V. M., & Arman, A. C. (2006). Effects of feature-based attention on the motion aftereffect at remote locations. *Vision Research, 46*(18), 2968–2976.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brouwer, G. J., & van Ee, R. (2006). Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Research, 46*(20), 3393–3402.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*(3), 308–313.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Dobbins, A. C., & Grossmann, J. K. (2010). Asymmetries in perception of 3D orientation. *PLOS ONE, 5*(3), e9553.

Egusa, H. (1983). Effects of brightness, hue, and saturation on perceived depth between adjacent regions in the visual field. *Perception, 12*(2), 167–175.

Goldreich, D., & Peterson, M. A. (2012). A Bayesian observer replicates convexity context effects in figure-ground perception. *Seeing and Perceiving, 25,* 365–395.

Goolkasian, P. (1991). The effect of size on the perception of ambiguous figures. *Bulletin of the Psychonomic Society, 29*(2), 161–164.

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure–ground perception of two-dimensional pictures. *Psychological Review, 104*(3), 618.

Huang, L., & Pashler, H. (2009). Reversing the attention effect in figure-ground perception. *Psychological Science, 20*(10), 1199–1201.

Kawabata, N. (1986). Attention and depth perception. *Perception, 15*(5), 563–572.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox–3. *Perception, 36*(14), 1.

Leonard, C. J., Balestreri, A., & Luck, S. J. (2015). Interactions between space-based and feature-based attention. *Journal of Experimental Psychology: Human Perception and Performance, 41*(1), 11.

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research, 49*(10), 1194–1204.

Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron, 55*(2), 313–323.

Long, G. M., Toppino, T. C., & Mondin, G. W. (1992). Prime time: Fatigue and set effects in the perception of reversible figures. *Perception & Psychophysics, 52*(6), 609–616.

Lustig, A. G., & Beck, D. M. (2012). Task-relevant and task-irrelevant dimensions are modulated independently at a task-irrelevant location. *Journal of Cognitive Neuroscience, 24*(9), 1884–1895.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London B: Biological Sciences, 263*(1367), 169–172.

Nothdurft, H. C. (1993). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research, 33*(14), 1937–1958.

O'Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision Research, 34*(12), 1595–1604.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Peterson, M. A., & Gibson, B. S. (1994). Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours. *Perception & Psychophysics, 56*(5), 551–564.

Peterson, M. A., & Salvagio, E. (2008). Inhibitory competition in figure-ground perception: Context and convexity. *Journal of Vision, 8*(16), 4.1–4.13.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience, 5*(7), 631–632.

Scimeca, J. M., & Franconeri, S. L. (2015). Selecting and tracking multiple objects. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 109–118.

Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature, 399*(6736), 575–579.

Vecera, S. P., Flevaris, A. V., & Filapek, J. C. (2004). Exogenous spatial attention influences figure-ground assignment. *Psychological Science, 15*(1), 20–26.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172.

Xu, L., & Franconeri, S. L. (2012). The head of the table: The location of the spotlight of attention may determine the 'front' of ambiguous objects. *Journal of Neuroscience, 32*(4), 1408–1412.