The viewing-from-above bias and the silhouette illusion

Nikolaus F. Troje and Matthew McAdam, Queen's University, Kingston, Ontario

Abstract: The silhouette illusion which was published online a number of years ago by the Japanese Flash designer Nobuyuki Kayahara has received substantial attention from the online community. One feature that seems to make it interesting is an apparent rotational bias: Observers see it spinning more often clockwise than counter-clockwise. Here, we show that this rotational bias is in fact due to the visual system's preference for viewpoints from above rather than from below.

Visual perception is riddled with ambiguities. In most cases, they are resolved by means of firm assumptions about the regularities of the visual world. Some of these are very well known and appear in every textbook on visual perception. For instance, most students of psychology are familiar with the demonstrations of bumps that change into troughs (and vice versa) when turned upside down – the switch in perceived relief being due to the visual system's very reasonable assumption that light normally comes from above rather than from below (Ramachandran 1988).

Other visual heuristics have not achieved the same degree of visibility even though their effects are at least as significant as the “light-from-above” heuristic. To this latter category belongs the “viewing-from-above” bias. It has been described in many different contexts but never explicitly enough to make it into textbooks and become established in the general awareness of students and scholars of vision science. This bias can affect the ambiguity of a Necker cube and other inherently depth ambiguous figures to a degree at which their perception becomes virtually unambiguous.

The ambiguity inherent in a Necker cube can be interpreted in terms of a decision between two possible viewpoints from which the cube is looked at. Observers tend to choose the viewpoint that looks at the object from above rather than from below. The origin of this insight remains obscure. Kornmeier and colleagues (2009) write that “It is well known that the cube-front-side bottom is the preferred initial percept of most observers” but none of the papers they cite in that context (Washburn et al 1929; Price 1967, 1969) provide any data in that respect. Data can be found, though, in more recent publications. Sundareswara and Schrater (2008) report that for long exposures of Necker cubes the view-from-above interpretation is seen for about 60% of the time. For very short presentations, the preference is almost 100%, a result that fits our own as shown in Fig. 1 (see also Troje 2010). Mamassian and Landy (1998) document a view-from-above bias for line drawings of surface shapes. Ooi and colleagues (2001) provide experimental evidence that relative height in the visual field is used as a depth cue which in turn implies the viewing-from-above assumption.

Here, we want to show that the viewing-from-above bias also explains “rotational” biases in the Silhouette Illusion, which has been created by Japanese artist Nobuyuki Kayahara and was published online on his website in 2003 (http://www.procreo.jp/lab/silhouette.swf). It shows the silhouette of the static posture of a female dancer pirouetting about a vertical axis. The dancer stands on one foot and stretches out the other. Like a Necker cube and other depth-ambiguous figures, this silhouette can be
interpreted as spinning either clockwise or counter-clockwise (Wallach & O'Connell 1953).

Since it was first published, Kayahara's silhouette illusion has travelled the internet quite extensively and it appears in the context of all sorts of weird theories and applications (for instance, as a “test” to distinguish “right-brainers” from “left-brainers”). One feature that may have sparked some of these ideas is the fact that even though both spinning directions can be seen, it is more often perceived to rotate clockwise than counter-clockwise. An online poll conducted by cognitiveDaily in 2008 (http://scienceblogs.com/cognitivedaily/2008/10/casual_fridays_tk421_why_cant.php) reports that about two thirds of the 1600 people who answered the poll see the dancer spinning clockwise and only one third counter-clockwise.

Looking at the animation carefully reveals that the direction from which it is rendered is not exactly perpendicular to the axis of rotation. Tracking the tip of the outstretched foot frame by frame results in an ellipse with an aspect ratio of 0.12, which translates into a camera viewpoint of either 6.8 deg above or below a horizontal plane. Assuming orthographic projection, the ambiguity between the spinning direction, on the one hand, and the sign of the camera elevation, on the other hand, can not be resolved. Therefore, the observed “clockwise bias” could as well be due to a viewing-from-above bias (Fig. 2).

To test this hypothesis, we re-created the silhouette animation using different camera elevations and measured the percentage of time in which the two possible percepts were seen, as well as the number of reversals that occurred. We also varied the speed with which the illusion is spinning in order to cover the wide range of spinning speeds at which the silhouette illusion appears on the internet.

Kayahara's original silhouette was based on a model provided by the commercially available rendering software Poser. We also used Poser (version 8, Smith Micro Inc.) and picked from the current version a model that was as similar as possible to the one used by Kayahara: The figure of a young woman adopting the same pose as Kayahara's dancer. The figure was rendered orthographically from four different viewpoints which only differed in camera elevation: +10 deg (i.e., 10 degrees from above), +3 deg, -3 deg, and -10 deg. In addition, we generated versions of the stimulus with three different spinning speeds: 25 deg/s, 100 deg/s, and 175 deg/s. Combining the two factors, our experiment contained a total of 12 displays. All renderings showed a black silhouette on a white background.

The 3D figure used to render the 2D stimuli was always rotated counter-clockwise. Note, however, that rendering a clockwise spinning figure with a camera looking from above would have created exactly the same sequence of images as our rendering with the camera from below looking at a counter-clockwise spinning figure. Likewise, a clockwise spinning figure shown from below is the same as a counter-clockwise spinning figure rendered from above.

Each of the 24 participants in the study was presented with each of the 12 displays once. Accounting for potential order effect we created 24 different orders based on a Latin square logic.

A single trial lasted 4 min. During the whole duration of the trial, subjects had to keep pressing one of two keys in order to indicate when the figure appeared to be spinning clockwise and when it was spinning counter-clockwise. When either no key was pressed or both keys were pressed, an unpleasant warning sound reminded the observer to keep exactly one key pressed.

Figure 3 shows data from one observer to illustrate typical behaviour and data analysis. If the camera is elevated 10 deg with respect to horizontal, this observer sees the figure spinning counter-clockwise for 60% of the time. That also means he sees the figure from above for about 60% of the time. When rendered with a camera looking from 10 deg below the horizontal, he sees the figure rotating clockwise
(and therefore from above) again for 60% of the time. The symmetry between these two cases indicates that this observer shows no rotational bias -- neither clockwise nor counter-clockwise. The viewing-from-above (VFA) bias is expressed in terms of the slope of the plotted function. In the case of this particular observer, this slope is 0.015 deg$^{-1}$ on average over the three spinning speeds. The number of reversals he experienced averaged to 3.4 reversals per minute.

On average over all 24 observers, the VFA bias was 0.012 deg$^{-1}$ (SEM 0.0015, p<10$^{-5}$, T-test). A two way ANOVA revealed no effects of rotation speed. The proportion of time seeing the figure spinning clockwise was 0.508 (SEM 0.012), again with no effects of rotation speed. On average, observers experienced 3.68 reversals per minute (SEM 0.25). This frequency was not affected by rotation speed, but we found somewhat less reversals if the camera elevation was 10 deg (mean 3.3 min$^{-1}$, SEM 0.40) as compared to 3 degrees (mean 4.0 min$^{-1}$, SEM 0.49; F(1,23) = 9.3, p<0.01). The VFA bias therefore stabilizes the percept.

The average slope of 0.012 deg$^{-1}$ predicts that at a camera elevation of ± 6.8 degrees, the one used in Kayahara's original silhouette illusion, we should see a “rotational bias” that favours one over the other spinning direction at a rate of 58.2%. The value of 66% reported in the online poll is even larger which is most likely due to much shorter exposure durations. At least for the Necker cube it has been shown that the viewing-from-above bias is much more pronounced for the initial percept and then reduces to smaller values for longer stimulus exposures (Sundareswara and Schrater 2008). At 4 reversals per minute, this initial percept probably dominates the data obtained in online studies like the one cited above.

Last but not least, we should say that Kayahara's original silhouette is in fact a much richer stimulus than it seems. Being a silhouette it is depth-ambiguous because it lacks any cues from self-occlusion. On the other hand, it does contain two subtle depth cues – however, with conflicting information. Tracing the end of the outstretched hand shows that the camera elevation with respect to the hand is only 6.0 deg which is a little less than the value for the foot. That means that the figure was rendered with a perspective camera positioned about 75 meters away from the figure (assuming a vertical distance of 1 m between foot and hand). According to this perspective cue, the view from above and therefore the clockwise rotation is the “true” rotation. However, another cue to the rotation of the figure is provided by the shadow which the feet cast on the ground. The ellipse circumscribed by the shadow of the outstretched foot – at least assuming that the ground is horizontal and that we are looking at it from above – clearly suggests counter-clockwise rotation. There are good reasons for the appeal that Kayahara’s silhouette figure has experienced in the media.

The main purpose of this study was to rectify the mystery that has developed around the silhouette illusion as it gained tremendous public attention as a tool to “probe your brain” on the internet. Reports of “rotational biases” for silhouettes or Necker cubes appear regularly even at established conferences on visual perception as could be experienced at ECVP in Lausanne this year. If the viewing-from-above bias would have received the same awareness and visibility as the light-from-above bias this would have been very unlikely to happen.

All experiments were carried out in accordance with the relevant institutional and national regulations and legislation and with the World Medical Association Helsinki Declaration as revised in October 2008.

We acknowledge funding from NSERC and the help of Stephen Scovil in creating the stimuli.
References

Kornmeier J, Hein C M, Bach M, 2009 "Multistable perception: when bottom-up and top-down coincide" *Brain Cogn* 69 138-47

Mamassian P, Landy M, 1998 "Observer biases in the 3D interpretation of line drawings" *Vision Research* 38 2817-2832

Ooi T, Wu B, He Z, 2001 "Distance determined by the angular declination below the horizon" *Nature* 414 197-200


Ramachandran V, 1988 "Perception of shape from shading" *Nature* 331 163 - 166

Sundareswara R, Schrater P R, 2008 "Perceptual multistability predicted by search model for Bayesian decisions" *J Vis* 8 12 1-19

Troje N, 2010 "Perceptual biases in biological motion perception and other depth-ambiguous stimuli" *Journal of Vision* 10 792

Wallach H, O'Connell D N, 1953 "The Kinetic Depth Effect" *Journal of Experimental Psychology* 45 205-217

Figure 1: Mean and standard errors of the mean of the percentage of “counter-clockwise” responses as a function of camera elevation. Each of 21 subjects was presented with a total of 196 trials. Each trial showed an orthographically projected wireframe cube rotating at 45 deg/s counter-clockwise about a vertical axis. Presentation time was 0.5 s with intertrial intervals of 2 s. The initial azimuthal viewpoint was randomized, but camera elevation changed systematically and adopted values between -30 and +30 degrees. Note that switching the direction of rotation of the 3D cube has the same effect on the 2D stimulus as switching the sign of the camera elevation. For instance, a counter-clockwise rotating cube rendered with a camera at -30 degrees elevation results in the same stimulus as the clockwise rotating cube rendered with +30 degrees camera elevation. Observers clearly prefer the interpretation that assumes a view from above.
Figure 2: The middle column shows four frames of the rotating silhouette of a human figure. The sequence can be interpreted in two ways, depending on the perceived order of the surfaces of the body in depth. The left column illustrates the interpretation in which we look at the figure's back in the first frame. As a consequence, the figure rotates clockwise as we go through the sequence of frames (from top to bottom), and we look at it from a slightly elevated viewpoint (in this case 10 deg). In the right column we illustrate the other interpretation. We look at the front of the figure in the first frame, the figure rotates counter-clockwise, and the viewpoint is 10 deg from below. An animated version of this illustration can be found at http://www.biomotionlab.ca/projects/silhouette.php
Figure 3: Data from a single, typical observer. Separately for the three different rotation speeds, we plotted the percentage of time during which the observer pressed the key indicating counter-clockwise rotation. Camera elevations are indicated with respect to a counter-clockwise rotating 3D figure. The facing-the-viewer bias is quantified in terms of the slope of linear functions fitting the data. Blue: 25 deg/s, red: 100 deg/s, green: 175 deg/s rotation speed.