DOI:10.6129/CJP.202009_62(4).0001
中華心理學刊 民 109，62 卷，4 期，459-472
Chinese Journal of Psychology 2020, Vol.62, No.4, 459-472

Reversion of the Kinetic Depth Effect: Interaction between Homogeneity and Familiarity

Chieh Min and Shojiro Sakurai
Department of Psychology, Kaohsiung Medical University

The kinetic depth effect (KDE) (Wallach & O'Connell, 1953) is a phenomenon in which an observer can perceive 3D rotation from continuously presented 2D projection images. Each frame of a projected 2D image does not contain enough depth cues, so the direction of rotation is unstable in the KDE, and the object is perceived as spontaneously reversing (oscillating). Blackburn (2011) reviewed numerous studies and proposed that adding a temporal or spatial dimension may disambiguate the KDE. Klopfer (1991) used masks of the human face as stimuli in his experiment and found that the masks appeared to oscillate rather than rotate. This phenomenon may be attributed to the familiarity factor, in that observers are more familiar with convex faces than with concave faces. However, the familiarity factor could not explain the perceived reversal of non-face KDE stimuli. This study aimed to identify the general factors that can explain the reversal of rotating masks and non-face stimuli. We hypothesized that KDE stimuli reverse during the alternation of two depth appearances, and that nonhomogeneous features alternate more than homogeneous features. The rotating mask illusion may be attributed to its nonhomogeneity. In this study, Necker cubes and masks were used as stimuli in two experiments to discern whether the relationship between the homogeneity factor and the familiarity factor influences the number of KDE stimuli perceived to reverse. The results of the two experiments support the hypothesis that both the homogeneity factor and familiarity factor influence the perceived reversal of KDE stimuli.

Keywords: *kinetic depth effect, depth inversion, rotating mask illusion*

Extended Abstract

The phenomenon that allows observers to perceive a 2D wire-frame or silhouette figure as a rotating 3D structure is called the kinetic depth effect (KDE) (Wallach & O'Connell, 1953). Without a depth cue, observers generally perceive the object's local and shape correctly, but the rotating direction of a KDE figure may be perceived as either clockwise or counterclockwise. Moreover, observers may perceive the object as spontaneously reversing direction (i.e., oscillations). In a review of a large number of studies, Blackburn (2011) found that adding temporal or spatial context could disambiguate the KDE. The required contexts varied with the rotation speed and complexity of the KDE figure. In other words, Blackburn showed that both rotation speed and complexity influence the perceptual depth reversal of KDE figures.

However, rotation speed and complexity do not explain the reversals of a rotating mask in what is called the rotating-mask illusion (Klopfer, 1991). Klopfer (1991) used both masks of the human face and unfamiliar objects as the stimuli in his experiment. Both masks and unfamiliar objects were shown upright and upside-down. He found that oscillation was more common with masks than with unfamiliar objects and more common with an upright mask than with an upside-down mask.

The results can be attributed to the fact that observers are more familiar with convex faces than with concave faces, and more familiar with upright faces than with upside-down faces. Thus, each time a rotating mask exposed its concave side, it appeared to reverse, leading to a high frequency of reversals. Although the familiarity factor may explain the rotating-mask illusion, it cannot explain the perceived reversals of non-facial KDE figures or unfamiliar objects. (e.g., Necker cubes.)

The purpose of this study was to identify the general factors that can explain the reversal of both rotating masks and non-facial stimuli. In recent studies, Blackburn (2011) revealed the importance of homogeneity in KDE. Non-homogeneous features disambiguate the KDE and alter the frequency of reversals. Schwartz and Sperling (1983) also investigated the homogeneity of the luminance of Necker cubes and found that observers perceived the bright lines as close and the dim lines as far away.

In the case of human faces, almost all of the details are in the front of a face. In other words, faces are non-homogenous features. We hypothesized that KDE figures reverse when the depth appears to change, and that non-homogeneous objects alternate more frequently than homogeneous objects because the more detailed part of an object is perceived as closer than the less detailed part.

Specifically, we hypothesized that homogeneity of features influences the frequency of reversals, and that objects with non-homogenous features would reverse more frequently than those with homogenous features. Thus, the non-homogeneity of the human face could explain the rotating-mask illusion. In this study, Necker cubes and masks of the human face were used as stimuli in two experiments to determine whether the interaction between homogeneity and familiarity influences the reversal frequency of a KDE stimulus.

Study 1

In Study 1, we used Necker cubes with different levels of homogeneity as the stimulus. The Necker cubes varied in a 2 (homogeneity) X 3 (complexity) factorial design. A thick borderline was added to either one face or two opposite faces of a cube. The former was the non-homogenous condition, and the latter was the homogenous condition. The complexity factor was manipulated by varying the number of borderlines from 1 to 3. More complexity increased the contrast between the homogenous and non-homogenous conditions. The two factorial design yielded six conditions. Each condition was presented once for a 90-second trial at a rotation speed of 60 rpm. The sequence of the conditions was randomized. All of the stimuli were generated with Blender and run under PsychoPy.

Twelve participants were used in Study 1. They were instructed to indicate the perceived reversals of the Necker cubes by pressing the space key and to report unclear percepts by pressing the S key on a standard keyboard. We calculated the frequency of reversals or unclear percepts and found that non-homogenous cubes reversed more frequently than homogenous cubes ($F(1, 11) = 5.531$, $p = .038$, $\eta^2 = .335$), which was consistent with our hypothesis. However, the results also showed that complexity had no significant effect on the frequency of reversals ($F(2, 22) = .512$, $p = .606$, $\eta = .045$).

Three participants reported unclear percepts in the homogenous condition. Although there were no significant differences in the number of unclear percepts between the conditions, the participants reported orally that they perceived the homogenous cubes as collapsing; specifically, the two opposite faces of a cube rotated in different directions.

Study 2

Building on Study 1, in Study 2 we used masks of the human face with different levels of homogeneity as the stimulus. Study 2 had a 2 (homogeneity) X 4 (familiarity) factorial design. We varied the homogeneity level by using one-sided (non-homogenous) or double-sided (homogenous) masks. We varied the familiarity levels by using four kinds of masks: normal masks, upside-down masks, upside-down messy masks, and upside-down messy masks with geometric organs. The closer a mask was to the faces we see every days, the higher the familiarity level. The two factorial design yielded eight conditions. We hypothesized that homogeneity would influence the frequency of reversals and that familiarity would have no significant effect.

Sixteen participants were used in Study 2, and the procedures were the same as in Study 1. We calculated the frequency of reversals or unclear percepts and found that homogeneity had no significant effect on the frequency of reversals in Study 2 ($F(1, 15) = 0.071$, $p = .793$, η^2

= .005). This result was inconsistent with the results in Study 1. Furthermore, familiarity played a significant role in reversals ($F(3, 45) = 15.004$, $p = .000$, $\eta^2 = .500$). The normal masks reversed much more frequently than the other masks. In conclusion, higher familiarity increased the frequency of reversals. Unfortunately, both these findings failed to support our hypotheses.

Three participants in Study 2 reported unclear percepts by pressing the S key in the homogenous normal mask condition. Although there were no significant differences in unclear precepts between the conditions, the participants reported orally cases in which the homogenous normal masks collapsed; specifically, the two opposite sides of a mask rotated in different directions.

Discussion

To explain the opposite effects of homogeneity in our two studies, we suggest that two different depth cues in a KDE figure compete with each other, and that the KDE reverses when the observer switches attention from one depth cue to another. Homogeneity influences the competition between the two images, because in a non-homogenous object one cue is stronger than the other, and thus is more likely to "win" the competition. Specifically, the part with more detail will always be seen as closer. A winning streak of one view of a KDE figure leads to a high frequency of apparent alternations in depth.

However, in the homogenous condition, we also must consider the intensity of the competition. For example, in the homogenous Necker cubes trial the competition between the two relative depths was weak, and there was no winning streak. Therefore, the homogenous Necker cubes rotated without reversal and appeared to rotate rather than oscillate. However, the result was very different for the homogenous masks. The two sides of the masks had highly distinct apparent depth, which increased the intensity of the competition between the sides. This high-intensity competition might have led to a high frequency of alternation in apparent depth, making the double-sided masks appear to oscillate rather than rotate. The familiarity factor might have contributed to the intensity of competition, as observers are more familiar with convex faces (front of the mask) than with concave faces (back of the mask), whereas the level of familiarity did not change between the opposite faces of a cube.

This could also explain the unclear percepts in both studies. For some observers, the 3D structures collapsed because of the high-intensity of the competition. Specifically, both sides of an object were perceived as being near the observer at the same time, making them appear to rotate in different directions as they become closer to the observers.

Thus, we can say that both homogeneity and familiarity influence the reversal of a KDE stimulus. Yin (1969) recently showed the importance of familiarity in facial recognition. The phenomenon that describes why comparing upside-down faces with upright faces is more difficult than doing the same task using non-facial objects is called the face inversion effect. The face inversion effect suggests that faces and non-facial objects are not processed in the same way. Klopfer (1991) also used differential familiarity to explain the rotating mask illusion. However, Chen (2009) revealed that familiarity is not the only reason for depth inversion. In other words, both bottom-up processing and top-down processing are involved in depth perception. Future studies investigating the influence of familiarity on depth perception must first make their operational definition of familiarity clear.

Recovering a 3D structure from a KDE figure is similar to the process used to transform retina images into a 3D world. Therefore, the KDE is an inverse of the problems in optics. Depth cues allow us to perceive depth in a 2D images (e.g., paintings or projections on the retina). Although the KDE figures lack a depth cue, the visual system can recover a stable 3D shape from an ambiguous projection using the information generated by the relative motion or perceptual tendency. For example, Schwartz and Sperling (1983) showed that observers tend to see the bright parts of an object as closer than the dim parts.

Similar effects were observed in rotating non-homogeneous Necker cubes and rotating masks. Our study shows that observers tend to perceive the parts of an image with more detail as closer than the parts without detail. In the absence of a depth cue, as in the recovery of retina images, the perceptual tendency (i.e., familiarity and level of detail) plays a more important role in recovering the 3D structure in the KDE. Conversely, our results clarify the importance of perceptual tendency in depth perception. Future studies should use diverse stimuli to identify the cooperative and competitive relationships between depth cues and the perceptual tendency of depth perception.