UNIVERSITY OF CALIFORNIA, SAN DIEGO

Rotation Reversals in the Kinetic Depth Effect

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Psychology and Cognitive Science

by

Shai David Azoulai

Committee in charge:

Professor Stuart Anstis
Professor Garrison Cotterll
Professor Seana Coulson
Professor Karen Dobkins
Professor Donald MacLeod, Chair

2014

The Dissertation of Shai David Azoulai is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2014

DEDICATION

To my family, who gave me unwavering support and love during my long graduate school career.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

KDE – Kinetic Depth Effect

RVF – Right Visual Field

LVF – Left Visual Field

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor Donald MacLeod for standing by me through thick and thin, for supporting me when I needed help, for pushing me when I needed encouragement, for all the work he put into helping me with my graduate career, and for being both a fantastic advisor and a wonderful friend.

And finally, I would like to thank my doctoral dissertation committee for the excellent feedback they provided which greatly improved the quality of the submitted work.

VITA

1999	Bachelor of Science, Computer Science, Georgia Institute of Technology
2003	Master of Science, Psychology, University of California, San Diego
2003-2013	Teaching Assistant, University of California, San Diego
2014	Doctor of Philosophy, University of California, San Diego

FIELDS OF STUDY

Major Field: Psychology (Cognitive Neuroscience)

Studies in Neural Representation of Body Image

Professor V.S. Ramachandran

Studies in Synesthesia

Professor V.S. Ramachandran

Studies in Vision and Perception

Professors Donald MacLeod & V.S. Ramachandran

ABSTRACT OF THE DISSERTATION

Rotation Reversals in the Kinetic Depth Effect

by

Shai David Azoulai

Doctor of Philosophy in Psychology and Cognitive Science

University of California, San Diego, 2014

Professor Donald MacLeod, Chair

Multistable visual illusions have been utilized by artists for hundreds of years, fascinating many people by the seemingly random ways that their perception can shift even when the stimuli they are viewing remain constant. Rather than being completely random, there are specific points at which people are more likely to perceive changes when viewing multistable stimuli, such as the Kinetic Depth Effect (KDE) silhouettes used in the following

experiments. This allows us to use KDE stimuli as a tool to explore the visual system to determine the properties of those stimuli which result in non-random perceptual rotation reversals. These properties can then give us insight into the inherent biases that may exist in the way that our brains process visual information.

Introduction

-What are multistable illusions and why are they interesting?

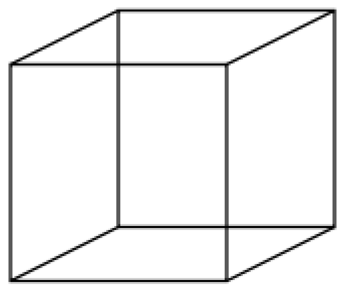


Figure 1: Necker Cube

The first multistable illusions that many people encounter is the Necker Cube shown in Figure 1 (Necker, 1832). The Necker cube is considered to be a multistable illusion because our perception of the cube can switch between alternative 3D interpretations despite the fact that the image itself does not change. In this case it is possible to see two different orientations of the cube depending on which side is perceived to be in front. The lack of a fixed depth for the faces of cube creates an ambiguity for our visual system and thus the visual 'problem' of the Necker Cube has two equally likely solutions. As our visual systems switches between these solutions we experience a change of perception and thus a multistable image.

The perceptual shifts that occur when viewing multistable images are of particular interest to vision scientists because they represent a change our conscious perception without any change in the stimulus. The images become multistable in our consciousness when the brain reinterprets the stimulus and our perception shifts from one valid interpretation to another. The reason for these perceptual shifts differs from illusion to illusion, but they all have one trait in common, which is that a single stimulus can produce multiple valid interpretations.

Currently there is a debate in the literature as to the exact cause of perceptual multistability. Within this debate there are two major questions:

1) What is the root cause of perceptual multistability?
2) Are all forms of perceptual multistability governed by the same processes or do different multistable illusions have different neurological bases?

In trying to answer the first question, one of the most studied forms of perceptual multistability is binocular rivalry, in which multistable perception is induced by presenting each eye with a different image (Wheatstone, 1838). Rather than combining these two images into something that may be incomprehensible, the brain instead produces a percept of the image from a single eye. As time passes the percept will switch between the images presented to each eye individually.

A research advantage to this type of multistable illusion is that the information presented to each eye can be carefully controlled and modified. This has allowed researchers the ability to determine whether or not certain stimuli are more likely to result in conscious perception by a subject as well as whether or not volitional control could modulate the percept (Breese, 1899; Breese, 1909). For example contrast, spatial frequency, and luminance have all been shown to strongly modulate perception of multistable stimuli (Alais & Blake, 2005; Fahle, 1982; Kang, 2009; Levelt, 1965).

The strong level of control of what is presented to each eye has also allowed for some very clever research aimed at understanding the neurological underpinnings of the perceptual switches.

Kang and Blake (2010) showed that the perceptual shifts from binocularly rivalrous stimuli may be due to a combination of adaptation to the existing percept and inhibitory cortical circuitry. They reached this conclusion by presenting subjects with alternating periods of rivalrous stimuli between the two eyes, and adaptation, where only a single eye was presented with a stimulus. During periods of adaptation the subject would always perceive the adapting stimulus. When the rivalry portion of the experiment began the subject would experience a period of dominance of the previously adapted stimulus.

Using different duration lengths for the adaptation portion of the experiment, the data showed that as the length of the adaptation *increased* the subjects experienced a *decrease* in dominance time for the adapted stimulus. This showed that adaption likely plays a role in binocular rivalry and that perceptual switches could not be explained by "neural noise" alone has has been proposed in other models (Moreno-Bote, Rinzel, Rubin, 2007).

The neural noise model is further weakened by recent twin studies which have shown genetic similarity can have a large impact on perceptual switches in binocular rivalry. Recently Miller et al (2010) released the results of a large twin study which explored the variance in binocular rivalry rates across monozygotic and dizygotic twins. Their data showed that switch rates of the multistable percept were highly correlated with genetic similarity and that up to 52% of the variance in binocular rivalry rates across subjects might be accounted for by genetics. In 2011 Shannon et al (2011) conducted a similar study which replicated the binocular rivalry results of the Miller study and also showed that monozygotic twins had a highly correlated switch rate for a normally viewed Necker Cube whereas dizygotic twins did not have similar switch rates for either binocular rivalry or a Necker Cube stimulus. Shannon's result is particularly interesting because it suggests an answer to the second big question related to multistable images; that there may be some shared neural circuitry for perceptual switching or at least that a similar mechanism may underlie different types of perceptual ambiguity.

Regardless of the exact reasons of why we experience perceptual switches, one common thread throughout the literature is that there is an inherent randomness to exactly when a perceptual switch will occur. While most of these studies are able to find correlations between switch rates or ways to modulate the stimuli to show when one is more likely to appear, there seem to be no hard and fast rules as to what will cause a perceptual switch or when that switch will occur.

-Is all multistable switching random?

Up until this point much of the research on multistable perception has focused on binocular rivalry and how other forms of multistable perception may share a neural basis with binocular rivalry. Much of this research talks about "switch-rate" and makes comparisons of that rate across subjects groups to prove a point. Other times "neural noise" and randomness are introduced to explain the mercurial nature of perceptual switching. But recently we came across a type of multistable image which may be behaving differently. In these images there appear to be very strong patterns across subjects as to exactly when a perceptual switch will occur. Furthermore as subjects are exposed to these images over time their “switch-rate” seems to change and the final “switch-rate” seems to be a factor of exposure to these images rather than an endogenous trait of perceptual multistability.

- *Kinetic Depth Effect*

Kinetic Depth Effect (KDE) illusions are particularly striking ambiguous images which are made from a rotating silhouette of a 3-Dimensional object and which have been recently popularized by a “spinning ballerina” illusion produced by Nobuyuki Kayahara in 2003 (Troje & McAdam, 2010). First described by Wallach & O’Connell in 1953 (Wallach & O’Connell, 1953), KDE images are most notable for the fact that perceptual reversals result in a change in the direction of rotation of the stimulus.

To elicit the basic phenomenon all one needs is a rotating silhouette of a 3D object of sufficient complexity. The visual system will easily pull out a full 3D structure from the rotation (Bennett et al, 1989; Huang & Lee, 1989). However since the object is a silhouette (a single solid color) most depth cues, including those due to shading and occlusion have been eliminated. Therefore the depth of the image is ambiguous and the viewer cannot distinguish which portions of the silhouette are in the distant hemisphere of the rotation and which are in the closer hemisphere (Marr & Nishihara, 1978; Richards, Koenderink, & Hoffman 1985; Kleffner & Ramachandran, 1992).

Further ambiguity is added if the stimulus is presented in a parallel projection as opposed to a perspective projection. In a perspective projection the size of the different parts of an image on screen are directly related to the distance of those parts of the image to your eye (closer = bigger). The parallel projection eliminates depth cues that come from a difference in image size of all portions of the object as it rotates. This means that the image of a portion of the object will remain a constant size as a portion of the object rotates into positions where it is close to you or far from you. In a parallel projection all the rays from the object to the viewport are parallel to one another, regardless of which part of the object they originate from. This results in a total loss of the depth perspective cue. In cases where a perspective projection is used instead of a parallel projection, rigid objects will appear to deform as they rotate when the viewer incorrectly reverses the depth polarity of the object. But when the depth is correctly assigned the object will appear to rotate rigidly. Thus the ambiguity of a KDE stimulus in a parallel projection comes from a complete lack of any depth information due to shading, occlusion, or perspective.

Every time there is a reversal in the assignment of depth cues there is a concurrent reversal in the direction of rotation of the KDE stimulus. Vice-versa a reversal in the direction of rotation of the stimulus will be accompanied by a reversal in the assignment of depth cues (except in cases where there is depth symmetry at the time of the direction reversal). The overall transformation that occurs is the same a mirror reflection in depth, which is equivalent to a left-right reflection and a 180-degree turn about the axis of rotation. This is why a direction reversal at a point in the revolution where there is depth symmetry does not require any transformation of the object, just a change in the direction of motion.

-Multistability from Geometry

Before proceeding any further it is important to consider the underlying causes of multistability of certain ambiguous figures. And it will be instructive to follow the approach that Marr used to solving the problems of vision.

Marr outlined a three level approach to approaching vision problems (Marr, 1982):

1) [Computational] - Identifying a useful computation to be solved, in this case a 2D to 3D transformation of an image

2) [Representation & Algorithm] - Finding a way to represent the problem and an algorithmic method of solving it
3) [Hardware] - And finally determining how the algorithm could be implemented in reality with physical hardware (biological or silicon)

Consider first the computation that must occur for a 2D to 3D transformation to take place. An inherent weakness of the human visual system is its reliance on producing 3-dimensional perception from 2-dimensional images projected onto the retina. When this projection of visual information from 3D to 2D occurs there is a loss of accurate depth information from the world. The visual system however is able to infer depth using various cues within the image on the retina (occlusion, motion parallax, relative size, perspective, shape from shading, etc.). All of these depth cues are based on assumptions that the visual system makes about the physical properties of the real world. It is these assumptions that allow the visual system to (mostly) accurately transform the one-to-many mapping of a 2-dimensional retinal image into a 3-dimensional percept.

However it is quite easy to fool the visual system by creating stimuli which intentionally break the assumptions it makes about the world. A famous example of such a stimulus is the Ames Room (Ames, 1952). The Ames room is designed to use the assumptions your visual system makes,

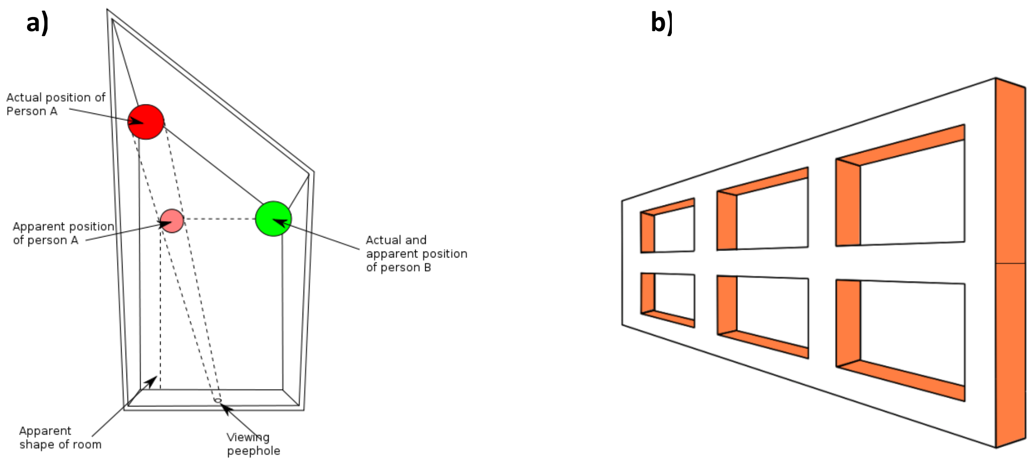


Figure 2: a) Ames Room b) Ames Window

about perspective and right-angles, to create an illusion of impossible size differences. In reality walls of the Ames room do not meet at 90 degree angles and objects or a gradient pattern are usually presented to give additional depth cues that correlate with the impression of a normal square room.

In a similar fashion the Ames window fools your visual system by using perspective to create the illusion of a rectangular object. In actual fact the shape of the window is a trapezoid. The illusion is created by rotating the window continuously in the same direction. Instead of perceiving a rotating trapezoid the visual system perceives a rectangular window whose direction of rotation changes from time to time. The illusion is so strong that a rod can be place through the window, and yet viewers are more likely to see the rod appear to go through a solid rectangular window rather than to perceive the correct shape of trapezoidal window and rod rotating in the same direction. This shows that misappropriation of a single depth cue (in this case perspective), is strong enough to result in a percept that is physically impossible.

These illusions occur because your visual system solves the 2D to 3D problem of depth incorrectly. When the visual system fails to correctly map the 2D retinal image to a 3D world the result is not a completely failed percept of the world, but a slightly broken one where size and distance are not what they should be. From Marr's perspective there was a very specific failure in the algorithm that was used to perform the transformation and the assumptions that it made about the world.

However understanding that the transformation failed here only gets us partway to understanding why that failure occurred. According to Marr the algorithm goes hand in hand with the way that the data are represented. To that end we must consider how 3-dimensional information is stored by our visual system. Sinha & Poggio, gave a clue to that in 1996 when they showed that learning could affect the percept of an ambiguous figure. Specifically they were able to demonstrate that pre-learned knowledge of the shape of a

3-dimensional object could influence perception when that object was presented in the same view in which it was learned. However, if that object was presented in a novel view the prior learning had a much lower effect on the ultimate percept. This suggests that 3-dimensional visual information is not directly stored in our visual system as full representation of the object, but rather that at least some of that information is implicit and view-based, potentially requiring the visual system to construct a 3-dimensional representation of an object from memory, in the same way that it would do so from a 2D retinal image (Sinha & Poggio, 1996). This idea is quite similar to one proposed by George Berkeley in the early 18th century. Berkeley suggested perception did not work in absolutes, and that instead the same cues that evoked size were also used to calculate distance. This meant that our final percept of concepts such as distance were made up of many interacting cues that all worked together create a representation of reality. This is in contrast to others who suggested that we have some pre-existing knowledge about the world and use that knowledge as a basis for making subsequent calculations of other properties (Ross & Plug, 1998).

This is all relevant to the work which will be presented here because it lends insight into the way that the visual system must process the 3-dimensional objects that will make up the stimuli in the coming experiments.

-What is so interesting about the Kinetic Depth Effect?

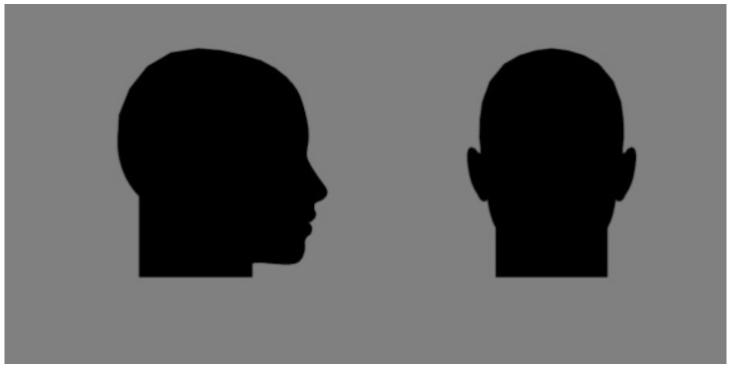


Figure 3: Profile and frontal view of a face silhouette

While the Kinetic Depth Effect is certainly a ‘cool’ phenomenon, why is it worth studying and useful for understanding how the brain processes multistable illusions?

From a probabilistic standpoint, let us assume for a moment that there are no biases which regulate switching of KDE images. If no biases are present then the image will have an equal probability of undergoing a perceptual switch at every point in the rotation and neither the position of the switch nor the new direction of motion would matter. However if biases do exist then they would likely be due to a combination of high-level and low-level processes related to object recognition, motion perception, and the amount of contour information in the image at any given time. For example, Figure 3 shows the silhouette of a face in profile vs. facing forward/backward. In profile one can clearly tell if a silhouette is facing left or right, but a front/back facing view is ambiguous). Since we can construct KDE images from any 3D object this allows us to introduce multiple levels of complexity into the images as well as to explore different high-level biases easily.

The work to be described shows how KDE ambiguous figures can reveal that perceptual shifts are probabilistic but to some extent systematic and that there may be multiple mechanisms at work which can increase or decrease the probability of a perceptual switch at any given time. Not only do KDE stimuli offer the viewer a much more drastic and obvious perceptual change than traditional static multistable images such as the Necker Cube, but because the change in perception always includes a reversal of the direction of rotation, it is very easy to detect a perceptual change in KDE image which is presented parafoveally. This gives us three major benefits. First it allows us to use KDE images to explore the possibility of a hemispheric difference in perception. Second, motion reversals of KDE image are very obvious even when not viewed directly, whereas subjects have difficulty reporting perceptual switches of static multistable images viewed non-foveally. Thirdly, it minimizes the effects of foveating on or attending to a specific part of the image, actions which are known to induce perceptual switches in other multistable images (Necker, 1832; Glen, 1940). In the Necker cube these perceptual switches have been shown to be strongly correlated with the eyes being near the distant corner of the current percept. When the percept switches this distant corner becomes the closest and if the eyes eventually move to the new "distant" corner this can once again induce a perceptual switch (Einhäuser, Martin, & König, 2004).

Our research will show that KDE movies have a very strong pattern of when perceptual switches occur. It will furthermore demonstrate that perceptual "switch-rate" is not a static property in the case of KDE moives, and this may require a rethinking of much previous research which used "switch-rate" as a major metric of multistability.

Furthermore we will show that these differences from other multistable illusions may be due to inherent biases in the way that our visual system processes images as minor changes to our stimuli were able to produce strong differences in when subjects experienced perceptual switches. Finally, we will introduce evidence which suggests that there may even be a hemispheric bias to how the brain processes KDE movies.

Simply put, the way that our perception changes when viewing a KDE movie is a window into the inner workings of our visual system and the biases and "shortcuts" that our brain uses to process visual information. Thus not only can the Kinetic Depth Effect give us insight on perceptual multistability, but it also allows us to explore how we process three-dimensional visual imagery.

Experiment Set 1 – The Basic KDE Phenomenon

Experiments

The goal of the following experiments is to explore whether or not KDE images undergo directional switches randomly, to explore possible low level and high level biases related to these directional switches, and to determine if there are any general rules related to how and why the brain's perception of an ambiguous stimulus will change.

The idea behind the first series of experiments is to show that directional switches of KDE images are non-random and to establish that patterns of directional switching are directly related to the types of images shown.

Experiment 1.1

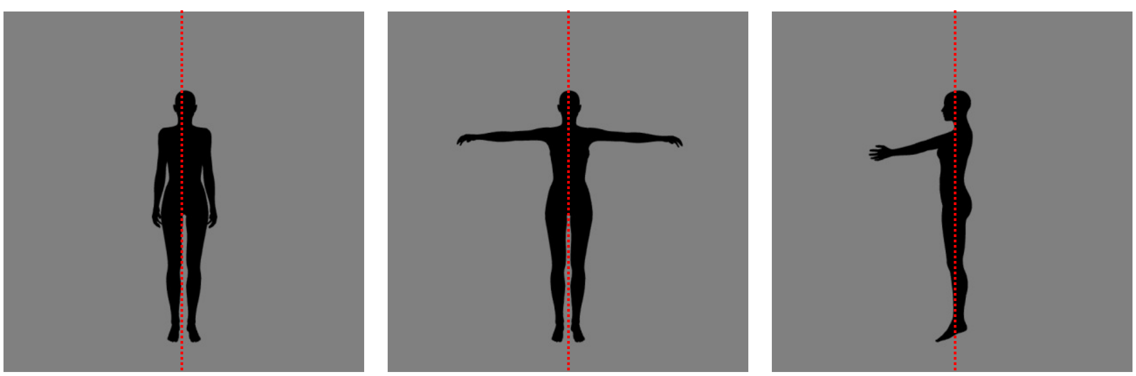


Figure 4: The three human poses used for Experiment 1.1. Each was a KDE stimulus which rotated along a vertical axis going through the midline of the body as indicated by the red dashed line

For this experiment we used KDE images of a human body in different poses. These were chosen because we felt that they might bring out some high level biases which could influence switching and would thus be more likely to show a non-uniform pattern of switches around the 360-degree rotation.

Experiment 1.1 - Methods

Stimuli consisted of a 3D human female model in three different postures with Left-Right symmetry but clear differences between front and back such that it was trivial to determine which way the image was facing. All of the images were computer generated using DAZ Studio 3. To avoid physical distortions due to perspective, all of the images, both for this experiment and the others were presented in a frontal orthographic projection.

Orthographic Projection

An orthographic projection is a means of presenting an image so that all projection lines are parallel to each other and orthogonal to the projection plane. The resultant image projected with an orthographic view will thus not alter the size of the objects presented relative to their distance to the viewer. This compares to a perspective projection where closer objects will appear to be larger.

By using orthographic projection we eliminated the possibility of image distortion due to size differences as the objects rotated. If the images had been presented with a perspective projection there would have been a noticeable change in size of portions of the silhouettes as they rotated and certain parts of the object moved towards and away from the subject.

Stimuli

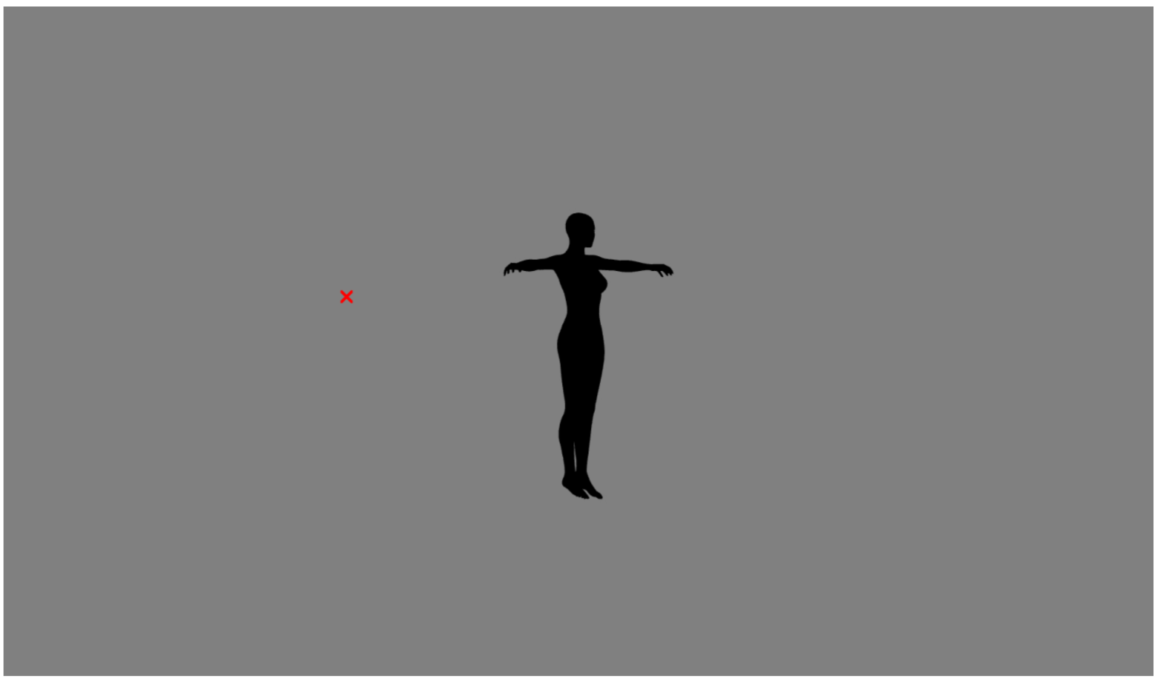


Figure 5: This is an example of how the stimuli presented in the experiment

Each stimulus consisted of 41 views (frames) representing a full revolution of a 3D object. This resulted in an angular rotation between successive frames of 8.78 degrees. The odd number of frames used was due to a limitation within the program used to create the images (DAZ Studio 3).

Later experiments were able to double the number of frames to 82, but the inherent limits of DAZ Studio 3 did not allow us to directly pick how many frames we wanted to use for each image series.

The frames were presented at a rate of 7.5 frames per second which resulted in one full revolution of the object every 5.47 seconds. This framerate was later doubled for the stimuli which contained 82 frames, which maintained the overall rotational speed of one revolution every 5.47 seconds.

The background color of the stimuli was a grey. The 3D silhouettes were a pure black.

Subjects were seated 26 inches from the monitor. At this distance the monitor subtended a viewing angle of 31.2 degrees horizontally and 14.9 degrees vertically. The images subtended between 2.2 to 8.8 degrees of horizontal visual angle depending on the frame, and 8.8 degrees of vertical visual angle.

Subjects

28 undergraduate students (10 males and 18 females) at UCSD were presented with a rotating silhouette that utilized the Kinetic Depth Effect to create an ambiguity in the direction of rotation. Subjects were compensated for their time with class credit for their current courses. Seven of the

subjects were eliminated from the study for failing to follow the experimental protocol.

Instructions

Subjects were instructed to indicate the initial direction of motion of the stimulus (clockwise or anti-clockwise as viewed from above) when it first appeared with a keypress [<] or [>]. Thus [>] indicated anti-clockwise rotation as viewed from above, and [<] indicated clockwise rotation as viewed from above. They were told to then press the [<] or [>] key to indicate the new direction of rotation whenever a directional change in motion occurred. They also had the option of pressing the [SPACEBAR] if the image started to do something that could not be described as rotating.

Image Presentation

Stimulus presentation order was randomized for each subject into blocks. Each stimulus was presented in the center of the screen for the duration of each trial block. A single block consisted of three viewing conditions each 120 seconds long. First the subject was allowed to freely view the stimulus for 120 seconds. The next two blocks were also randomized and lasted 120 seconds during which time the subjects would have to foveate on a fixation cross to the left or right of the stimulus. There was a one second interval, where the screen would blank out, except for the fixation cross, between stimulus conditions in a single block. One second after the screen blanked out the ambiguous figure would reappear. After each block of Center/Left/Right viewing the subject was given an option to briefly take a break and continue the experiment when they were ready.

All of the experiments present the images to the subjects in a free-viewing condition and in conditions where they are foveating on a fixation point to the left or right of the image. The reason for this is that research has shown that fixating on specific parts of a multistable image can induce or suppress perceptual switches. To avoid this confound data was collected in left visual field, right visual field, and central viewing conditions.

Data Collection

The data collected was the current frame being presented when the computer registered the subjects' response indicating they noticed a change. Thus there was a slight lag between the time the subjects noticed a change, decided to press a button, and completed the button press. There was likely a reaction time delay of about 400-500ms before the button press was registered (Kornmeier & Bach, 2012). This translates into a delay of at least 4 frames of the stimulus. Thus the peaks in the data presented below occur slightly *after* the subjects experienced a directional switch.

To confirm whether or not subjects were responding slowly due to a reaction time lag a control experiment was conducted which presented subjects with a single KDE image that was presented at two different speeds. The results of this experiment, shown later, confirm that there seems to be a consistent response-time lag which can at least partially account for subject response peaks at non-cardinal points of the rotation. There is also likely a small delay from image processing.

The total run time for each subjects was 18 minutes (6 minutes per stimulus), this allowed for 66 full rotations of each stimulus with a total of 198 full rotations across all stimuli.

The total run time of the experiment across all subjects whose data was used (21 of 28 subjects) was 378 minutes, the total number of revolutions across these subjects was 4,146.

At the end of the experiment subjects were asked to debrief the experimenter. This consisted of verbal descriptions of what they saw and questioning as to specific positions where the images seemed to change direction.

Statistics Notes

The majority of the data analysis was done using several circular statistical tests.

One of the primary tests used for circular statistical analysis is Rayleigh test for non-uniformity of circular data. The Rayleigh Test works on the principle of considering circular data points to be individual vectors. The test asks how large the resultant summed vector length must be to indicate a non-uniform distribution. This test detects only a unidirectional imbalance, and does not distinguish between a uniform distribution and a symmetrically bilobed distribution for example. It is the optimal test if the deviation from uniformity takes the form of a von Mises distribution (I_o is a modified Bessel function of order 0),

$$f(\theta;\mu,\kappa) = \frac{e^{\kappa\cos(\theta-\mu)}}{2\pi I_0(\kappa)}$$

Whereas the circular uniform distribution is:

$$U(\theta) = 1/2\pi$$

While the Rayleigh test is powerful for unimodal distributions can easily fail to find non-uniformity in multi-modal distributions, particularly if there are peaks at opposite poles. In these cases we used the Omnibus/Hodges-Ajne (Zar, 1999) test for non-uniformity of circular data. The null hypothesis of Hodges-Ajne is that the data is uniformly distributed around the circle and test is sensitive to any non-uniformity as an alternative hypothesis. It tests for non-uniformity by considering the least number of data points that can be found in any 180 degrees of the data set. It then compares this value to a critical minimum number based on the total number of data points, and if the number of data points found is less than the critical minimum a non-uniformity is inferred. Hodge-Ajne is not as powerful as the Rayleigh test for detecting non-uniformity of a unimodal distribution. While it is slightly less powerful than the Rayleigh test, it does not make any assumptions about the underlying distribution of the data and therefore is works well for data where we cannot make assumptions of a von Mises distribution or unimodality.

The final major test that we used was the Kuiper two-sample test. This tests whether or not two sample circular data sets differ significantly. The difference can be in any property, such as vector mean, location, and dispersion. It is a circular analogue of the Kolmogorov-Smirnov test. Like the Kolmogorov-Smirnov test, Kuiper uses the discrepancy statistics D^+ and D^- defined as the largest positive and negative differences of two cumulative distribution functions. However the Kuiper test sums these two unsigned statistics to make the test sensitive along the entire distribution. This test was useful in cases where we expected to find minor differences between two similar data sets, especially when those data sets were not unimodal.

A final note, there are two ways to define standard deviation in circular statistics. Angular deviation which is bounded within $[0,\sqrt{2}]$, and circular standard deviation which is unbounded and can be between $[0,\infty)$. There seems to be a preference for angular deviation because of the limited bounds and therefore this is the measure used (Zar, 1990).

Angular deviation "s" is computed as follows:

First the data is transformed into unit vectors and averaged

$$r_i = \begin{pmatrix} \cos \alpha_i \\ \sin \alpha_i \end{pmatrix}$$

$$\bar{r} = \frac{1}{N} \sum_i r_i$$

This gives the mean resultant vector $\bar{r}$ which has directionality and a length of between 0 and 1. The length of the mean resultant vector ($\bar{r}$) is R $= |\bar{r}|$. Because R varies between 0 and 1 it is a useful measure as an R value close to 1 indicates a highly concentrated distribution at a single modality.

Finally angular deviation is calculated using R as follows, s = $\sqrt{2(1-R)}$. The larger the value for R, the more highly concentrated the distribution, and the smaller the angular deviation. It is this final calculation which gives angular deviation a range of 0 to $\sqrt{2}$.

In determining which test to use, the following logic was applied:

The most powerful of these tests is the Rayleigh Test for non-uniformity and this would ideally be what we use wherever possible. In order to determine whether or not the Rayleigh test would be appropriate we first looked at the circular distribution of the data. In cases where there appeared to be a unimodal distribution Rayleigh was used. Normally it might be suspect to decide to use a statistical test post-hoc after looking at the data. However because of the way that the Rayleigh test works we would actually be severely punished in statistical power for an incorrect assumption of a unimodal distribution. Therefore if we use this test with incorrect assumptions it is extremely unlikely we would find a significant result. However just to be sure of the results, every data set which showed statistical significance with a Rayleigh test was also tested using the less-powerful Hodges-Anje test. . It turned out that in every case where the Rayleigh test showed significance the Hodges-Anje did as well.

For data sets that appeared to have a bimodal or multimodal distribution we can easily use the Hodges-Anje test for non-uniformity. In this case we again looked at the data prior to choosing the test, however this test requires no underlying assumptions about the data and is significantly weaker than the Rayleigh test. So once again we felt safe using it post-hoc.

Finally for the Kuiper test, it was a simple matter of wanting to know if two distributions were similar. In this case we didn't need to make any underlying assumptions about the data and Kuiper simply tests for any major difference between two distributions examined.

Data Confirmation

Finally it should be noted that each of the positive results reported in Experiment Set 1 and Experiment Set 2 here were confirmed via a subject by subject analysis. This was done two ways. First identical statistical tests to those performed on the entirety of the data set were conducted on the data of individual subjects for each condition.

<u>For Experiment Set 1</u>

131 of 176 cases the results remained significant even for a single subject. Additionally the individual subject means angles were centered, as expected, around the group mean angle. Of the 45 cases where the results were not significant: 25 had a 'small n', due to a low number of perceptual

switches (less than 10); 5 had a p<=0.10; and the remaining 15 were likely due to noise.

As an additional test the mean sin and cos were calculated for each individual subject on each experimental condition. If the subjects did have a bias in switch direction then either the mean sin or mean cos would be significantly different from zero on a per subject/per condition basis. This was confirmed for every single positive result with p<0.025 or better and in the expected direction.

<u>For Experiment Set 2</u>

The mean sin and cos were calculated for each individual subject on each experimental condition. If the subjects did have a bias in switch direction then either the mean sin or mean cos would be significantly different from zero on a per subject/per condition basis. This was confirmed for every single positive result with p<0.025 or better and in the expected direction. The only exception was for the movie of a human with one arm out to the side. However a further analysis of this data revealed a bimodal distribution of perceptual switch angles separated by 180 degrees. This was not sufficient to overcome the bias in the overall mean direction, but did significantly weaken the cos/sin t-test.

The per subject / per condition statistical tests run on experiment set 2 had a large number of non-significant results. These results seem to be due to the fact that most of the data in Experiment Set 2 either had bimodal distributions which prevented useful statistical testing. Additionally many of the conditions had $n<5$ perceptual switches making it very hard to reach statistical significance of any kind.

However the consistency of the mean cos/sin results suggests that the less-than-useful individual r-tests were primarily a result of the distribution of the data and not an overall inconsistency between total mean results and individual results.

<u>**How to read the results**</u>

-*<u>Anti-Clockwise Reversal</u>* means that a subject saw a directional switch and the NEW direction of motion was anti-clockwise as viewed from above.

-*<u>Clockwise Reversal</u>* means that a subject saw a directional switch and the NEW direction of motion was clockwise as viewed from above.

1st Row: A single image of the stimulus is shown. The circular graphs presented show a graph which indicates how many times subjects reported a directional switch at a particular point within the rotation. The data presented in the graphs are binned totals of the actual number of rotation reversals the subjects experienced. The bin widths are 8.8 degrees, which cover a full single frame of the image. The clockwise and anti-clockwise reversals are each presented in a separate graph.

2nd Row: The graphs with the dashed red line show an overlay of the linear image motion at every point in the rotation as calculated using a frame-by-frame difference model described towards the end of the document. The dashed red line represents the relative linear image motion at each frame and has been scaled to fit on a graph with the subjects' responses.

3rd Row: The frames which indicate the peak point (statistical mode) at which subjects reported a directional switch. Below these peak images are presented some basic statistical properties of the graph above, the mode and mean are presented in degrees, and finally p-value is presented for the relevant circular statistic indicated, in this case the Rayleigh test.

4th Row: The four images at the bottom of the figures indicate the orientation of the stimulus at the 0, 90, 180, & 270 degrees of rotation.

Experiment 1.1 - Results & Conclusions

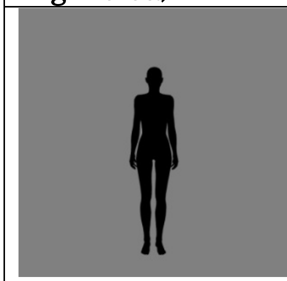
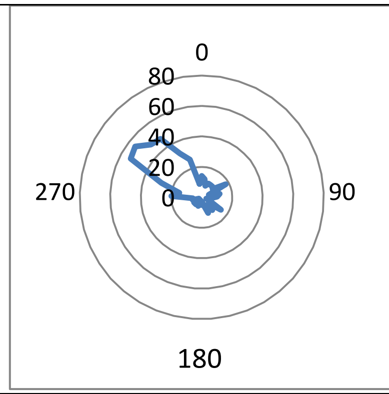
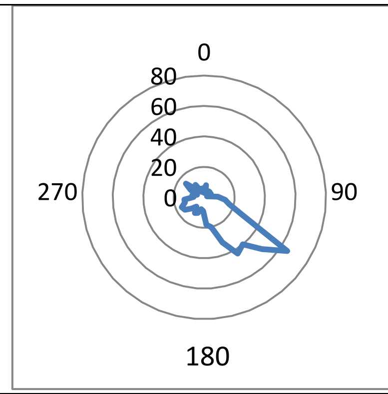
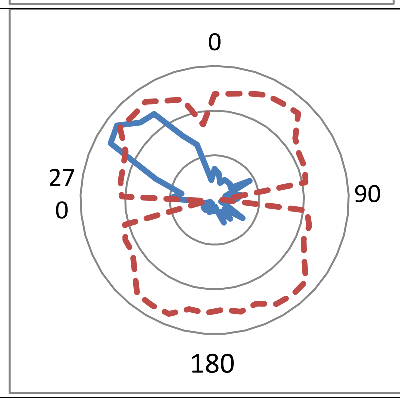
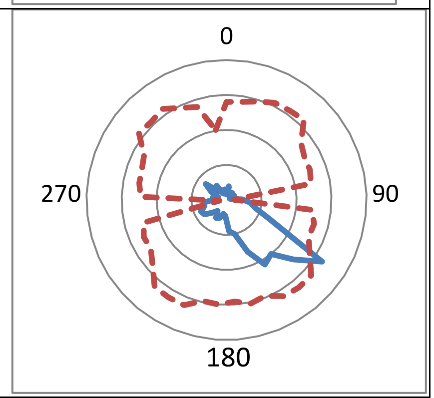
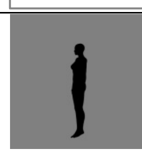
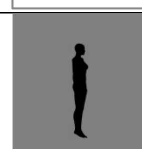
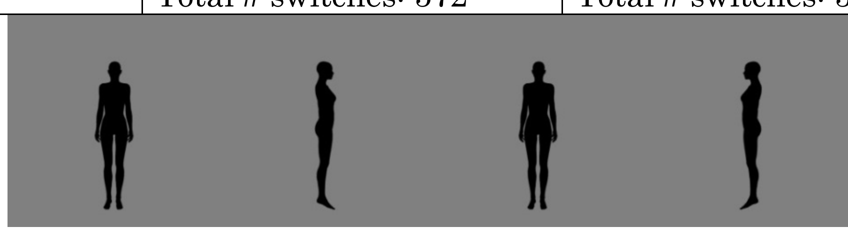
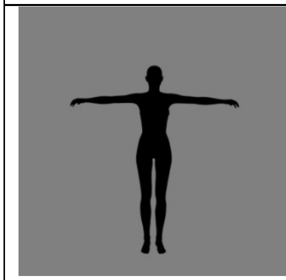
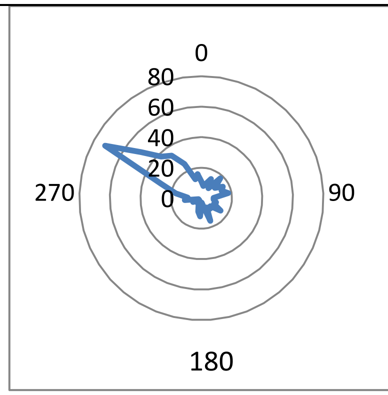
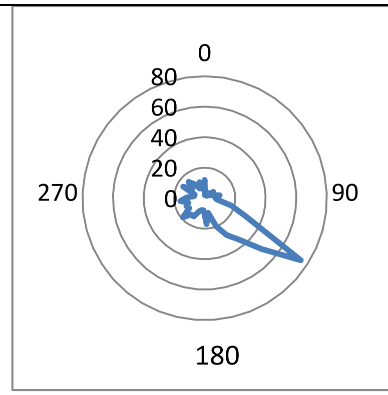
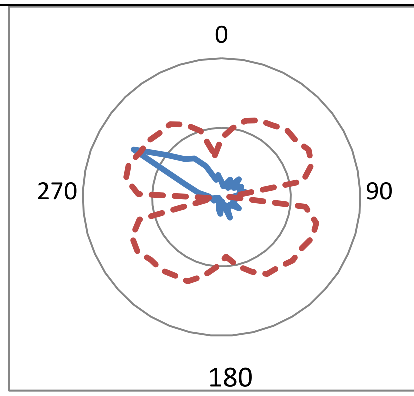
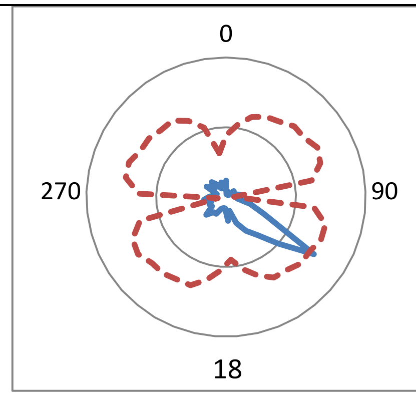
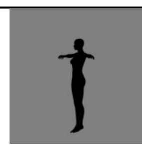
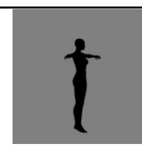
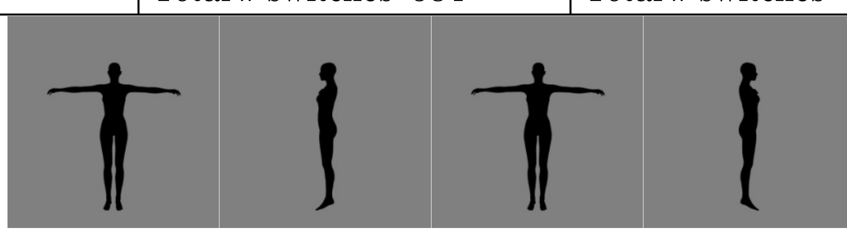
<table>
<tr><th>Figure 6a)</th><th>Anti-Clockwise Reversal</th><th>Clockwise Reversal</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td>Statistical Test: Rayleigh test for non-uniformity</td><td>
mode: 308
mean: 325
p-value = 2.88 e-052
Total # switches: 572</td><td>
mode: 123.2
mean: 153.1
p-value = 4.74 e-047
Total # switches: 571</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 6 a: KDE silhouette with arms at side

Figure 6b)	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode: 299.2 mean: 330.8 p-value =1.79 e-031 Total # switches: 584	 mode: 123.2 mean: 154.1 p-value = 8.18 e-028 Total # switches: 578



Degrees: 0 | 90 | 180 | 270

Figure 6 b: KDE silhouette with b) arms outstretched to side

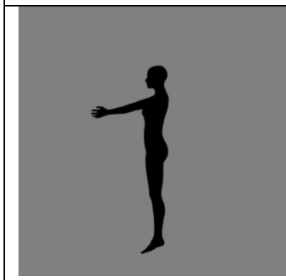
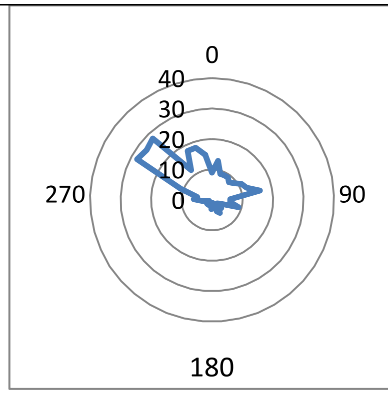
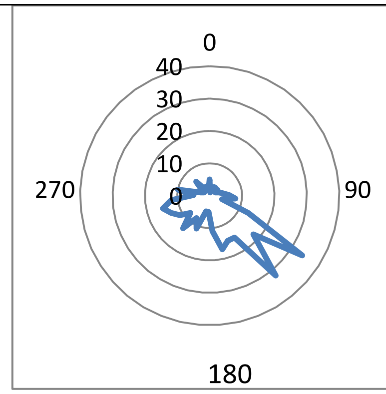
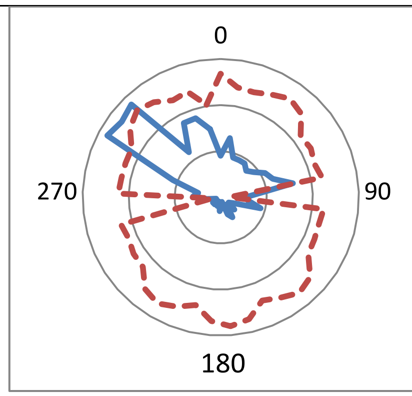
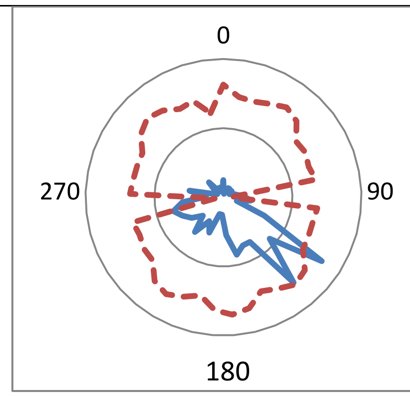
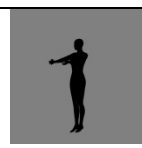
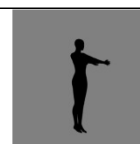
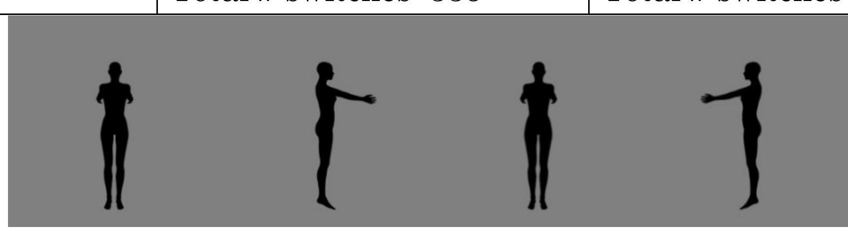
Figure 6c)	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode: 299.2 mean: 351.6 p-value = 1.01 e-032 Total # switches: 335	 mode: 123.2 mean: 170.0 p-value = 2.04 e-026 Total # switches: 341
 Degrees: 0 | 90 | 180 | 270		

Figure 6 c: KDE silhouette with arms forward in front of the body

The results of experiment 1.1 clearly show that not all positions along the 360 degree rotation of KDE images are equally likely to induce a directional switch. Also, once we account for a reaction time delay, see section E – Control Experiment, we can see that 1) the directional switches seemed to primarily occur when the silhouette had just passed through a profile position. 2) From the profile position subjects could either see the silhouette turn 'away' from them or turn 'to face' them. Note in figure 6(a,b,c) that the peak for "Anti-Clockwise Reversal" and "Clockwise Reversal" are at different locations along the 360 degree rotation. The "Anti-Clockwise Reversal" peaks when the silhouette is facing to the observer's left and the "Clockwise Reversal" peaks when the silhouette is facing to the observer's right. If a silhouette is facing rightward, and the subject indicates that the direction of rotation has changed so that the silhouette has started rotating left, the silhouette is now rotating to face the subject once again. Thus these peaks indicate that, not only were the perceptual switches centered around the profile positions, but the switches tended to occur so that the silhouettes would continue to face towards the subjects.

Thus directional switches occurred most frequently under the following conditions:

> Initially, before the rotation reversal was reported, the silhouette appeared to be facing towards the subject and to be rotating into a

profile position facing left or right. Once in the profile position the silhouette appeared to change direction and once again turned to face towards the subject.

This indicates a clear bias not only for position of the directional switch but also entails that the silhouette is preferentially seen as ‘facing’ the subject.

If the profile position alone was the important factor in inducing switches then each of the graphs would have two peaks indicating a direction change at each profile position instead of a single peak.

This result may be related to the work on point-light walkers done by Vanrie, Dekeyser & Verfaillie in 2004. Vanrie showed that when subjects were presented with an ambiguous figure of a point-light walker which could either face towards or away from them, there was a strong preference to perceive it as facing towards them, over 80% of the time, even when the point-light walkers were designed to appear to be walking backwards when the subjects perceived them as facing forward (Vanrie, Dekeyser, & Verfaillie, 2004).

The data show that there is a strong preference for seeing these stimuli face towards the subject. However visual inspection of the graphs seems to suggest that these distributions may have a second lobe of increased switches 180 degrees opposite to the main lobe. If this lobe exists it would indicate that while the majority of the time there is a preference for seeing directional switches which maintain a forward-facing stimulus, there are occasions where a directional switch will occur that maintains an looking-away from the subject.

To test for this possibility the data from figure 6a & 6b was transformed and retested (6c was not included because of the peaks crossed a much larger portion of the whole rotation and the data would have had to be drastically altered to conduct this test).

First a 75.4 degree arc of data, centered around the statistical mean of each stimulus was identified, this accounted for 9 frames of the stimulus and contained the frames with the highest number of directional switches indicated by the subjects. The average number of switches/frame was calculated using the remained of the dataset and this average value was used to replace the 75.4 degree arc which previously contained the peak values. This retained the position of all data within the set, but effectively removed the main peaks we had already identified. If there is a second peak 180 degrees opposite the main peak, then a Rayleigh test would once again prove to be the most useful statistical analysis.

The results of the altered and reanalyzed data did show a non-uniformity in the distribution, the weakest of which was $p < 3.7 \times 10^{-5}$. While this is strong it does not compare to the original results which were on the order of $p < 8.18 \times 10^{-28}$. Furthermore the peaks identified were nowhere near where they would be expected to appear if a secondary peak was 180 degrees opposite to the main peak. For figure 6a, the peaks were (Right: 308, Left: 123) and the identified peaks after the secondary analysis were (Right: 34, Left: 256), a difference from the expected peak of (Right: 86 off, Left: 133 off) Similarly for figure 6b, the peaks were (Right: 299, Left: 123) and the identified peaks after the secondary analysis were (Right: 53, Left: 239), a difference from the expected peak of (Right: 114 off, Left: 116 off).

It is fairly safe to say, that at least for the current data sets there is no secondary peak 180 degrees opposite to the main peak. The lack of such a secondary peak suggests that subjects were unlikely to see directional switches occurring at profile positions that would maintain a "facing-away" pose.

Subjects experienced a total of 4,146 revolutions of the stimuli over the course of the experiment, and if there really are preferred switching locations, each revolution provided 2 opportunities for a perceived directional switch to occur this results in (4,146x2) = 8,292 switch opportunities. Overall 2,981 reversals were experienced across all subjects, this comes out to be 2,981/8,292 = 35.95% of the total possible switches assuming a maximum rate of two direction reversals per revolution. There is a notable difference in the number of directional switches produced by 6c (676) as compared to 6a (1,143) & 6b (1,162). However the significance of that difference, if any is currently unclear.

Experiment 1.2

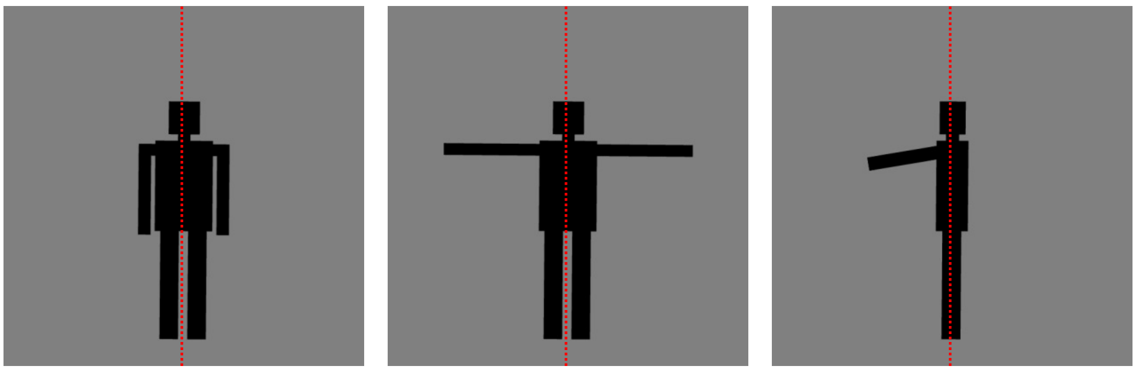


Figure 7: The three robot poses used for Experiment 1.2

Experiment 1 clearly showed that KDE images have preferred switch locations thus indicating that there are some visual processes which can bias where those switches will occur. Furthermore those perceptual switches imply at least one high-level bias: they were more likely to occur at locations which resulted in the 3D human figure being perceived as looking *towards* the subject.

In this experiment we wanted to look more closely at this possible high-level bias so we created 3D images of a humanoid-shaped robot made up of boxes which had both left-right symmetry and front-back symmetry for two out of the three images. Thus as compared to the previous images there were no facial features in profile, feet pointing forward, etc to define the direction the robot was "facing". The extra symmetry in these images means that as they rotate, the images from the first 180 degrees of rotation will be identical to those from the second 180 degrees. Thus we would expect that these two stimuli should produce a distribution of directional switches which repeats every 180 degrees.

Experiment 1.2 – Methods

We repeated the experimental method used above but substituted the rotating 'robot' with both front-back and left-right symmetry for the rotating human figure.

<u>Experiment 1.2 - Results & Conclusions</u>

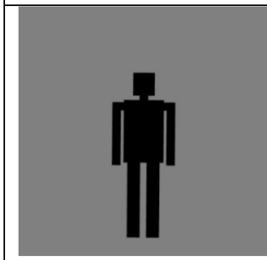
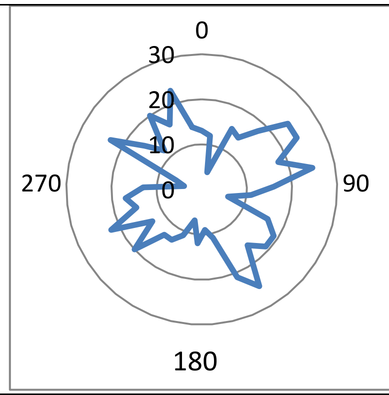
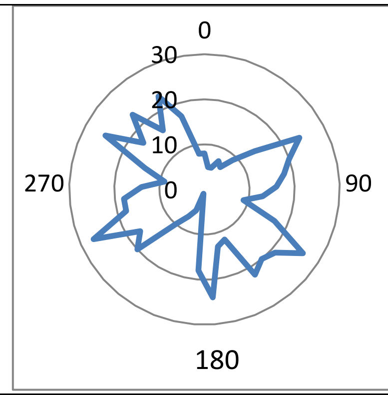
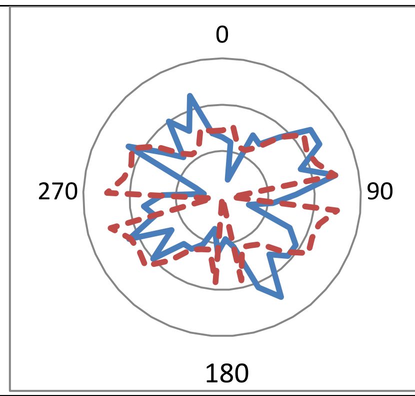
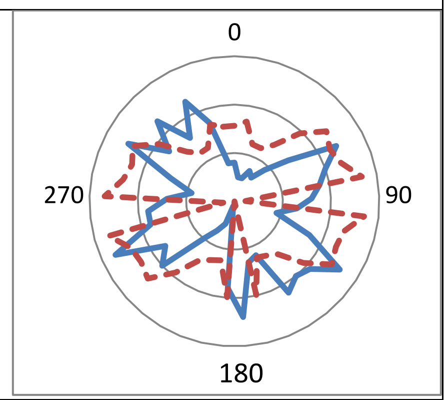
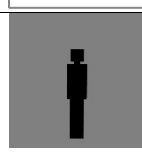
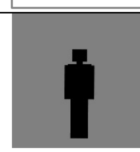
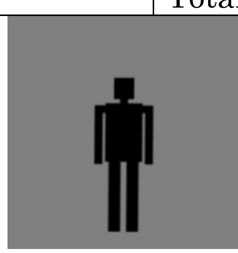
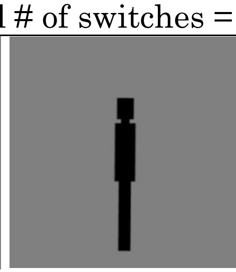
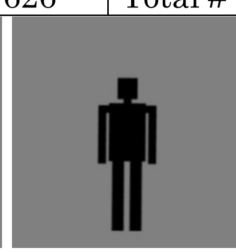
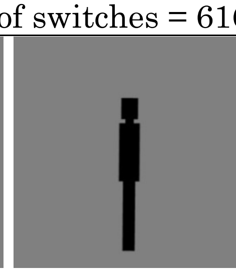
<table>
<tr><th>Figure 8a)</th><th>Anti-Clockwise Reversal</th><th>Clockwise Reversal</th></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td></td><td>4 Peaks~ 150, 246, 326, 52</td><td>4 Peaks~ 150, 246, 318, 62</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test: Omnibus Hodges-Ajne Test for non-uniformity</td><td>
52 & 246 Degrees
Omnibus Hodges-Ajne
mean = 68.27
p-value = 0.13
Total # of switches = 626</td><td>
150 & 318 Degrees
Omnibus Hodges-Ajne
mean = 196.08
p-value = 0.26
Total # of switches = 616</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 8 a: Experimental results of KDE silhouette of "Robot" with arms down

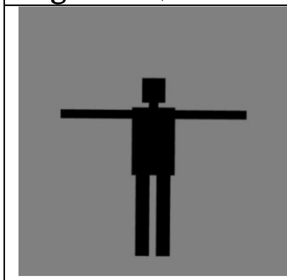
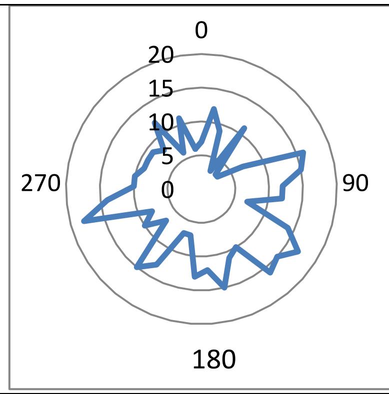
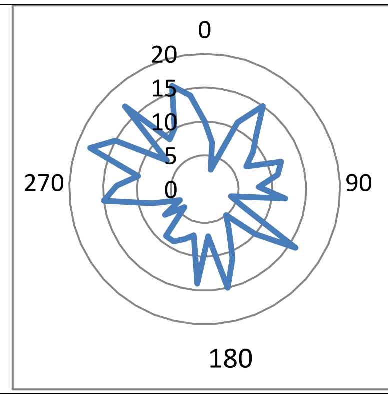
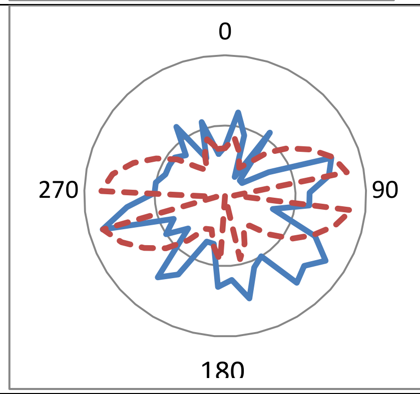
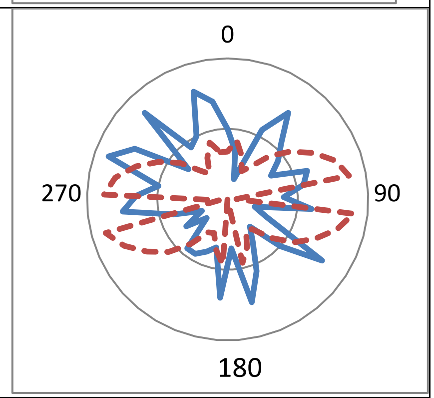
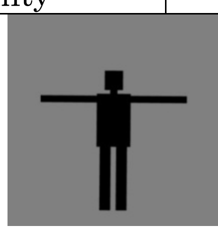
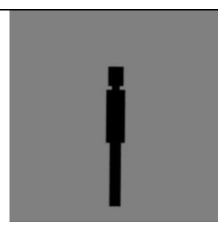
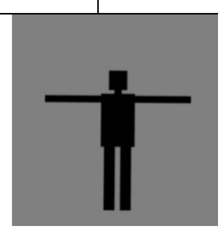
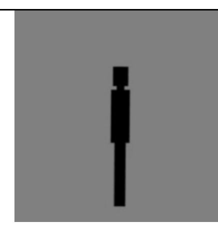
<table>
<tr><th>Figure 8b)</th><th>Anti-Clockwise Reversal</th><th>Clockwise Reversal</th></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test: Omnibus Hodges-Ajne Test for non-uniformity</td><td>Omnibus Hodges-Ajne
mean = 169.34
p-value = 0.01
Total # of switches = 429</td><td>Omnibus Hodges-Ajne
mean = 331.89
p-value = 0.11
Total # of switches = 411</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 8 b: Experimental results of KDE silhouette of "Robot" with arms outstretched to the side

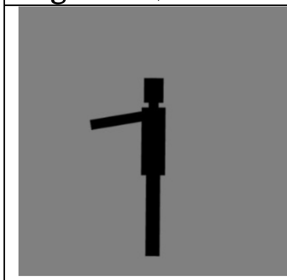
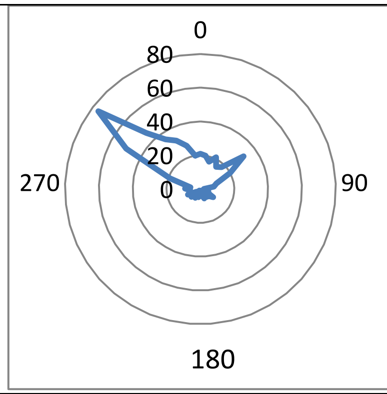
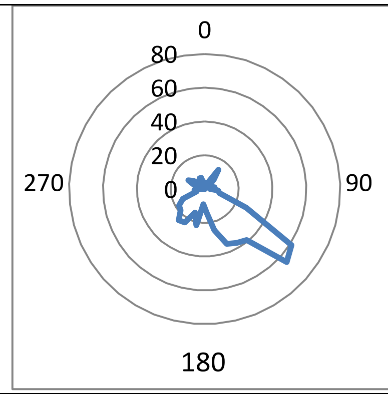
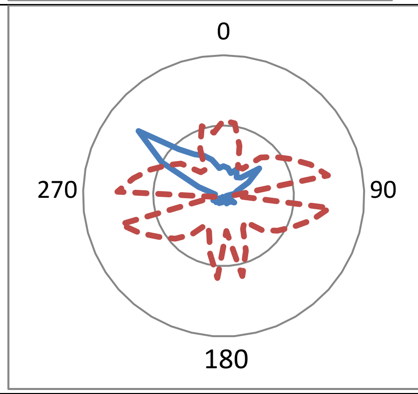
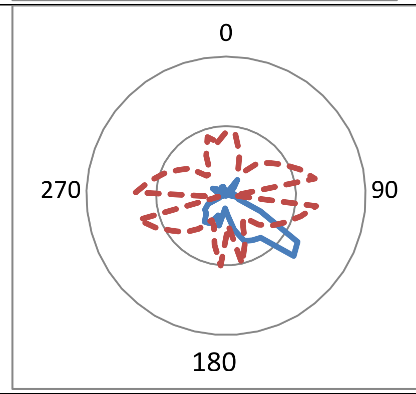
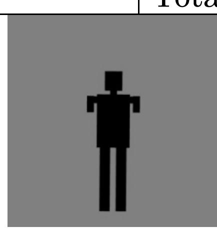
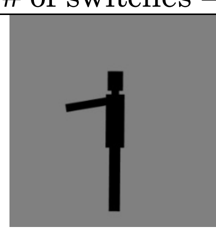
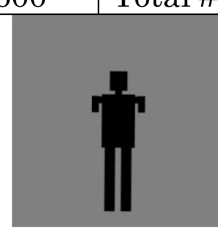
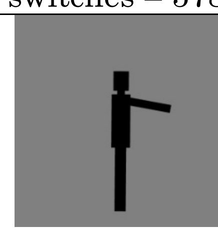
<table>
<tr><th>Figure 8c)</th><th>Anti-Clockwise Reversal</th><th>Clockwise Reversal</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td>Statistical Test: Rayleigh test for non-uniformity</td><td>
mode = 308.0
mean = 338.72
p-value = 1.36 e-078
Total # of switches = 600</td><td>
mode = 132.0
mean = 157.51
p-value = 1.96 e-069
Total # of switches = 578</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 8 c: Experimental results of KDE silhouette of "Robot" with arms outstretched in front of the body

8a) A visual inspection of results for figure 8a suggest that there are four positions within a full rotation where subjects are likely to experience directional switches. However a statistical analysis of the data from both the Clockwise and Anti-clockwise responses was unable to reach a level of significance to reject the idea that the data are not uniformly distributed (Right, $p = 0.13$), (Left, $p = 0.26$). Compare this result to the Anti-Clockwise Reversal data from 8b which was shown to differ significantly from a uniform distribution ($p < 0.01$) and it is clear that something odd is happening.

To understand this result it is imperative to consider how the Omnibus/Hodges-Ajne test non-uniformity of circular data works. The test was designed to be useful in cases where a Rayleigh Test was inappropriate due to a non-unimodal distribution or a non-normal distribution of the data. The test functions by making the assumption that if the data is non-uniform, then there will be some 180 degree portion of the data set which has a smaller number of data points relative to the whole set. It then tests every 180 degree hemisphere that can be created and looks for a case where the total number of data points on one side of the distribution is sufficiently small so as to indicate a significant non-uniformity, and in most cases this test can successfully find any non-uniformity that may exist. However in this situation the Hodges-Anje test is failing because of the symmetrical nature of the distribution of the data across any arbitrary hemisphere of the rotation.

Since the data is symmetric there is no hemisphere which has a sufficiently low number of data points to meet the requirements of Hodges-Anje for non-uniformity.

This case is especially odd because while Hodges-Anje cannot find a non-uniformity for 8a (Left or Right), it does find a non-uniformity for 8b (Right). A visual comparison of the two graphs seems to suggest that 8a has a four-lobed pattern which is repeated almost identically for both the Clockwise and Anti-Clockwise responses. 8b on the other hand seems to contain more chaotic data and there is no clear similarity between 8b Left and Right.

This leads to the conclusion that there is quite possibly a strong non-uniformity in the data from 8a, but that Hodges-Anje has failed to find it. As a way to look for this missing significance an ad-hoc statistical analysis was devised. First it was assumed that this was a four-lobed distribution with the lobes peaks occurring at 90 degree intervals. The data was then transformed to wrap back on itself every 90 degrees. Thus the data points from 5, 95, 185, & 275 were all summed and binned together. The starting point for this transformation was randomly generated so as not to introduce any bias, and if there is in fact a regular four lobed pattern to the data it will not matter what the starting point is. Finally the data which now consists of only 90 degrees of rotation was reallocated to make a new 360 degree data set

(22.5-->90, 45-->180, 67.5-->270, 90-->360). This new data set was examined with standard Rayleigh test for uniformity, yet even this modified test did not produce a statistically significant result. However it can be noted that data from figure 8a (robot with arms relaxed at its side) and 8b (robot with arms held out) failed with nominally different p-values. The distributions for 8a, which we expected to contain four lobes, had p-values of $p = 0.12$ (Right) and $p = 0.21$ (Left). While the distributions for 8b much higher p-values of $p=0.47$ (Right) and $p=0.43$ (Left).

Despite these failures of statistical significance, what follows is an analysis of the data based on a visual inspection that indicates a four-lobed distribution of the data from 8a. This analysis should not be too strongly considered, given the lack of a strong statistical test for significance.

Similar to experiment 1 there are some positions in the 360 degree rotation where direction reversals are more likely to occur. In the case of the humanoid figures from the previous experiment, we proposed that the preference for reversals in the profile view was associated with a perceptual preference for a ‘facing’ orientation of the figure. On that view it may seem surprising that profile rotation reversals were as highly prevalent as front-back reversals given the symmetrical nature of the robot from Figure 8a. This suggests that subjects may not be assigning and arbitrary front/back to the robot, and that perceptual reversals could be modulated by lower level features of the stimuli such as linear image velocity or depth symmetry.

The reversals near the two 'profile' positions are similar to what was seen in the humanoid figure from experiment 1.1. The switches near the forward/backward positions were not seen in experiment 1.1, however at a low-level these are four positions in the rotation where the linear velocity of the image contours reached a point of 'minimum-motion'.

There are three points in the rotation where the linear velocity of the image should be noted:

1) Linear velocity was at a global maximum just prior to when the object reached a profile position. $\text{Velocity}_{p1} = 0.4006$ mm/deg.
2) Linear velocity was at a local minimum as the object entered a full frontal/backwards view. $\text{Velocity}_{f} = 0.0673$ mm/deg.
3) Linear velocity was at a global minimum for two frames of the image in the profile as the "fat" body of the object occluded the "arms". $\text{Velocity}_{p2} = 0.0168$ mm/deg.

It also should be noted that all four switch locations occurred at points where the 3D image was symmetric in both the left-right and front-back planes. This combination of minimum motion and two-way symmetry makes it impossible to determine which image properties were most relevant to the perception of directional switching.

Finally the ‘robot’ has no defined front/back and it should be noted that this lack of front/back information inevitably results in the “Clockwise Reversal” and “Anti-Clockwise Reversal” having similar frequency at the four peaks of switch location along the 360 degree rotation.

8b) The results from figure 8b did not provide any illumination regarding how subjects perceived perceptual switches. During the debriefing segment of the experiment almost every subject reported that this stimulus was “weird”. Specific reports of exactly what subjects experienced generally fell into two categories.

1) A 2-dimenionsal robot-like figure which was extending its hands outwards and inwards.

2) The image would appear to be a 2-dimenionsal blob that was morphing over time but which bore no resemblance to a humanoid figure.

3) A 3-dimensional robot figure that was rotating, but which would regularly turn into one of the other two percepts. Those subjects who did experience a rotating silhouette of a robot were very clear that this percept could not be maintained.

These reports are reminiscent of Ullman's discussion of rigid motion (Ullman, 1984-A; Ullman, 1984-B), where he suggests that deformation rather than rigid motion is quite common in cases where no 3D structure is perceived.

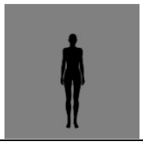
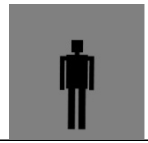
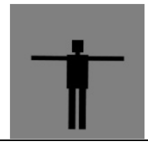
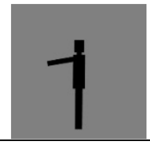
The subject reports make sense when considering how the image of this 3-Dimensional object changed from frame to frame. Consider the two extreme frames of this stimulus. In one case the object is in a full frontal view "facing" towards or away from the viewer and there is a small vertical gap of negative space between the legs. In the other extreme the object is in a profile view and looks like a vertical black line of varying thickness. In the transition between these two extremes, going from the profile to frontal view, there is very little visual information provided to the subject so as to determine how the object is changing. Instead, the subject merely sees a horizontal lengthening of the image, with some areas lengthening more than others. The first real bit of new information comes when the gap appears between the legs and at this point the object is almost in a full frontal view. Then as quickly as that gap appears, it closes, and the object now appears to shrink back to a vertical line. The paucity of 3-dimensional information here results in a series of images which could just as easily be a 2-dimensional transformation over time as a robot that is rotating in 3-dimensions. The only portion of the image sequence which does not correlate perfectly with horizontal dilation over time is the gap between the legs that can be seen in the frontal view. However this gap is only visible in 14 of 41 frames and by itself does not provide enough information to guarantee seeing a 3D object. This is a similar result to that found by Wallach and O'Connell (1953) when they presented subjects with a horizontal wire as a silhouette. Rather than experiencing a 3D percept Wallach's subjects indicated that they saw a 2D object which was growing and shrinking in length. One possibility for why the robot from figure 8a did not also exhibit this trait is that for figure 8a there was additional visual information contained within the gaps between the arms and body which did not exist in figure 8b. This small bit of additional detail may have been sufficient for the visual system to be able to experience the image as 3-dimensional.

8c) The most interesting result from this experiment comes from figure 8c when the robot has its arms "forward", because we immediately see the same pattern of rotational switching as with the humanoid figure. This result may be due to the fact that in typical humanoid figures when the arms are positioned as they are here it indicates a biologically plausible direction in which the figure is facing. This could provide very strong evidence that the "facing-bias" discussed earlier is not only present for KDE images, but does not require a fully contoured body, merely the a few visual cues to indicate a forward-facing direction. If so this would correlate well with the point-light walkers which present a “facing-bias” that were discussed earlier (Vanrie, Dekeyser, & Verfaillie, 2004).

The main difference between the robot and the humanoid is that the humanoid figure had clear cues to indicate which direction it was facing when in the profile position. Those cues do not exist at all on the robots from Figure 8a & 8b. This means that for those stimuli the “left profile” position is identical to the “right profile” position. So regardless of the perceived direction of rotation there is no cue to bias the robot to “turn towards” the subject like the humanoid did. This also provides a clue as to why the 3rd robot image showed the same pattern as the humanoid, and such a different pattern from the other robots in 8a and 8b. In this case the arms were positioned as if they were held up in front of the robot. This provided a cue which implied which side would be “forward” in the profile positions if it was going to be a biologically plausible figure. This minor change seems to have had a major effect and thus subjects experienced the majority of directional switches so that the robot would continue to ‘face’ towards them.

Additional Analysis – Total # of Switches

Table 1: Total number of directional switches, Left and Right for each stimulus

	6a 	6b 	6c 	8a 	8b 	8c 
Anti-Clockwise	572	584	355	626	429	600
Clockwise	571	578	341	616	411	578
Total	**1143**	**1162**	**696**	**1242**	**840**	**1178**

It is worth comparing the total number of directional switches across all the humanoid and robot stimuli to look for any major difference in the overall rate of directional switching which may indicate some underlying difference between the stimuli (Table 1).

The first thing to note is the large difference in directional switches between figure 6c which has roughly 40% fewer directional switches than either 6a or

6b. The second major difference is a comparison of 8b which has roughly 30% fewer directional switches than either 8a or 8c.

The difference in total switches between the various robot figures could possibly be explained by the simple fact that figure 8b was a stimulus that produced a fundamentally different percept than either 8a or 8c. Since most of the subjects did not strongly experience a rotating 3-Dimensional percept for 8b a comparison of the total number of switches to 8a and 8c is unlikely to have any significant meaning.

On the other hand the difference in total switches between the humanoid figures cannot be dismissed so lightly. In all three cases the data showed a strong preference for directional switches in the profile position and a forward facing bias. So how is stimulus 6c so different from the other two that subjects saw directional switches 40% less often? The only obvious major difference is that stimulus 6c has a larger front-back asymmetry than the other two stimuli. This means that if a direction reversal were to occur when the image was facing directly towards or away from an observer, there would need to be a much larger total change in the perceived depth of the image as compared to the other stimuli. This would be due to the fact that as the direction switch occurred, there would be an accompanying depth reversal of the image, which does NOT happen in the profile positions. The extended arms of the figure would need to translate much farther in depth for this type of reversal than the arms in either of the other two images. Both 6a and 6b would still undergo a depth reversal from front to back, but only 6c would have a depth reversal that required greater overall magnitude of depth change as the percept switched.

Additional Analysis – Linear Image Velocity & 3D Object Information

There are three basic object/image level considerations that need to be considered as possibly affecting how directional reversals of motion occur.

1) Linear Image Velocity

As each of the 3D stimuli rotate, the image presented to the subject changes over time. The low-level properties of these images are directly measurable and may be contributing factors to determining when higher-level direction reversal will occur. The linear image velocity is a measure of how quickly the contour edges of the silhouette move across the screen. Linear image velocity need not directly relate to the angular velocity of the 3D object; however changes in linear velocity are a major input component into the visual mechanisms which ultimately create a 3D construct from the image. Additionally, the interpretation of linear image velocity from frame to frame is a problem with many possible solutions (Hildreth & Ullman,1982). Thus linear image motion may contribute to perceptual reversals of direction.

It should also be noted that linear image velocity will be most strongly determined by the portions of the 3D object which are furthest away from the axis of rotation, in these cases the positioning of the arms. For stimuli 6a, 6b, 8a & 8b the front/back pose is where linear image velocity will be at its slowest, and the profile position is where linear image velocity will peak. Conversely for stimuli 6c & 8c the minimum image velocity will be at a minimum in the profile positions and the image velocity will peak in the front/back pose. These differences are purely related to the geometry of the stimuli. Stimuli 6a, 6b, 8a, & 8b have their most extreme points, from the axis of rotation, to the left/right of the body, while stimuli 6c & 8c have their most extreme points "in front" of the body.

A final note about linear image motion of 3D silhouettes. The linear image velocity will typically follow a sinusoidal pattern as the object spins, however in some cases the extreme points, furthest from the axis of rotation, which are driving linear image velocity may be temporarily occluded resulting in a drastic change of linear image motion. A normally lit 3D object does not have this problem, but the lack of shading in a silhouette means that an image contour can be occluded by a portion of the object that is "behind" it.

2) 3D Object Motion

The actual direction of rotation experienced by the subjects is directly related to how they perceive the 3D object to be moving. Every time a subject experiences a perceptual reversal of motion it will either cause, or be caused by, a reinterpretation and reversal of the angular velocity of the 3D object. Thus any factors which influence how the 3D object motion is interpreted may influence when reversal of direction of experienced (Marr, 1982).

3) <u>3D Depth Reflections/Mirror Reversal During Rotation Reversal</u>

Most of the discussion thus far has focused on how subjects will experience a reversal of the direction of motion of the KDE stimuli. However it is often the case that when a motion reversal occurs there is a concurrent depth reinterpretation of the 3D object which results in a mirror depth reversal (Figure 9 & Figure 10). This depth reversal occurs due to the lack of shading information in a silhouette. However it will not result in a new interpretation of the 3D object if the reversal occurs when the image of the object has front-back symmetry at the moment of the rotation reversal. This may be a significant factor in determining when direction reversals will occur because any image frame within the rotation of the 3D object which is front-back symmetric will not incur mental cost to reinterpret the position or shape of the object.

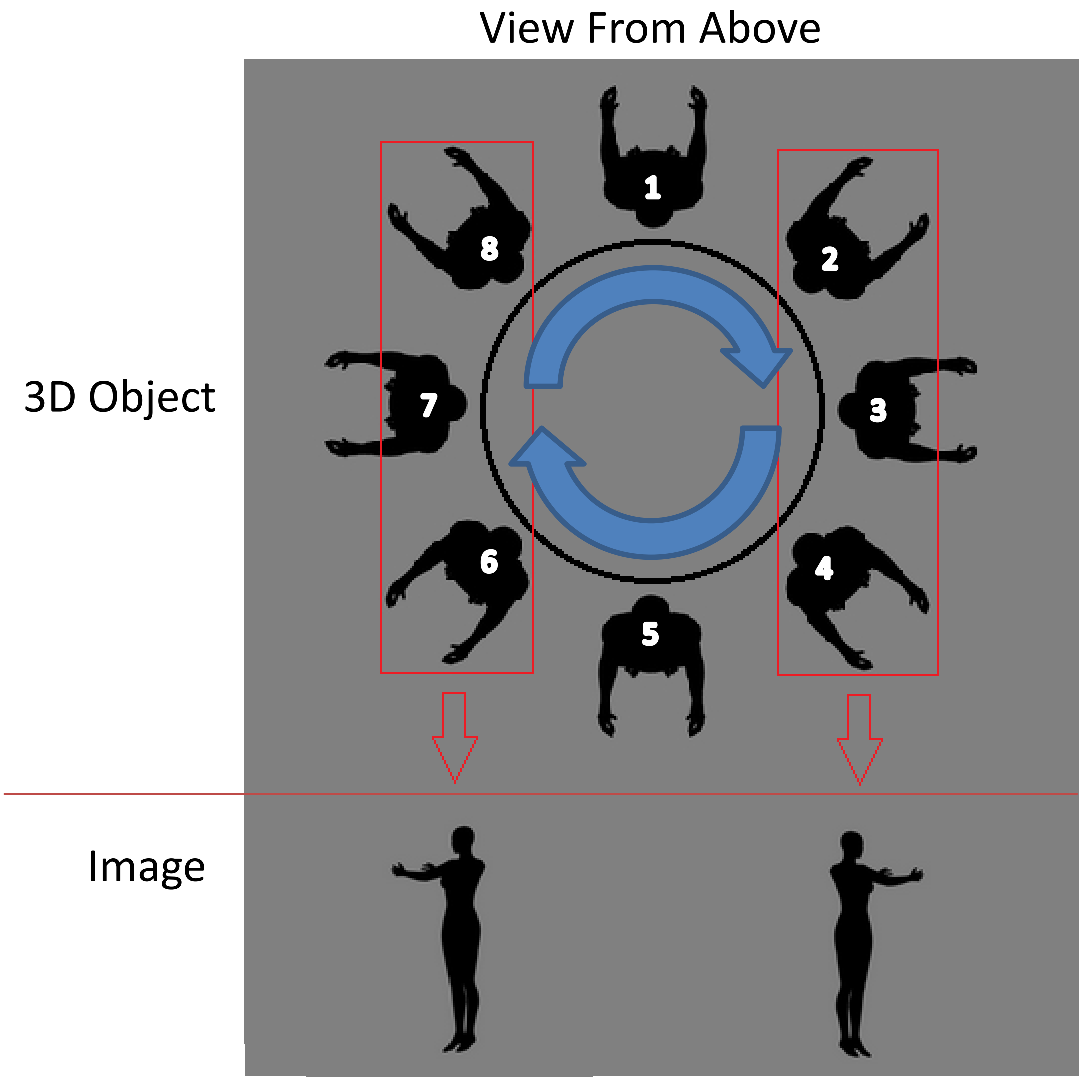


Figure 9: An example of depth reinterpretation. As an object rotates there will be various positions where the image produced by the object can be interpreted in two ways. Here you can see that positions 2 & 4 produce the same image as do positions 6 & 8. Rotation reversals at these positions require a depth reinterpretation. But no such reinterpretation is required at positions 3 & 7.

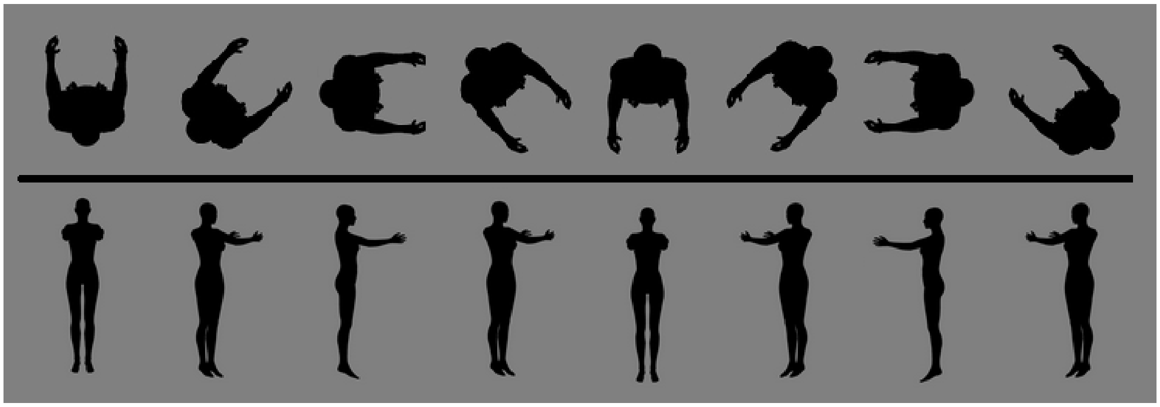


Figure 10: A top and side view of a KDE stimulus as it rotates. Notice that different top views will produce the same side view.

A case where linear image velocity may be affecting where rotation reversals arises from the four peaks of figure 8a. As the robot KDE image spins, the linear velocity of the edges of the image moving across the screen changes as a function of velocity=sin(Θ), where Θ is the angle of the orientation and Θ = 0 is the starting position of facing forward/backwards. However the sinusoidal function breaks down in cases where one portion of the object occludes the other. At the forward/backward positions the image edge velocity reaches zero and then reverses direction. This zero-velocity point is where a slow image edge velocity can imply a slowing of the rotation of the 3D object and thus there may be higher probability that a perceptual switch will occur because there would be a smaller reversal in angular object velocity.

Interestingly at the profile positions the linear edge velocity is at a maximum, followed by a brief minimum velocity due to occlusion, and then immediately back to maximum. This indicates that not only may there be a high level bias enabling switches at that location, but that it would be strong enough to also overcome a low-level bias for continuous motion.

When the image contour velocity approaches zero in the succession of views of a rotation at constant 3D velocity, this is also roughly consistent with a slowing of the 3D rotation to a standstill. In the latter case, a slowing to an instantaneous rotation standstill could form a smooth transition into the opposite direction of rotation, as would occur if angular acceleration were constant .In the former case, and in the actual experimental image sequence, the angular velocity is constant and there is no implied rotation reversal .Thus the two interpretations both have some physical plausibility, as well as implying a momentary standstill of the image contour. *A priori* considerations together with the image data are therefore relatively supportive of a perceptual rotation reversal at these points, and this may be the reason for the frequent reversals reported in the frontal poses of the robot in Figure 8a.

Notice that the two 3D perceptual interpretations considered here are still in principle distinguishable on the basis of the higher derivatives of image contour motion: with rotation at uniform velocity the image contour velocity is approximately linear through its zero crossing, whereas if the 3D rotation itself decelerates to perform a 'California stop' at the point of image standstill, the image contour velocity is approximately proportional to the square of the time difference from the moment of standstill. But these differences are likely to be too subtle for subjects to detect with high sensitivity (indeed the visual system is generally rather insensitive to acceleration in general (Nakayama, 1982).

Interestingly at the profile positions of images 8a & 8b the linear edge velocity is at a maximum, as shown previously. This indicates that not only may there be a high level bias enabling switches at that location, but that it would be strong enough to also overcome a low-level bias for continuous motion. However a confounding point is self-occlusion of the thin arms of robot by the thick body at the profile position. Which creates a brief, but large, drop in linear image motion just as the robot reaches the profile position. This confound is not present in stimulus 8c, as the thin protruding arms of the robot are not occluded by the thick body. However in 8c the minimum and maximum points of image motion are different due to the extension of the arms forward. In 8c the point of minimum image motion coincides with the profile position and the point of maximum image motion coincides with a front/rear view (also there is no loss of image motion due to occlusion as there was in 8a & 8b).

When rotation reversals occur in the profile view, the image contour velocity, is large but nearly constant, and is interpreted as a rotation velocity that is constant but with an abrupt sign reversal. This is made possible by reassigning each silhouette contour element to a depth-reflected part of the 3D object. Such perceptual reassignments of image features to object parts are often necessary during observation of rotating objects under natural conditions (Marr, 1982) so there is nothing artificial about this depth 'reflection' which affects only the correspondence between 3D structure and image features, and not the perceived 3D shape itself. What is arguably unnatural about the instantaneous perceptual rotation reversal in the profile view is that it implies a large and physically implausible discontinuity in angular momentum, as is always the case when rotation reversals associated with non-zero 2D image contour velocity.

The perceptual implausibility of this large change in angular momentum may be being partially mitigated by two factors. 1) As stimuli 6a, 6b, 8a, & 8b rotate into the profile position their linear image velocity increases right up until the point where the body of the image occludes details of the arms. This results in a very large drop in linear image motion just as the object enters into the profile position, resulting in a reduction of what would otherwise be a large change in angular momentum. 2) In the cases of stimuli 6c & 8c, the frontally extended arms alter the profile of the linear image velocity such that the profile positions coincide with minimum linear image motion. Thus, while there is still a drastic change in the overall angular momentum of the 3D object, the slowing down of the linear image motion may act as a cue towards reinterpretation of the 3D object motion which then facilitates a perceived change in rotational direction.

For the robot of Figure 8a, the opposed lobes corresponding to each profile view and the front/back views are roughly balanced in frequency between 'left' and right' responses. This is inevitable given the front/back symmetry of the depicted robots, which makes the perceptual assignment of the frontal view quite arbitrary, so that it is impossible to know whether a 'left' or a 'right' response is associated with perceptual rotation towards or away from the observer.

<u>Experiment 1.3</u>

Experiment 1.1 showed that direction reversals of KDE images were not random and that high level biases, such as a facing-towards bias may affect where in the rotation directional switches occur. Experiment 1.2 further supported the existence of a high level facing-towards bias. It also suggested that low-level factors related to image motion and 3D object interpretation may be factors which can inhibit or induce perceptual reversals of direction. For experiment 1.3 we decided to use simple very images which still produced a kinetic depth effect to explore the low level biases which might affect perceptual direction reversals.

The images we used were simple triangles that were tilted back slightly to provide enough visual cues for a kinetic depth effect to occur, upright triangles simply look like an object that is changing shape rather than spinning (Wallach & O'Connell, 1953).

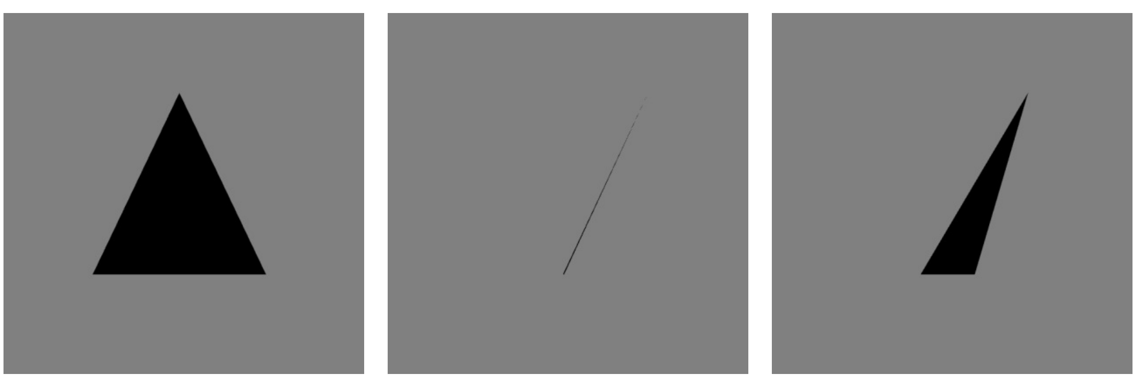


Figure 11: Single Triangle (a single 2D triangle which was tilted back and with an axis of rotation at the center of the base)

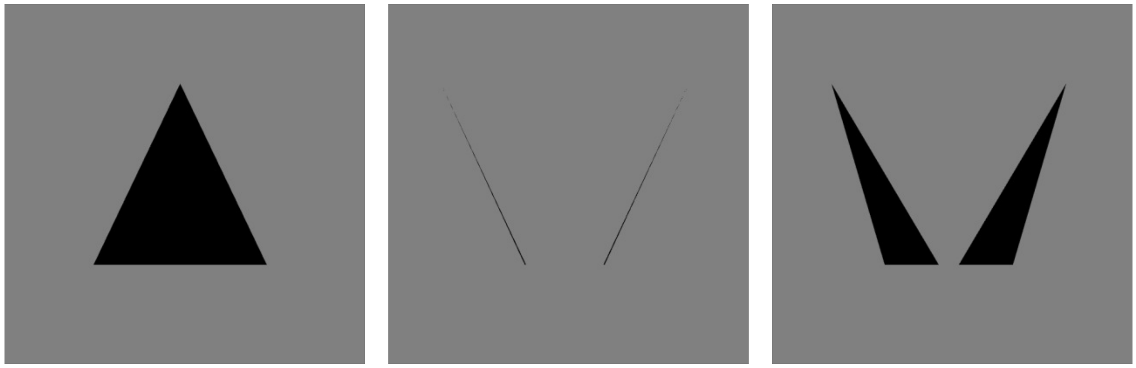
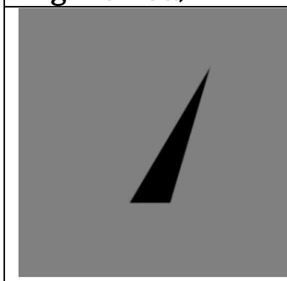
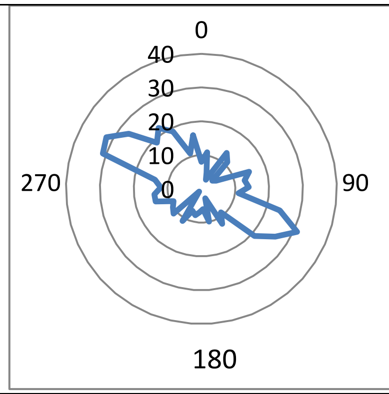
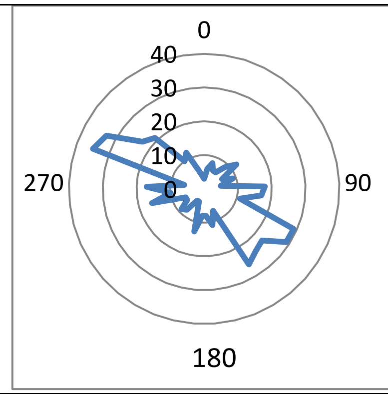
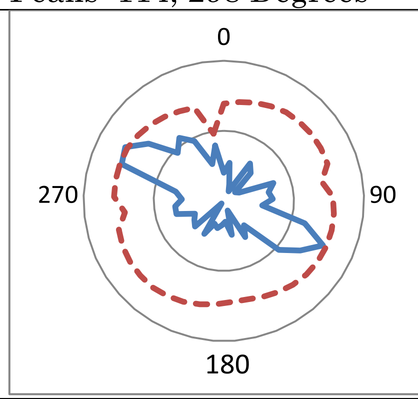
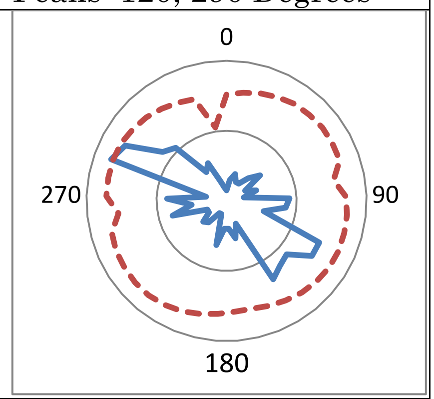
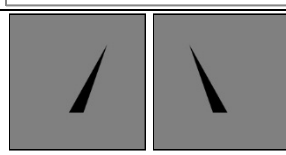
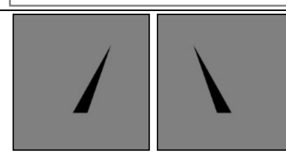


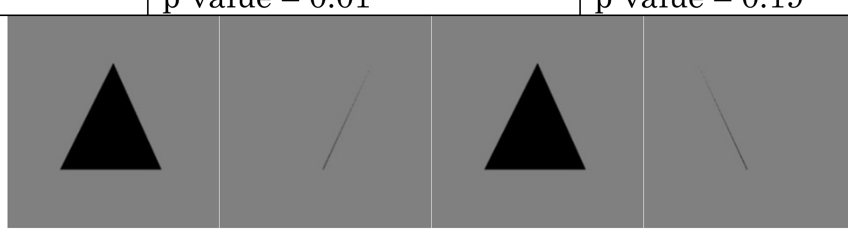
Figure 12: Two Triangles (same as the one triangle experiment but with a second triangle directly opposite the first and each shifted 1 degree of visual angle off the axis of rotation)

Experiment 1.3 – Methods

We repeated the experimental method used above but used simple geometric shapes instead of complex figures.

Experiment 1.3 - Results & Conclusions

Figure 13a)	Anti-Clockwise Reversal	Clockwise Reversal
		
	Peaks: 114, 298 Degrees	Peaks: 120, 290 Degrees
Frame-by-Frame analysis of linear image motion		
Statistical Test: Omnibus Hodges-Ajne Test for non-uniformity	 mean = 330.37 p-value = 0.01	 mean = 168.69 p-value = 0.19



Degrees: 0 \| 90 \| 180 \| 270

Figure 13 a: Experimental results of a single KDE triangle

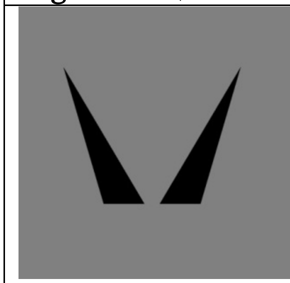
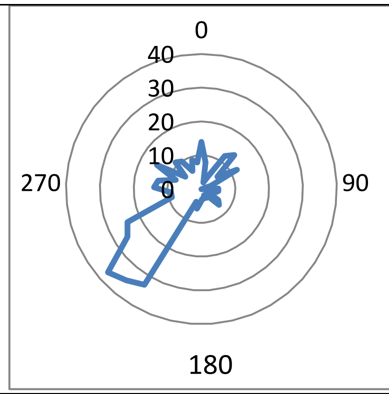
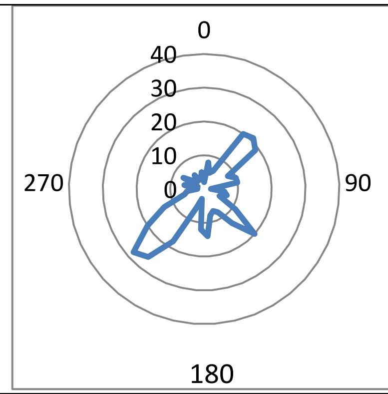
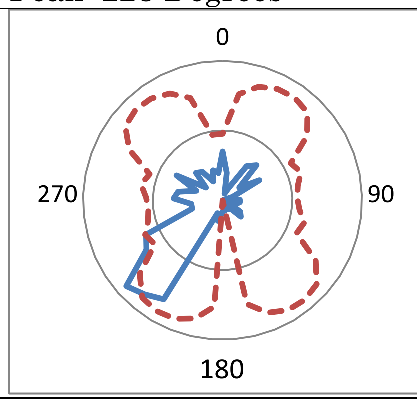
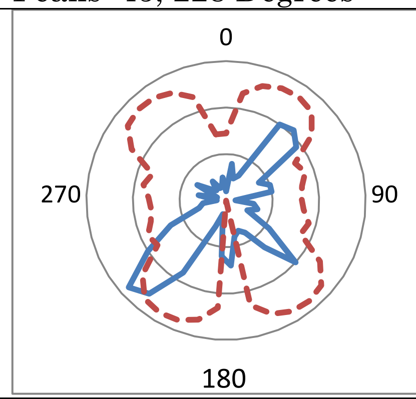
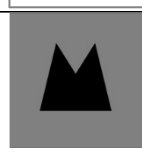
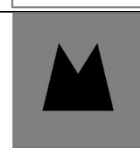
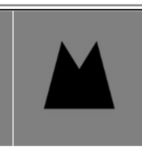
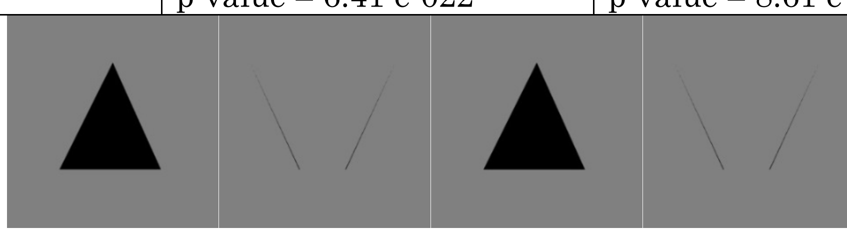
<table>
<tr><td>Figure 13b)</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td></td><td>Peak: 228 Degrees</td><td>Peaks: 45, 228 Degrees</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td>Statistical Test: Omnibus Hodges-Ajne Test for non-uniformity</td><td>
mean = 257.72
p-value = 6.41 e-022</td><td>
mean = 165.64
p-value = 8.61 e-006</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 13 b: Experimental results of two KDE triangles

The single triangle behaved as expected with two clear peaks for direction reversals at the positions where it was a 'knife-edge' and the image edge velocity was zero. This adds some support to the notion that perceptual rotation reversals preferentially occur when motion in the image is minimal. However, as in experiment 1.2 these reversals coincided with positions where the 3D object had front-back symmetry and thus no additional reinterpretation cost from depth reversal. The coincidence of minimal linear motion and zero-cost depth reversal, once again makes it impossible to disambiguate which of these factors might be more influential in modulating a perceptual reversal of direction.

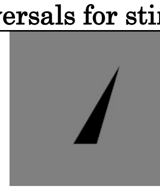
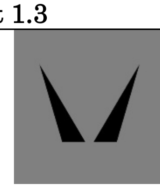
The double triangles on the other hand behaved in exactly the opposite manner, with the majority of direction reversals coming as linear velocity of the triangle images peaked. Why would the double triangle behave so differently? A key consideration is that at the moment in question, the silhouettes of the two triangles come precisely into register, the 'nearer' one obscuring the farther (so it looked like a single triangle).

With one object obscured by the other there is ambiguity as to which object generates a particular image contour. This ambiguity allows for a rotation reversal through a reassignment of each image contour to the opposite triangle at the moment of the reversal. In this way the image sequence can be perceptually interpreted in terms of an unchanging (but rigidly rotating) 3D configuration of two triangles. This is not possible at other moments during the rotation cycle (except for the where the image is a knife edge), because when the images of the triangles are visibly out of register, a perceptual rotation reversal requires that the 'nearer' and 'farther' triangles, now separately identifiable in the image, must change places in the perceived 3D configuration that is undergoing rotation. Just as with the robots of Figs 8a and 8b, this reassignment of image features to different 3D sources (that are momentarily in register in the image) occurs quite generally in natural situations during the perceptual solution of the inverse problem of constructing 3D interpretations from moving images, and the only unnatural aspect of the interpretation is the implied discontinuity in angular momentum.

One problem is that the hypothesis above does not explain why the two-triangle image would produce *fewer* directional switches at the minimum linear motion points when the triangles are at a knife edge. This is particularly worrying because this image frame has front-back symmetry just like the frame where the triangles obscure one another. If there is a benefit to directional switches occurring in image frames with front-back symmetry this benefit should be equally conferred to triangle as they are in a knife edge orientation.

One remaining possibility lies in the post-experiment debriefing where subjects reported that this stimulus would often appear to look more like a 2-Dimensional morphing object rather than a 3-Dimensional representation of two triangles rotating about a common axis. This stimulus had the second-most number of subject reports (after figure 8b) as appearing to be a morphing 2-Dimensional image rather than a 3-Dimesional object. This may also account for the vastly different distributions of reversals of direction experienced when comparing Clockwise vs Anti-clockwise responses. In fact the differences in the shapes of these distributions make very little sense as the 'Anti-Clockwise Reversal' distribution is almost definitely unimodal and the 'Clockwise Reversal' distribution appears to be tri-modal and yet there should be no difference between 'Clockwise Reversal' and 'Anti-Clockwise Reversal' since the stimulus has both Left-Right and Front-Back symmetry.

Table 2: Total # of direction reversals for stimuli in experiment 1.3

Total number of direction switches		
	1074	791

Finally, an examination of the relative number of directional switches between the two stimuli shows that the 2-triangle stimulus has about 25% fewer perceived directional switches as compared to the single-triangle stimulus. Is this reduced subject response due to an inherent stimulus property that is resulting in fewer direction reversals or, as in the case of figure 8b, could this just be due to the fact that subjects did not see as much 3-dimensionality in this image as they did for the single-triangle stimulus?

In the end it is very difficult to pull any useful information from these results and there is a high likelihood that the erratic nature of the data is due to an inherent weakness in the 3-dimensionality of the stimulus. This would account for subjects' reported confusion over this image and the odd data distributions.

Experiment Set 2 – KDE with conflicting stimuli

The first set of experiments clearly demonstrated that directional switching of KDE stimuli is not random, and in fact may be affected by numerous low-level and high-level biases. Unfortunately, the left-right symmetry of the stimuli from experiment set 1 often resulted in situations where the object and image parameters of biases that might be affecting directional switches came to a maximum/minimum at the same image frames within an object's revolution. This made it impossible to disambiguate which, if any, of these biases were responsible for the directional switches. Additionally, this overlap made it impossible to determine the relative strength of each bias or make comparisons between them.

In this second set of experiments the goal is to find ways to separate out the overlapping biases from experiment set 1 by using KDE stimuli which have been designed to 1) have bias parameters that do not have coincident maxima/minima at the same frame of rotation; 2) differ in magnitude for a single bias parameter which may alter perceptual direction reversals; or 3) have multiple bias parameters which could compete against each other so that a determination can be made as to which bias will more strongly affect perceptual reversals of object rotation.

<u>Experiment 2.1</u>

The first experiment contains four stimuli which are similar to the stimuli from Experiment Set 1, however the left-right symmetry has been removed. So we now have:

- Human one arm forward
- Human one arm outstretched sideways
- Robot one arm forward
- Robot one arm outstretched sideways

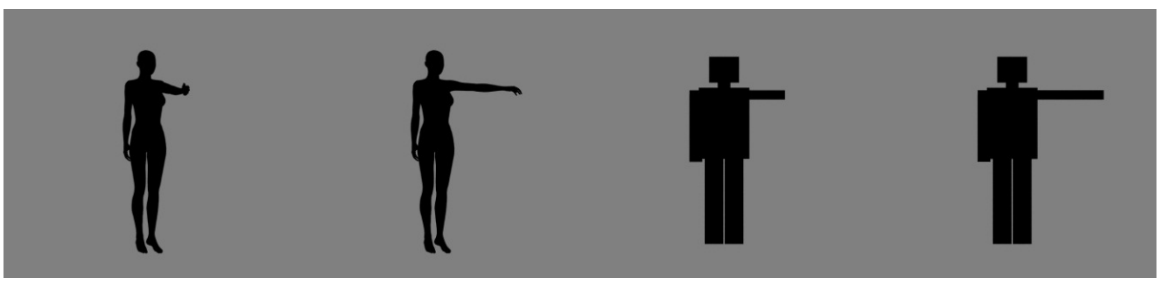


Figure 14: The images are shown at non-cardinal angles so that it easier to see the asymmetry

Once again the same reasoning for using robot-like and human figures applies, however in this case the difference in Left-Right symmetry will allow us to look for a new bias based on depth reinterpretation. For all KDE images, including stimuli used in experiment set 1, the perceptual multistability of the stimulus is experienced as a directional switch between clockwise and anti-clockwise rotation. In Experiment Set 1 the peak locations of the directional switches occurred at points in the object rotation in which there was depth symmetry in the frontal plane. This meant that those directional switches were experienced as merely a change in rotational direction with no reinterpretation of the object in depth. However, for 3 out of 4 stimuli in the following experiment there are no points in the object rotation where front-back symmetry exists. The result is that any time a subject experiences a directional switch for those stimuli there MUST also be a corresponding mirror depth reversal and 3-dimensional reinterpretation of the object. Additionally both the humanoid and the robot have two different poses (Arm-Forward & Arm-Outstretched-Sideways) which will have different amounts of total depth reinterpretation depending on the position in the revolution where the directional switch occurs.

Any KDE stimulus will have a different amount of 3-Dimensional depth depending the distance of the extreme points of the object from the axis of rotation. As the object spins the most extreme point will vary in its distance from the frontal plane sinusoidally. The Arm-Forward and Arm-Outstretched-Sideways stimuli in this experiment reach their depth maxima and minima at different points in the rotation. (Wallach & O'Connell, 1953)

For the Arm-Forward stimulus the depth maxima occurs when the object is in a facing-towards or facing-away position, and the depth minima occurs in the profile positions. Conversely for the Arm-Outstretched-to-Side object the depth maxima occur in the two profile positions and the depth minima occur in the facing-towards and facing-away positions.

Thus a competition has been set up between two potential biases. The tendency is for direction reversals to occur in the profile positions which result in the object tending to face "towards" the subject, as was found in experiment set 1. And a potential depth translation bias, in which it might be assumed that the visual system would have a preference for less overall reinterpretation of depth.

Experiment 2.1 - Methods

Stimuli

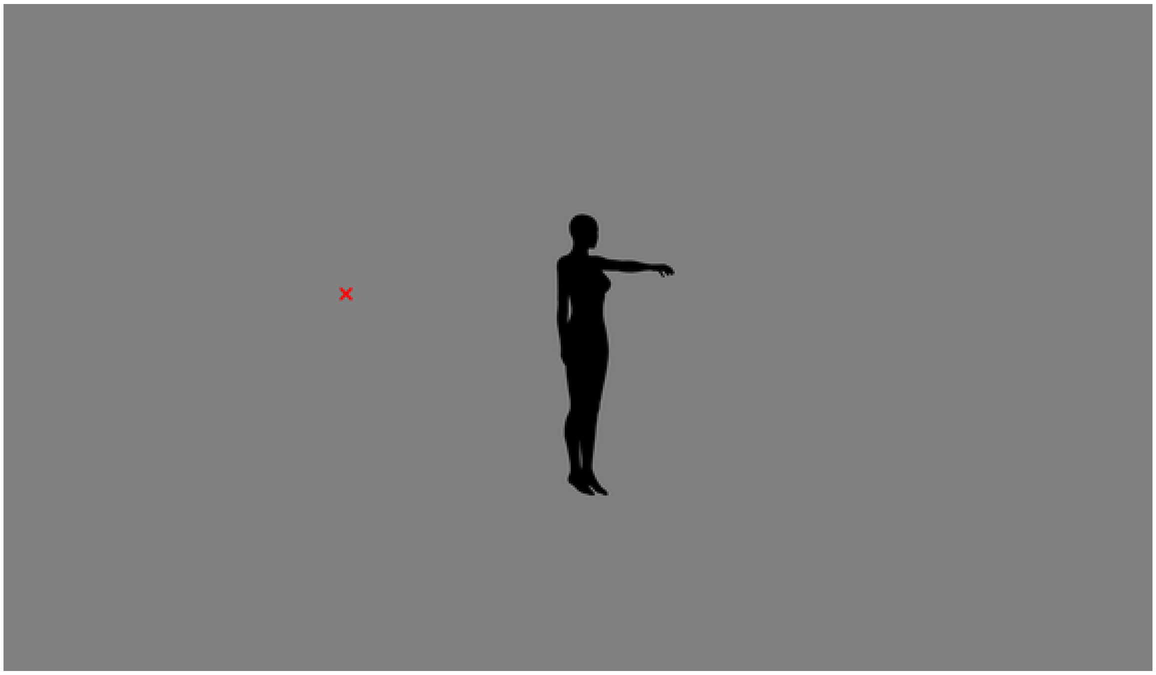


Figure 15: Sample stimulus from experiment 2.1

As in Experiment Set 1, DAZ Studio 3 was used to create the stimuli.

Each of the four stimuli consisted of 82 views (frames) representing a full revolution of a 3D object. This resulted in an angular rotation between successive frames of 4.39 degrees. The frames were presented at a rate of 15 frames per second which resulted in one full revolution of the object every 5.47 seconds.

To avoid physical distortions due to perspective the all of the images were presented in a frontal orthographic projection.

The background color of the stimuli was a mean grey. The 3D silhouettes were a pure black.

Subjects were seated 26 inches from the monitor. At this distance the monitor distended a viewing angle of 31.2 degrees horizontally and 14.9 degrees vertically. The images subtended between 2.2 to 5.2 degrees of horizontal visual angle depending on the frame, and 8.8 degrees of vertical visual angle.

Subjects

37 undergraduate students at UCSD were presented with a rotating silhouette that utilized the Kinetic Depth Effect to create an ambiguity in the direction of rotation. Subjects were compensated for their time with class credit for their current courses. Seven of subjects' data were not included in the final analysis. Six of these subjects were removed for not following the experimental protocol, while the seventh was removed for falling asleep on the keyboard, resulting in a very large, but ultimately useless data set. This subject was directed to the location of a sleep research laboratory.

Instructions

Subjects were instructed to indicate the initial direction of motion of the stimulus (clockwise or anti-clockwise as viewed from above) when it first appeared with a keypress [<] or [>]. Thus [>] indicated anti-clockwise rotation as viewed from above, and [<] indicated clockwise rotation as viewed from above. They were told to then press the [<] or [>] key to indicate the new direction of rotation whenever a directional change in motion occurred. They also had the option of pressing the [SPACEBAR] if the image started to do something that could not be described as rotating.

Image Presentation

Stimulus presentation order was randomized for each subject. During the experiment each stimulus was presented in the center of the screen for the duration of each trial block. During a single block the subject would first be allowed to freely view the image for 120 seconds. After that viewing period the screen would temporarily blank out for 1 second as a fixation cross would randomly appear to the left or right of the screen (9 degrees off center). The subjects had been previously instructed to foveate on the fixation cross. One second after the screen blanked out the ambiguous figure would reappear. After each block of Center/Left/Right viewing the subject was given an option to briefly take a break and continue the experiment when they were ready.

Data Collection

The data collected was the current frame being presented when the computer registered the subjects' response indicating they noticed a change.

Thus there was a slight lag between the time the subjects noticed a change, decided to press a button, and completed the button press. There was likely a reaction time delay of about 400-500ms before the button press was registered (Kornmeier & Bach, 2012). This translates into a delay of at least7 frames of the stimulus. Thus the peaks in the data presented below occur slightly *after* the subjects experienced a directional switch.

To confirm whether or not subjects were responding slowly due to a reaction time lag a control experiment was conducted which presented subjects with a single KDE image that was presented at two different speeds. The results of this experiment, shown later, confirm that there seems to be a consistent response-time lag which can at least partially account for subject response peaks at non-cardinal points of the rotation. There is also likely a small delay from image processing.

The total run time for each subjects was 24 minutes (6 minutes per stimulus), this allowed for 88 full rotations of each stimulus with a total of 264 full rotations across all stimuli.

At the end of the experiment subjects were asked to debrief the experimenter. This consisted of verbal descriptions of what they saw and questioning as to specific positions where the images seemed to change direction.

Experiment 2.1 - Results & Conclusions

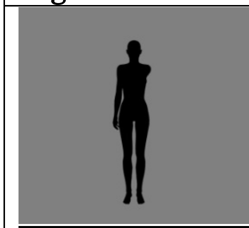
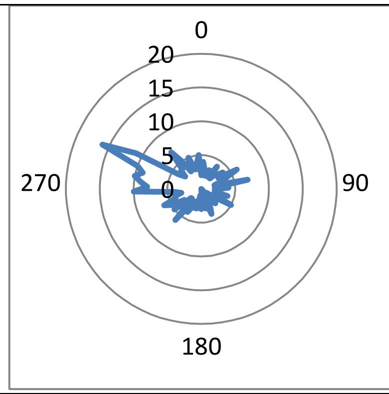
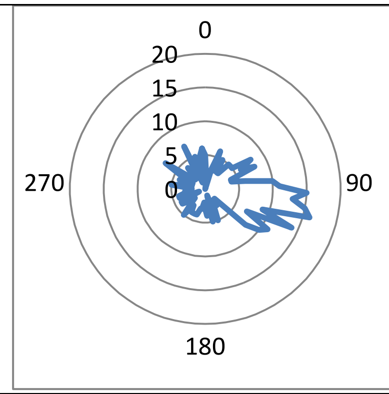
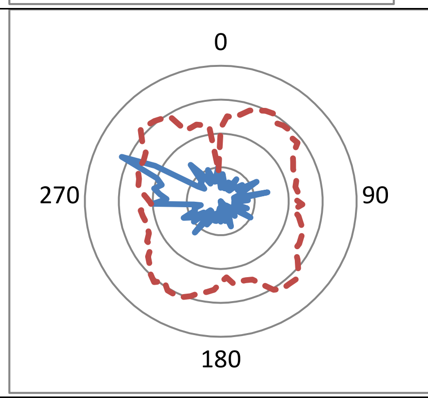
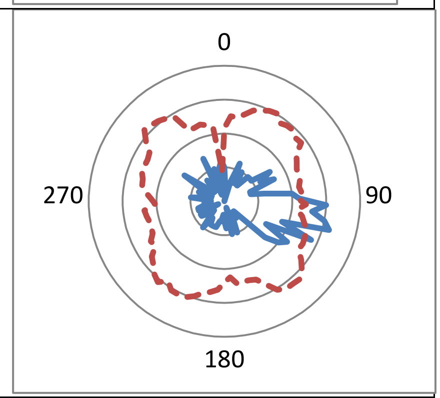
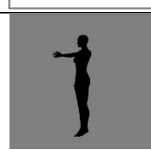
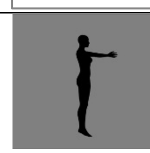
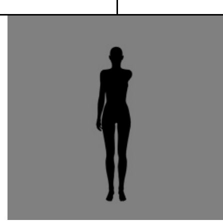
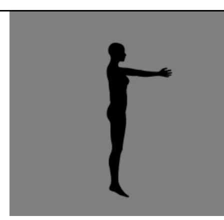
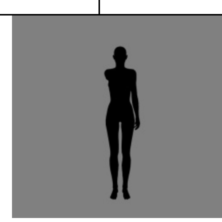
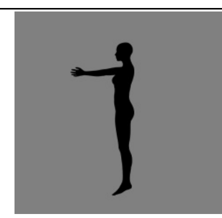
<table>
<tr><td>Figure 16a</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test: Rayleigh test for non-uniformity</td><td>
mode: 298.7
mean: 300.4
p-value = 9.21 e-011
Total # switches: 307</td><td>
mode: 109.5
mean: 99.7
p-value = 2.04 e-026
Total # switches: 337</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 16 a: Results from humanoid with one arm forward stimulus

16a. To best understand and analyze the results of this second experiment, they must be compared to the results from the corresponding stimulus in experiment 1.

In experiment 1 stimulus 3c (a humanoid figure with both arms forward) tended to induce directional switches when it was in a profile position and about to start turning to face away from the subject. While it was clear where in the revolution directional switches were most likely to occur it was not clear why those points were preferred by the visual system. In this case there are two possible suggestions. 1) This was a point in the rotation in which the stimulus was about to being “facing away” from the subject. As was noted earlier there is evidence that the visual system has a “facing towards” preference for ambiguous stimuli, so the directional switches could be occurring at this position to maintain that preference (Vanrie, Dekeyser, & Verfaillie, 2004). 2) Alternatively it is possible that directional switches were occurring at the location because it was a point in the revolution where the image had front-back symmetry. Thus a directional switch at this location does not require any reinterpretation of the 3D object.

The current stimulus, 16a, is identical to that from experiment 1, 3c, save for the fact that rather than both arms being forward, only a single arm is forward. The effect of this manipulation is that every point in the rotation of the stimulus, even the profile position, will require some amount of 3D reinterpretation of the object in depth.

The data show, that despite the fact that a depth reinterpretation must occur simultaneous with the directional switch, subjects will still experience directional switches when the stimulus is in the profile position (90 & 270 degrees). As shown in the Figure 16a the points of the revolution where subjects' directional switches peaked are at 109.5 & 298.7 degrees. Also, as in experiment 1, the peaks for Left-Response and Right-Response were roughly 180 degrees to one another (189.2 degrees) and indicate that subjects still tended to experience directional switches which would keep the stimulus facing towards them.

However this data alone cannot disambiguate why directional switches are occurring most often in the profile view. Specifically the following two points must be considered:

1) While a reinterpretation of depth IS occurring, it is possible that the profile positions are points within the rotation where those reinterpretations result in the least amount of total reinterpretation of the object. This is due to the fact that at the profile position the most distant points of the object, from its axis of rotation (the forward arms), are very close to the frontal plane of the image. This means that when the depth reversal occurs there is a smaller total distance in depth that each part of the object

needs to undergo. Compare the profile position of this object to the front-back position. In the front-back position the extended arm will be at its furthest forward or backwards position in depth. If a direction reversal were to occur at the front-back position it would require a much larger total distance transformation in depth of the extended arm. One moment it would be coming out towards the subject, and the next moment it would be behind the rest of the 3D object and pointing away from the subject. Thus, while a transformation of depth must still occur at the profile position, it may be the "least costly" point of the rotation for the depth transformation to occur (Knill, Kersten, & Yuille; 1996).

2) As discussed previously it is possible that linear image motion may be a factor in determining when a directional switch will occur. In the case of stimulus 16a, the profile positions are two points in the object's revolution where the linear image velocity has a local minimum (as seen in the graph of linear image motion in Figure 16a). It is possible that the slowing of the local motion signal is also a factor in when a directional switch is likely to occur. This idea is further explored in experiment 3.

As a final note, as in experiment 1 this stimulus had reversal peaks which occurred slightly after the positions which seem to be triggering these reversals. See Section E for the Control Experiment measuring response time.

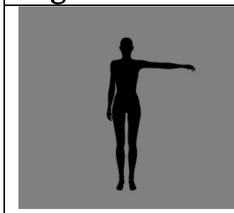
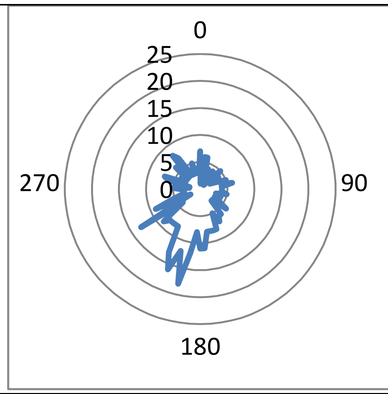
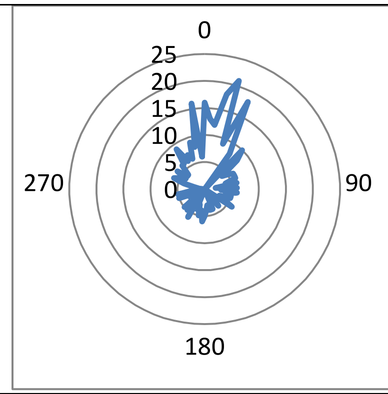
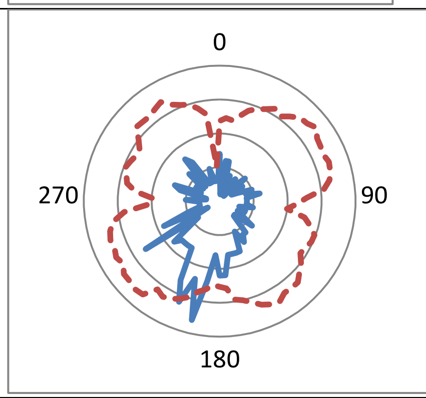
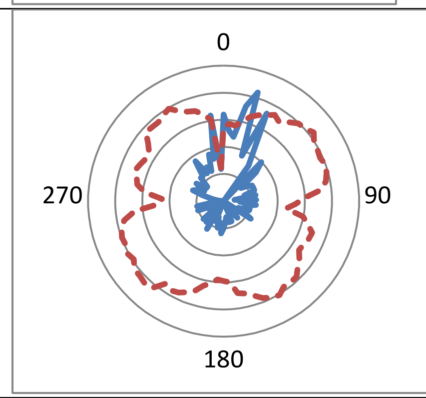
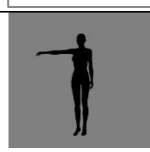
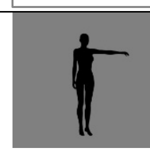
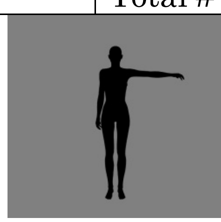
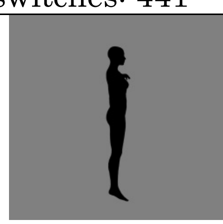
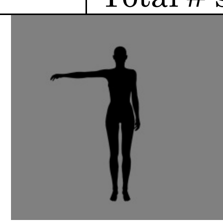
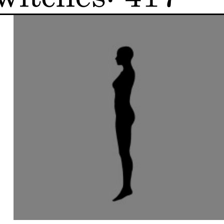
Figure 16b	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode: 197.6 mean = 202.17 p-value = 1.22 e-014 Total # switches: 441	 mode: 22.0 mean = 17.59 p-value = 2.86 e-024 Total # switches: 417
Degrees:	 0 |  90	 180 |  270

Figure 16 b: Results from humanoid with one arm out to side stimulus

16b. The second experimental stimulus is once again a non-symmetric analogue of a figure (Figure 6b) from experiment set 1, but this time the figure has one arm held out to its side. In the last figure considered, with single arm extended forward (Figure 16a), the profile position minimizes the spatial displacement of the extended arm in a depth reflection. But in Figure 16b, where a single arm is extended sideways, the deviation from symmetry about the frontal plane in the profile position is very large, and it is the facing positions that minimize translation of the extended arm and hand in a depth reflection. On the other hand, a depth reversal in the front or rear facing position incurs an added cost in that front and back must be swapped, a cost that does not arise in the profile position, where only the arm is affected. Thus consideration of spatial arrangement does not clearly predict a preference either for profile reversals or for frontal reversals.

The result here is clear, however. When a single arm is held sideways (Figure 16b) subjects tended to see perceptual rotation reversals most often in the front facing position. This behavior differs sharply from that of all the other human figures considered thus far, which all showed a strong preference for reversal in the profile view. In this case the number of reversals peaked at 197.6 degrees & 22 degrees (the front and back positions were at 0 degrees & 180 degrees).

This could suggest that the cost of depth-reflecting the entire figure front-to-back (necessary when any figure is reversed in the frontal position) is less than the cost of translating the outstretched arm in depth (when the asymmetrical figure is reversed in the profile position).

But alternatively, the typically frequent profile view rotation reversals may be absent in Figure 16b because the image velocity information here discourages profile reversals. The image of the single arm stretched out sideways sweeps at a sustained high velocity through the profile view, with only a momentary (two-frame) occlusion, and this steady apparent movement is naturally consistent with a correspondingly progressive rotation. For an observer to experience perceptual rotation reversal in the profile view, the 3D depth reflection that is allowed by the momentary ambiguity of the profile image would have to be imposed with precisely appropriate timing and at the price of a perceptually implied but physically impossible discontinuity of angular momentum. Here then is a case where the cue of image velocity may be exerting a controlling influence, by clearly indicating a steady rotational velocity, in contrast to the situation where rotational reversals are associated with image standstill.

It has already been noted that this stimulus was more likely to undergo a direction reversal in a front-back, rather than profile, position. But more specifically that reversal occurred primarily when the stimulus appeared to be facing the subject. Thus when the direction reversal, occurred the stimulus immediately went from facing-towards to facing-away from the subject. If there is a bias for preferring to see the stimulus facing towards the subject, would it not also be expected that a significant number of reversals would occur such that the stimulus initially appeared to be facing-away from the subject and after reversing direction was facing directly towards the subject? Why did this type of reversal not occur more frequently? In terms of depth reflection the cost to from front-to-back is identical as the cost to going back-to-front. There would also be an identical perceived change angular velocity of the object regardless of where the reversal started. Thus there must be some factor that either is inhibiting a reversal which would result in a forward facing stimulus, or which is greatly benefiting a reversal that results in an away-facing stimulus. One possibility is that object reversals from front-to-back result in the outstretched hand remaining in the hemisphere of rotation that is nearest to the viewer. However the results from stimulus 16d (the robot with an arm outstretched to the side) show no such preference for the outstretched limb being close-to or far-from the viewer. Perhaps the robot is not human-enough for the visual system to care whether or not its arm is close or far, or perhaps the extra image contours that the humanoid figure has compared to the robot are adding an extra sense of depth which is sufficient to create a preference for seeing the arm swinging close to the subject.

If we compare the results of Figure 16a (single arm forward) to the left-right symmetric counterpart with both arms extended forward (Figure 3c), we can see that the directional switches of the non-symmetric image are completely out of phase. So the simple change of having one arm outstretched rather than two has shifted where directional switches will occur within the 3D rotation. This the reduction in ambiguity provided by the progressive movement of the arm across the profiled torso Figure 16b could generate a reduced number of profile reversals rather than an abundance of frontal reversals. However the absolute frequency of reversals in Figures 16a and 16b contradicts this as there are 33% more reversals for figure 16b (644 reversals in 16a, 858 reversals in 16b. While this result is counter-intuitive a possible explanation will be addressed at the end of this section as this data is compared to that of the two robot stimuli.

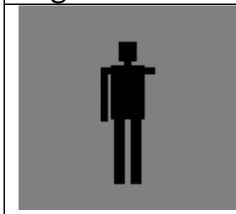
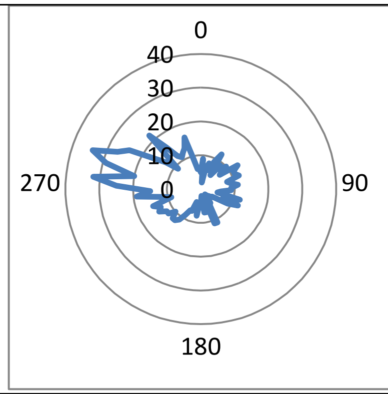
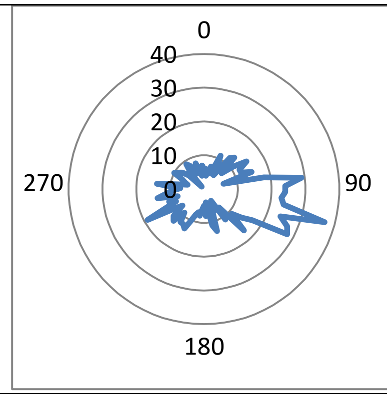
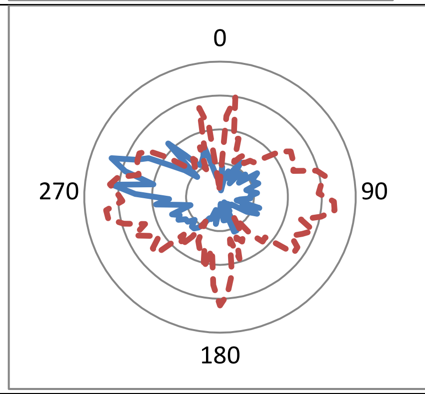
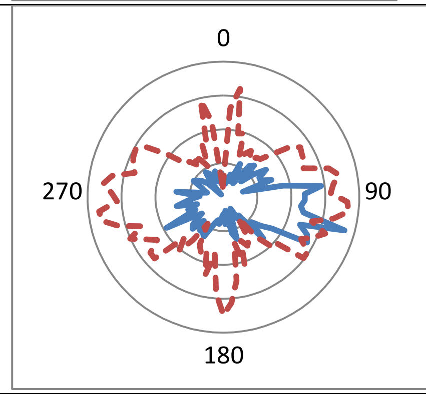
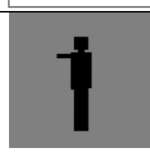
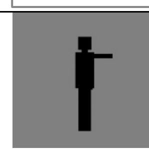
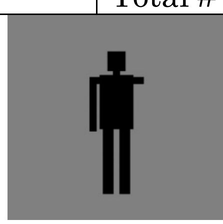
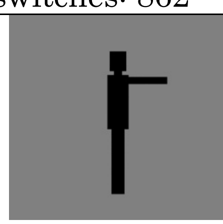
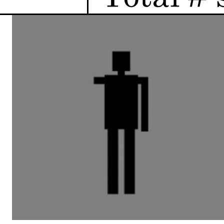
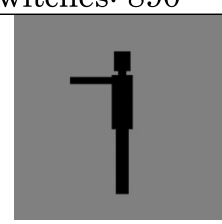
<table>
<tr><td>Figure 16c</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test:
Rayleigh test for non-uniformity</td><td>
mode: 294.1
mean = 298.00
p-value = 4.40 e-029
Total # switches: 862</td><td>
mode: 109.8
mean = 120.04
p-value = 6.12 e-021
Total # switches: 890</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 16 c: Results from "Robot" with one arm forward stimulus

16c. The results from the third experimental stimulus (16c) provide even more insight into exactly what may be happening with internal biases and a cost function for directional switches. A major difference between this stimulus and that of the human figure 16b, is that the robot figure has no discernible features to indicate which side is its front or back. The robot also an analogue of the one from the first experimental set (Figure 3c), but this time with only a single arm held forward.

As with figure 16a, subjects reported the majority of direction reversals in the profile positions. Also, like stimulus 16a, the profile positions correlate with the points in the revolution that would result in the least amount of depth reinterpretation of the object, as well as the points where linear image velocity reaches a local minimum. Thus the preference for directional switches in the profile position is expected, given the similarity between 16a and 16c

A result that was quite unexpected is the fact that the total number of directional switches for 16c (1752 switches) is massively larger than that of 16a (644 switches). This represents an over 270% increase in total switches. There were 1978 revolutions of 16c across all subjects (3956 half-revolutions). This means that on average 88.6% of revolutions contained at least one directional switch.

A detailed analysis of what may be causing this large difference in total number of switches is included at the end of this section.

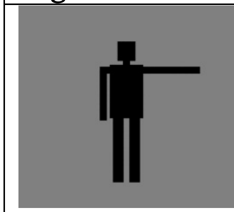
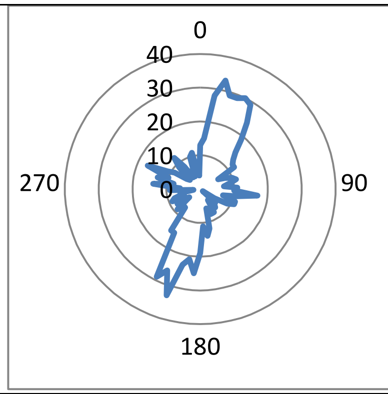
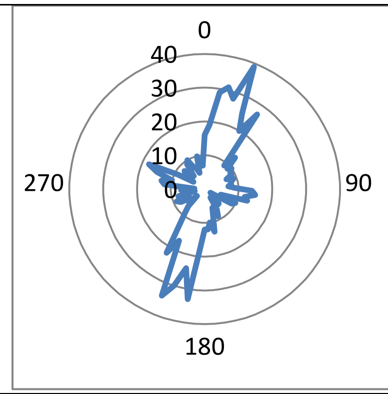
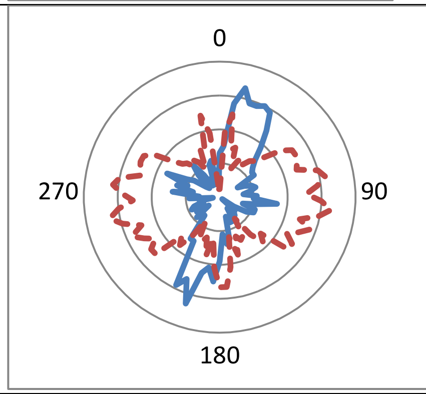
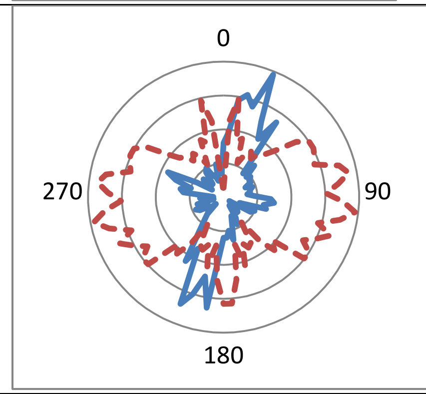
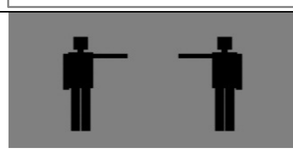
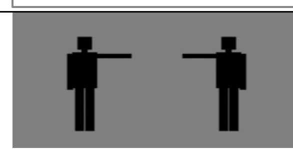
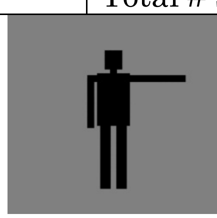
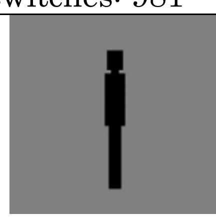
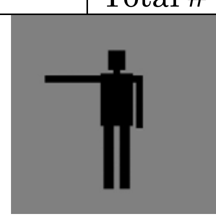
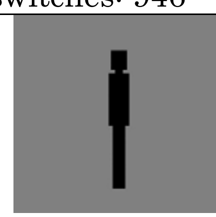
<table>
<tr><th>Figure 16d</th><th>Anti-Clockwise Reversal</th><th>Clockwise Reversal</th></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td>Statistical Test: Omnibus Hodges-Ajne test for non-uniformity</td><td>
Omnibus Hodges-Ajne
bimodal: 26.3, 206.3
mean = 54.07
p-value = 0.01
Total # switches: 981</td><td>
Omnibus Hodges-Ajne
bimodal: 26.3, 206.3
mean = 39.83
p-value = 0.05
Total # switches: 946</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 16 d: Results from "Robot" with one arm out to side stimulus

16d. The final stimulus, 16d, is related to both stimulus 8b from experiment 1 (they are both robot figures with the arm position outstretched sideways) and stimulus 16b in the current experiment (since there is only a single arm outstretched).

The first thing to note is that unlike figure 8b, this stimulus does have clear points in the revolution which result in a directional switch. It is impressive that a simple addition of left-right asymmetry was able to eliminate the problems subjects had in being able to experience figure 8b as a 3D object, rather than a 2D shape that was expanding and contracting horizontally over time. There are likely two factors that contribute to this change.

1) Firstly the asymmetry of 16d results in the front and back views of the object being different; in one case the outstretched hand is towards the right, in the other case towards the left. Whereas in figure 8b these positions in the rotation were identical and thus a full 360 degree rotation resulted in images from the first 180 degrees being identical to images from the second 180 degrees.

2) The second factor which likely helped to eliminate the loss of 3-dimensionality was the fact that one arm was held down at the side. This created an additional image contour between the arm & body. It has been shown, that for silhouettes, contours are a major factor in determining whether the object is a seen as a 3D rotation or a 2D scaling (Ullman, 1979).

An examination of the overall distribution of the data for this stimulus shows two clear peaks for directional switches, as well as a possibility of two minor peaks 90 degrees from the major ones. The fact that each image has two major peaks instead of one is unsurprising due to fact that there is no discernible “front” to the stimulus. This also explains why the Left-Response and Right-Response graphs are basically identical.

The two major peaks for directional switches (26.3 degrees & 206.3 degrees) correlate with the front and back orientation of the stimulus similar to figure 16b, as opposed to a peak for switches in the profile position that are seen in 16a.

The positions of the major peaks show, that as in 16b, the direction reversals are correlated with two stimulus properties, minimum linear image velocity and minimum depth reinterpretation (in this case zero).

Yet visual inspection suggests that there may be two minor peaks in the data as well (116.3 degrees & 290.3 degrees), which represent image frames of the stimulus close to the profile position. This is odd because a directional switch at these positions will not only require a large depth reinterpretation of the object, but also the linear image velocity is at a maximum here (save for three frames where the thin outstretched arm is occluded by the thick body). This suggests that despite the large depth reinterpretation that must occur, there was still some small preference for directional switches in the profile position.

Comparison of total # of directional switches

Table 3: Total number of directional switches, Clockwise and anticlockwise for each stimulus from experiment 2.1

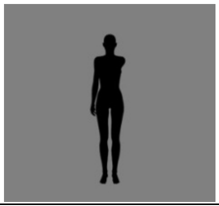
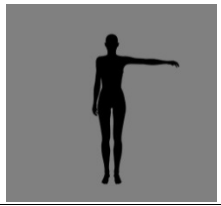
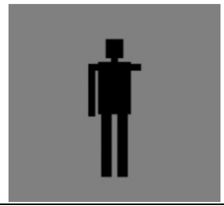
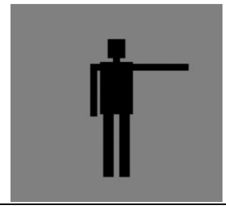
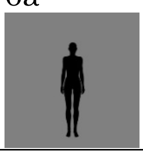
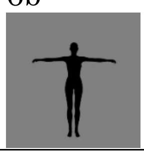
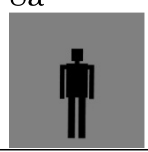
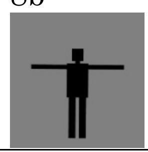
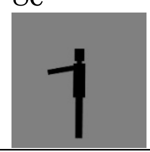
	16a 	16b 	16c 	16d 
Anti-Clockwise	307	441	862	981
Clockwise	337	417	890	946
Total	**644**	**858**	**1752**	**1927**

Table 4: Total number of directional switches, Clockwise and anticlockwise for each stimulus from experiment 1.1 & 1.2

	6a 	6b 	6c 	8a 	8b 	8c 
Anti-Clockwise	572	584	355	626	429	600
Clockwise	571	578	341	616	411	578
Total	**1143**	**1162**	**696**	**1242**	**840**	**1178**

Across all subjects there were a total of 1978 Revolutions per stimulus or 3956 Half-Revolutions.

1) The data indicate that robot stimuli had massively more switching than humanoids for the experiment 2.1. This can be compared to stimuli 6c (696 switches) and 8c (1178 switches) in experiment 1 which

show a similarly massive difference. The robot in 8a also had more directional switches (1242) than the humanoid in 6a (1143), however the scale of that difference is far smaller than what was found elsewhere, thought this may have been the result of subjects not seeing 8a as strongly in 3D as 8c, both due to the limited contours of the stimulus and self-occlusion of the gaps between the legs and arms for most of the revolution. A comparison of 6b to 8b doesn't really apply due to the significant problems subjects' experienced in seeing 8b as a 3D object.

2) Ullman (1979) demonstrated that the complexity of image contours directly affect the ability to see a KDE stimulus as a rotation 3D object vs a 2D deformation. It is possible that the larger amount of contour information for the humanoid figures gave not only a stronger perception of 3D rotation, but also a more complex and complete internal 3D representation. Consider this in the context of the evidence shown here that objects are more likely to reverse in positions where less 3D reinterpretation is required. It therefore follows that if the internal representation of the humanoid figures it more robust, then those figures will always require more 3D reinterpretation during a depth reversal and thus they will undergo fewer directional reversals than the less contoured robot stimuli.

Summary Analysis of Experiment 2.1

All of the stimuli from this experiment have direction switches which peak at points in the revolution that correspond to both minimum object reinterpretation and minimum linear image velocity. Thus, with the results of experiment 1. three potential factors have been identified that might affect where a directional switch is likely to occur:

1) Directional switches occur frequently in positions where a minimum amount of reinterpretation of the object in depth is required.
2) Directional switches occur frequently when the object is in the profile position and the new direction of rotation will keep the stimulus facing towards the subject. However this was mostly eliminated in cases where a large depth reinterpretation was required for a directional switch in the profile position.
3) Directional switches occur at points in the object rotation where linear image motion is at a minimum. However, when depth reinterpretation was not a factor, directional switches were more likely to occur in the profile, even if that position was where linear image motion was at a maximum.

Thus presumed order of importance of the potential factors identified are:

1) Minimum depth reinterpretation
2) Profiles switches to maintain forwards facing object
3) Linear image motion at a minimum

Experiment 2.2

The second experiment was originally going aimed at using stimuli which consisted of only a head or only a body to determine which aspect of the image might have a greater impact on perceptual switching. However, during the course of running pilot experiments a surprising result emerged. When viewing a smaller head by itself subjects still had a bias for seeing directional switches in the profile position, but there was a roughly equal likelihood of the head being seen as facing-towards or facing-away. This directly contradicted the notion of a general forward facing bias from the full-body stimuli in the previous experiments. Further pilot experiments revealed that when the subjects were shown a larger head the preference for seeing it face towards them re-emerged. A full experiment was conducted on this Large head/ Small head difference to clarify if the difference seen in the pilot experiments held up and to explore why size might be making a difference at all.

Experiment 2.2 - Methods

See Experiment 2.1

Stimuli

As before, DAZ Studio 3 was used to create the stimuli.

Each of the two stimuli consisted of 82 views (frames) representing a full revolution of a 3D object. This resulted in an angular rotation between successive frames of 4.39 degrees. The frames were presented at a rate of 15 frames per second which resulted in one full revolution of the object every 5.47 seconds.

Figure 17: Large head stimulus

The large head varied in visual angle horizontally from 8.41 degrees (in profile) to 6.51 degrees (front-back) and was 7.15 degrees vertically.

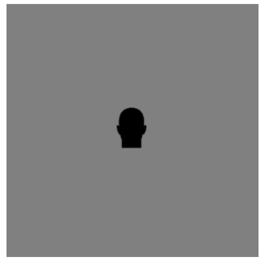
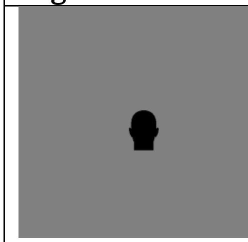
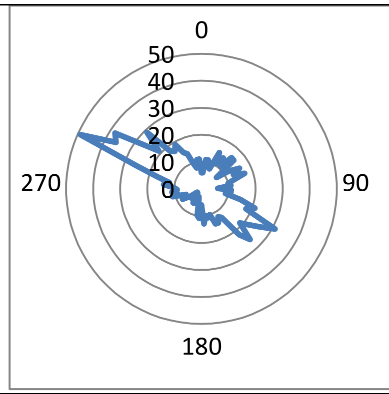
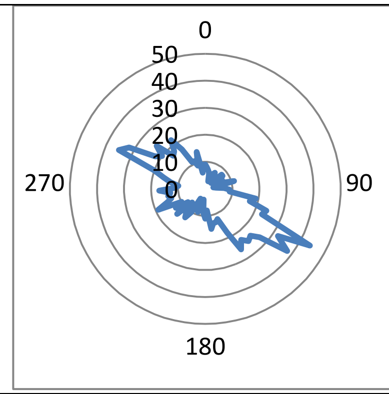
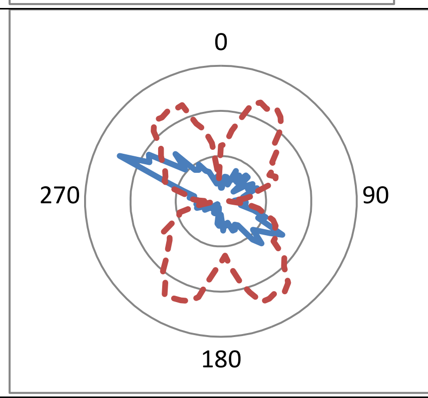
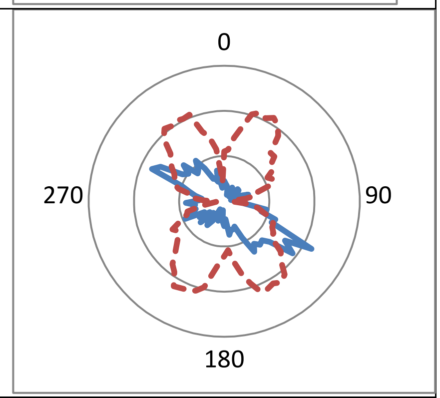
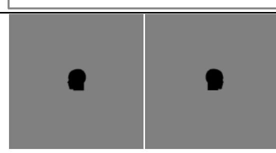
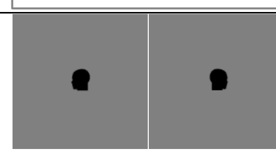
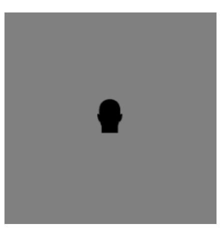
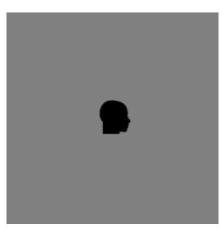
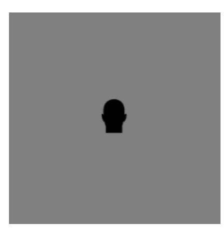
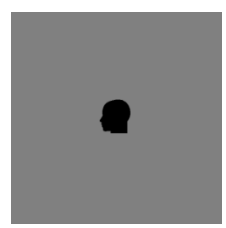


Figure 18: Small Head Stimulus

The small head varied in visual angle horizontally from 2.07 degrees (in profile) to 1.64 degrees (front-back) and was 1.79 degrees vertically.

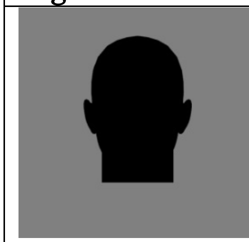
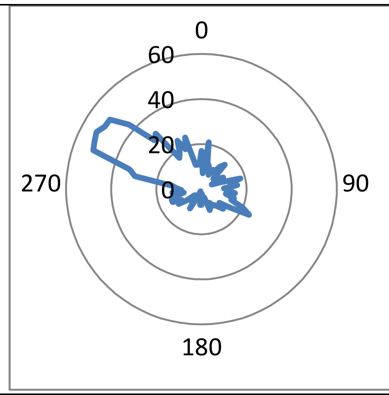
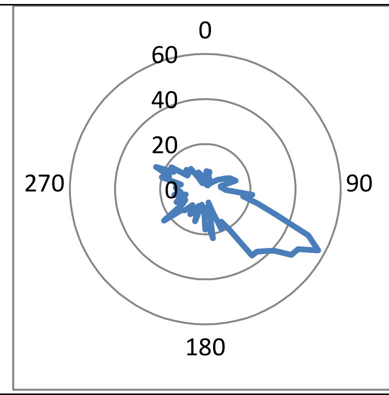
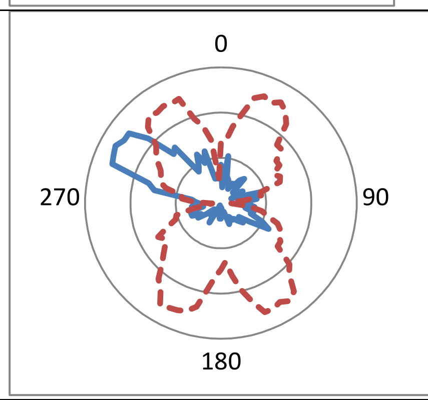
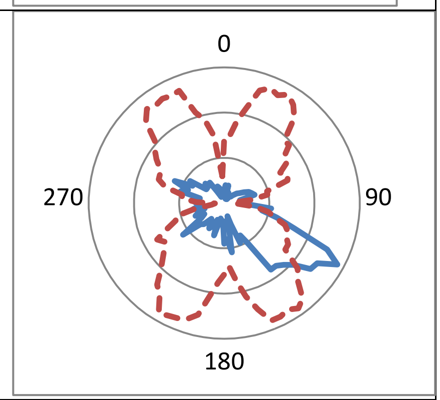
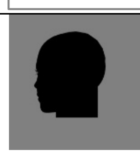
Experiment 2.2 - Results & Conclusions

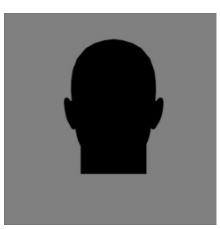
Figure 19a	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Omnibus Hodges-Ajne test for non-uniformity	 Peaks: 298.5, 122.9 Omnibus Hodges-Ajne mean = 7.61 p-value = 3.26 e-005 Total # switches: 1079	 Peaks: 298.5, 122.9 Omnibus Hodges-Ajne mean = 199.47 p-value = 2.08 e-005 Total # switches: 1097

Degrees: 0 | 90 | 180 | 270

Figure 19 a: Results from small head and large head in experiment 2.2

Figure 19b	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 Omnibus Hodges-Ajne mode = 302.9 mean = 324.17 p-value = 1.79 e-052 Total # switches: 1180	 Omnibus Hodges-Ajne mode = 122.9 mean = 154.71 p-value = 8.15 e-049 Total # switches: 1173

Degrees: 0 | 90 | 180 | 270

Figure 19 b: Results from small head and large head in experiment 2.2

19 a & b. Both the small and large head can be discussed together. In the case of the larger head we see the typical response pattern of directional reversals occurring most often in the profile position. Additionally we see the usual bias of reversals occurring so that the head tends to face towards the subject.

The surprising result comes when we look at directional reversals for the smaller head. These two stimuli were absolutely identical except for their size. However, when viewing the smaller head, subjects were far more likely to experience rotation reversals that would cause the head to remain facing AWAY from them, rather than the typical reversal pattern which tended to maintain a forward facing head. The bi-lobed nature of the distribution is made clear by the fact that both clockwise response and anticlockwise reversal distributions have peaks that occur at 298.5 & 122.9 degrees. Yet the mean resultant vectors for these distributions are 7.61 degrees (Anti-Clockwise Reversal) and 199.47 (Clockwise Reversal). These mean resultant vectors are so far off from the peaks because the peaks are roughly 180 degrees opposite of one another and when the mean resultant vector is calculated the peaks cancel each other out.

It is not yet clear what is causing this difference, however we speculate that several factors could be involved.

1) The larger head is very close in size to the previous stimuli used (Large Head = 8.41x7.15 degrees) (Experiment 1 Stimuli ~ 8.8x8.8 degrees). This suggests that there may be an effect of receptive field size relative to the overall size of the image.
2) At a higher level there may be less of an expectation for a smaller (i.e. further away) head to be facing towards you.
3) Linear image motion for the smaller head will necessarily be slower than the larger head, but these directional switches primarily occurred in the profile view where linear image motion was near a minimum.
4) One reason that the opposite of this result might be expected is that, like all silhouettes, these stimuli show no facial features in the facing-toward view. Under natural conditions, when a head is far enough away to subtend a sufficiently small visual angle, such features would be easily obliterated by failure of visual resolution. But for larger head images, those facial features would be clearly visible. Therefore the absence of these features in the larger silhouette might be expected to suggest a facing-away orientation. This however is not the case, which suggests that either this high-level expectation is unimportant, or is it is being over-ridden by some other visual cues which cause the large head to appear to face towards the subjects more readily.

It is interesting to note that even though the smaller head has a higher likelihood to undergo a directional switch in more than one position, the other likely position is still a profile view, possibly because this switch point is symmetric and requires no reinterpretation in depth of the object, just its 3D angular velocity.

To further show that the small head and large head were giving different response distributions a Kuiper two-sample test was run comparing the results of each experiment to each other. The large head Clockwise Reversal was compared to the small head Clockwise Reversal, and the large head Anti-Clockwise Reversal was compared to the small head Anti-Clockwise Reversal. In both cases the result was that the two distributions were dissimilar with $p < 0.001$. Note also that this was the maximum possible statistical significance that the Kuiper test can provide, due to the use of tabulated values in the CircStat for Matlab Toolbox which was used to perform the analysis (Berens, 2009).

Finally the statistics were re-run after removing the data from the central viewing condition. This is because the added acuity of the central viewing condition may have been a factor in the difference seen between the big-head and small-head results. This new analysis resulted in no major difference across conditions, suggesting that smaller and more accurate receptive field size in the central viewing condition is not likely to be a factor in the resultant data.

An analysis of the total number of directional switches of the two stimuli relative to one another shows a slightly higher (8%) number of reversals for the large head. This difference is minimal compared to previous differences found, but other (unreported) pilot experiments also suggested that smaller stimuli might undergo fewer directional reversals. Further studies would be necessary to determine if size is indeed a significant factor in the total number of directional reversals.

Table 5: Directional switch comparison between Small Head and Large Head

Number of Directional Switches	Small Head	Large head
Anti-Clockwise	1079	1180
Clockwise	1097	1173
Total	**2176**	**2353**

<u>Experiment 2.3</u>

The final experiment in this set was designed to see what would happen when the bias for head facing forward and the bias for body facing forward were set up to compete against one another. The stimulus was a human figure with her leg out and head turned.

Figure 20: Competing head/body bias with head in profile

Figure 21: Competing head/body bias with body in profile

The main goal of this experiment was to see if competing stimuli like this would create additional reversal points based on where the head or body was positioned, and to determine whether it was the head or body that exerted more control over our final experience of where the perceptual switches occurred.

Experiment 2.3 - Methods

See Experiment 2-1

Subject Instructions

Subjects were instructed that there would be three separate viewing conditions: Free viewing, fixating on the head of the stimulus, or fixating on the body of the stimulus.

In all cases subjects were instructed to use the keyboard to indicate the initial direction of rotation of the object as well as any subsequent changes in the direction of rotation.

Subjects were first instructed to freely view the object. The viewing period began when the subject pressed a key and ended after 120 seconds.

Next a screen would appear instructing the subject to fixate on either the head or the body (randomly). Once again the viewing period began when the subject pressed a key and ended after 120 seconds.

Finally a screen would appear instructing the subject to fixate on either the head or the body (whichever was remaining). Once again the viewing period began when the subject pressed a key and ended after 120 seconds.

<u>Stimulus</u>

In an effort to further explore what factors might be involved in creating directional reversal points an image sequence was created that could potentially have two different reversal points. It has been noted earlier that, for both a head alone and for a full body, rotation reversals will often be centered around the profile view (Figs 6a,b&c and 11a&b). This stimulus has the head turned about 90 degrees with respect to the body so that the head and body will not simultaneously be in profile. Additionally a single leg is tilted slightly backwards to create further asymmetry within the object so that any directional reversals will require a full image reinterpretation in depth.

Experiment 2.3 - Results & Conclusions

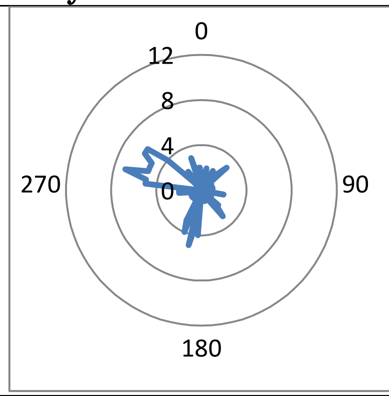
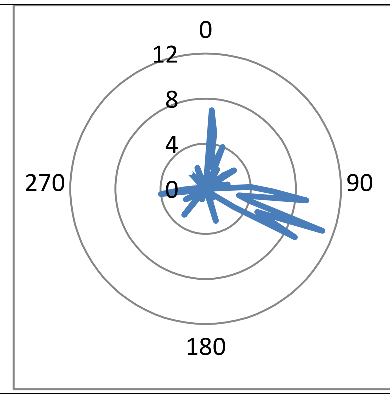
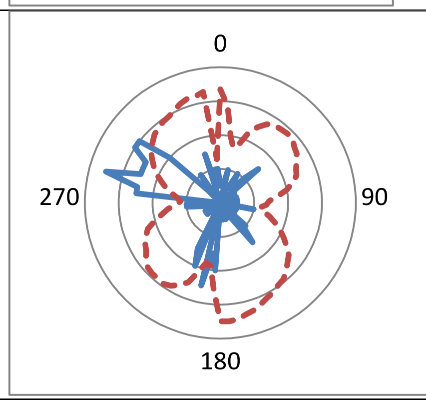
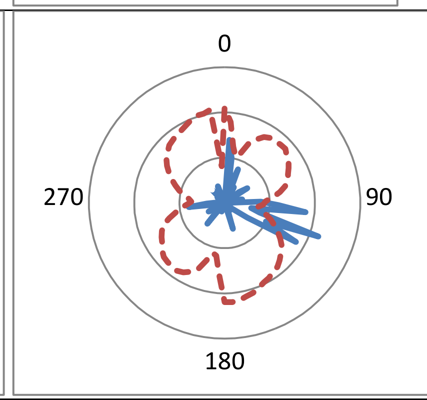
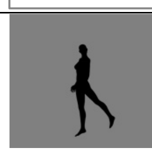
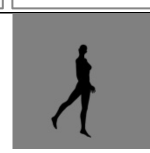
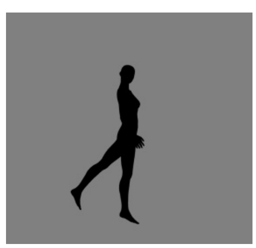
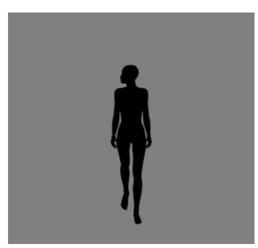
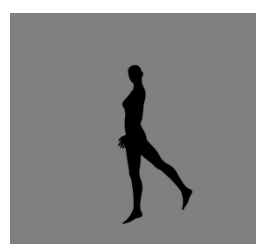
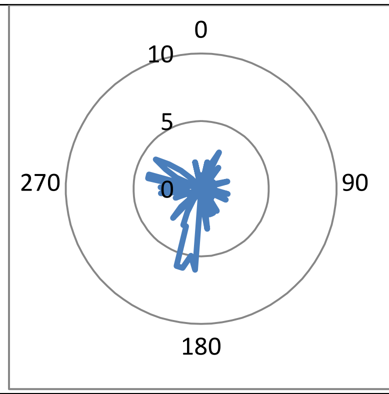
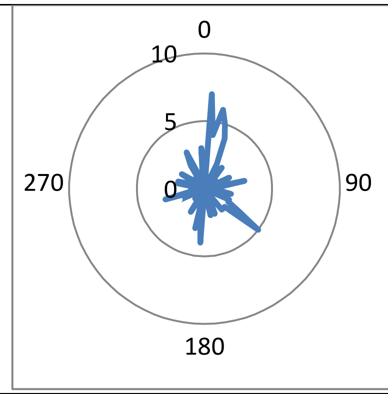
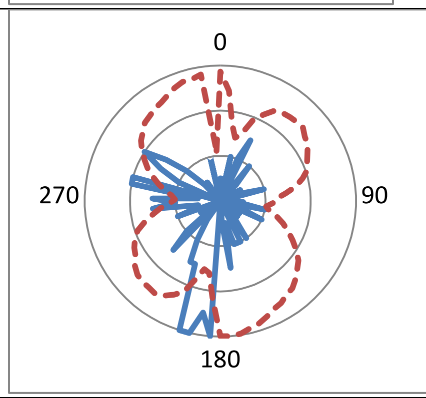
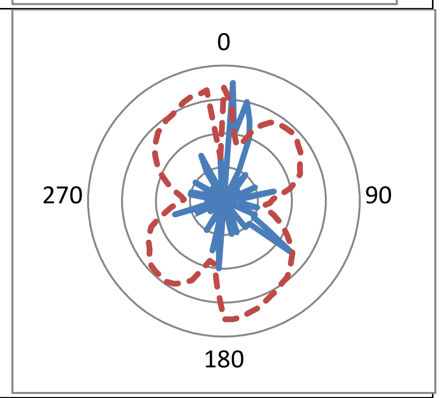
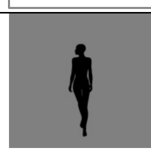
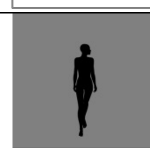
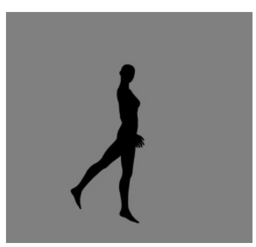
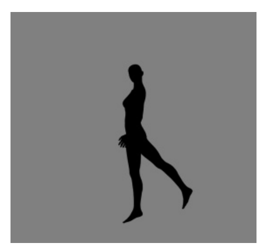
<table>
<tr><td>Figure 22a</td><td>Anti-Clockwise Reversal - Body</td><td>Clockwise Reversal - Body</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td>Statistical Test: Omnibus Hodges-Ajne test for non-uniformity</td><td>
Omnibus Hodges-Ajne
mode = 289.9
p-value = 5.92 e-006
Total # switches: 127</td><td>
Omnibus Hodges-Ajne
mode = 114.1
p-value = 1.03 e-009
Total # switches: 126</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 22 a: Results from experiment 2.3 while foveating on the body

Figure 22b	Anti-Clockwise Reversal - **Head**	Clockwise Reversal - **Head**
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Omnibus Hodges-Ajne test for non-uniformity	 Omnibus Hodges-Ajne mode = 188.9 p-value = 3.70 e-006 Total # switches: 117	 Omnibus Hodges-Ajne mode = 8.8 p-value = 0.02 Total # switches: 115

Degrees: 0 | 90 | 180 | 270

Figure 22 b: a) Results from experiment 2.3 while foveating on the head

22 a & b This experiment yielded a very unexpected result. Our expectation was that subjects would experience directional reversals based on both the position of the head and the body, this was verified in the experimental results below. However, upon debriefing, subjects also reported seeing the head and body moving independently of one another in a non-rigid motion. Furthermore this independence could be so complete that the head would appear to rotate completely around the body exorcist-style with no regard for the natural limits of bodily motion.

This not only shows that there are clearly multiple competing biases within the image, but also that the brain is capable of experiencing multiple simultaneous directional reversals within a single object.

The results show that regardless of the fixation point there are four total points within the rotation where rotation reversals are likely to occur. These reversal points correspond to each of the two profile positions for both the body and the head. When foveating the body the most common direction reversals occurred when the stimulus was in a profile position. As before, the new direction of motion indicates a preference for having the body remain facing towards the subject. Similarly, when foveating the head, subjects tended to see direction reversals when the head was in the profile position, also with the usual preference for seeing the head face forward. Together, these results show a very nice dichotomy between the viewing conditions.

Directional switches seem to occur in all four orientations for each foveation point, but the specific viewing condition result in more reversals occurring in concert with the profile position of the foveated region.

During the experimental debriefing subjects were asked to clearly describe what they experienced when viewing this stimulus. Their responses suggest that the stimulus would often undergo a rigid rotation reversal, similar to all other images shown. However subjects also reported occasional non-rigid motion in which the body or head would change direction independently of the rest of the object. This percept was so strong that at times the head would appear to rotate completely independently of the body.

While the results of these studies are novel, there do seem to be some areas of potential conflict with current research. Take the 2012 work of Pastukhov, Vanau, and Braun. They used a series of three-dimensional stimuli generated via a structure from motion paradigm of moving dots that were attached to the surface of an invisible sphere (Pastukhov, Vanau, & Braun, 2012). While these stimuli are not the exact same type of illusion as the Kinetic Depth Effect they do share some significant similarities. Firstly all of these stimuli are inherently multistable because the dots on the front and back surface of the sphere are interchangeable. And similar to KDE images a depth reversal of the positions of the dots from front-to-back will also result in a concurrent change of direction of rotation of the sphere as a whole.

One of the major findings of this study was that multistable reversals were governed by physical plausibility. However our results from Experiment 2.3 (a human figure with a leg out and head turned (figure 22) clearly showed that the visual system was more than willing to forgo physically plausible motion when the figure's head would spin freely of the body Exorcism-style.

To be fair though none of Pastukhov's stimuli contained the types of high level features that exist in humanoid figures. So there is a strong possibility that both results are correct … but that the high level bias found here was simply stronger than their result. However it does mean that their statement about "multistable reversals being caused by physical plausibility" is overly strong. And not only does this result violate the 'rigidity assumption' of Ullman (Ullman, 1984-A; Ullman, 1984-B; Ullman, 1987) but it also violates any reasonable constraints on plausible non-rigid biological motion.

Experiment Set 3 – The Effect of Linear Image Motion

This set of experiments was designed to look at the effect of the linear image motion across the screen.

As discussed earlier, the linear motion of any particular point in a rotating KDE image will vary sinusoidally over the course of its rotation. The slowing down and instantaneous stop may facilitate a change in perceived direction of rotation. By creating stimuli where the previously explored biases for front-facing and minimum depth reinterpretation are either congruous or incongruous with these instantaneous stops, it is possible to determine how changes in linear motion affect directional switching.

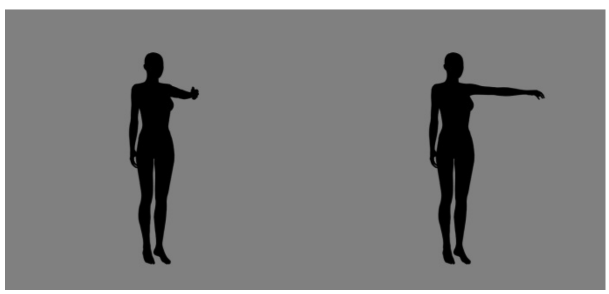


Figure 23: Stimuli for experiment 3

Experiment 3 - Methods

Stimuli

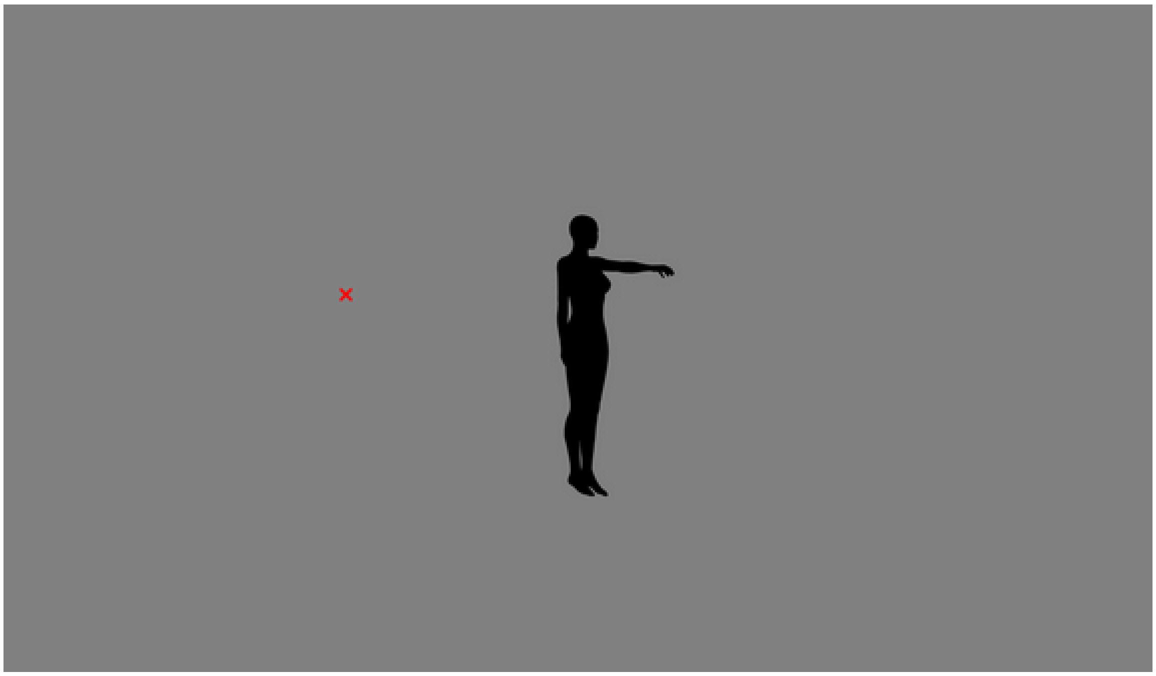


Figure 24: Sample experimental stimulus from experiment 3

There were two silhouettes used for this set of experiments, the humanoid figure with an arm forward and the humanoid figure with an arm outstretched sideways. Each of these objects was used to create a new KDE stimulus which would slow down at one of the four cardinal points of the revolution (profile-left, profile-right, front, & back), come to a very brief stop, and then continue to rotate. This resulted in a total of eight experimental stimuli.

Asymmetric stimuli were used, even though they showed more variability in terms of where directional switches would occur, compared to the left-right symmetric images from experiment set 1. This resulted in stimuli with the following properties:

1) One stimulus where the linear image motion and depth reinterpretation biases were in competition (Humanoid with an arm outstretched sideways). In experiment 2.1 this stimulus was shown to preferentially undergo directional switches in the front-back orientations.
2) A second stimulus where the linear image motion and front-facing biases were in competition (Humanoid with an arm forward). In experiment 2.1 this stimulus was shown to preferentially undergo directional switches in the profile orientations.

Each stimulus consisted of 82 frames, and a single revolution took 5.47 seconds.

To avoid physical distortions due to perspective the all of the images were presented in a frontal orthographic projection.

The background color of the stimuli was a mean grey. The 3D silhouettes were a pure black.

Subjects were seated 26 inches from the monitor. At this distance the monitor distended a viewing angle of 31.2 degrees horizontally and 14.9 degrees vertically. The images subtended between 2.2 to 5.2 degrees of horizontal visual angle depending on the frame, and 8.8 degrees of vertical visual angle.

Subjects

23 undergraduate students at UCSD were presented with a rotating silhouette that utilized the Kinetic Depth Effect to create an ambiguity in the direction of rotation. Subjects were compensated for their time with class credit for their current courses. Two of the subjects' data were not included in the final analysis. These subjects were removed for not following the experimental protocol.

Instructions

Subjects were instructed to indicate the initial direction of motion of the stimulus (clockwise or anti-clockwise as viewed from above) when it first appeared with a keypress [<] or [>]. Thus [>] indicated anti-clockwise rotation as viewed from above, and [<] indicated clockwise rotation as viewed from above. They were told to then press the [<] or [>] key to indicate the new direction of rotation whenever a directional change in motion occurred. They also had the option of pressing the [SPACEBAR] if the image started to do something that could not be described as rotating.

Image Presentation

Stimulus presentation order was randomized for each subject. During the experiment each stimulus was presented in the center of the screen for the duration of each trial block. During a single block the subject would first be allowed to freely view the image for 120 seconds. After that viewing period the screen would temporarily blank out for 1 second as a fixation cross would randomly appear to the left or right of the screen (9 degrees off center). The subjects had been previously instructed to foveate on the fixation cross. One second after the screen blanked out the ambiguous figure would reappear. After each block of Center/Left/Right viewing the subject was given an option to briefly take a break and continue the experiment when they were ready.

Data Collection

The data collected was the current frame being presented when the computer registered the subjects' response indicating they noticed a change. Thus there was a slight lag between the time the subjects noticed a change, decided to press a button, and completed the button press. There was likely a reaction time delay of about 400-500ms before the button press was registered (Kornmeier & Bach, 2012). This translates into a delay of at least7 frames of the stimulus. Thus the peaks in the data presented below occur slightly *after* the subjects experienced a directional switch.

To confirm whether or not subjects were responding slowly due to a reaction time lag a control experiment was conducted which presented subjects with a single KDE image that was presented at two different speeds. The results of this experiment, shown later, confirm that there seems to be a consistent response-time lag which can at least partially account for subject response peaks at non-cardinal points of the rotation. There is also likely a small delay from image processing.

The total run time for each subjects was 48 minutes (6 minutes per stimulus), this allowed for 176 full rotations of each stimulus with a total of 528 full rotations across all stimuli.

At the end of the experiment subjects were asked to debrief the experimenter. This consisted of verbal descriptions of what they saw and questioning as to specific positions where the images seemed to change direction.

Experiment 3 - Results & Conclusions

The results indicate that linear image motion had only a minimal effect on directional switching and that overall linear image motion was NOT strong enough to overcome the biases found in experiment set 2 related to seeing the image face forward and minimum 3D object reinterpretation. Instead the stimuli tended to switch as predicted from the second set of experiments.

The only exception is that when the motion stop-point corresponded with the position of expected peak of directional switches from experiment 2. In those cases there seems to be a slight increase in directional switches at the expected peak as compared to the cases where there was no overlap. This difference can be seen in the length of the mean resultant vector. As discussed earlier, the mean resultant vector has a length between [0,1]. The closer that the mean resultant vector is to 1, the more highly consolidated the data-set around a single point in the rotation. There were four graphs where the expected peak for directional switches aligned with the stop-point of the stimulus (25c-anti-clockwise, 25d-clockwise, 26a-anti-clockwise, 26b-clockwise). Those four cases represented the four highest mean resultant vectors for the entire data set (0.417, 0.530, 0.473, 0.490). The rest of the mean resultant vectors were between 0.231 - 0.304 with a mean value of 0.266. This could suggest that linear motion has an effect on how we perceive KDE images, but that it is simply a weaker effect than those found earlier.

However it is possible that the slightly enhanced directional switching at the predicted positions is not due to an effect of linear image motion, but rather the fact that the true 3D object motion is also coming to a standstill and thus when a directional switch occurs the visual system does not need to reassign a new, and opposite, angular velocity to the rotating object was required for the previous experiments.

Note that the graphs used in these experiments differ from those in the rest of the document. This is due to the fact that the rotational velocity of the stimuli is not constant. As before a full rotation around the graph indicates a full rotation of the image shown to the subjects. However, because the images slowed down and came to a stop, there is not a direct correlation between the cardinal points of the graph and the image facing directly Left, Right, Forward, or Backward. Therefore in this experiment set the 0, 90, 180, and 270 degree points on the graph do not correspond to a particular position of the 3D silhouette, but rather to the amount of time that the image has been rotating. To avoid confusion these degree numbers have been replaced with percentages which indicate a percentage of the total revolution time (0%, 25%, 50%, 75%). Since each rotation slows down at a particular point, the stimuli are still periodic within a single experiment each specific point of the stimulus rotation will always correspond to the same point on the circular graph. However this correspondence does not hold true across different experimental images. Finally a red arrow has been added to each graph which indicates where the stimulus came to a brief stop. To make it easier to interpret the data, and for ease of comparison across stimuli, the data has been reorganized so that all of the images stopped at the same point relative to one another on the graph (0%) as indicated by the red arrow.

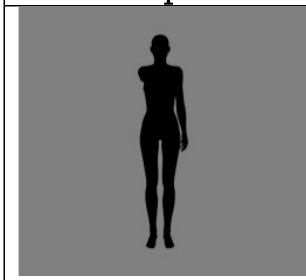
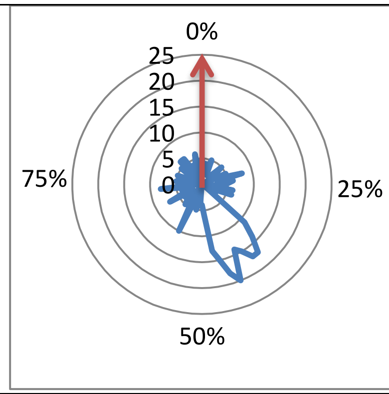
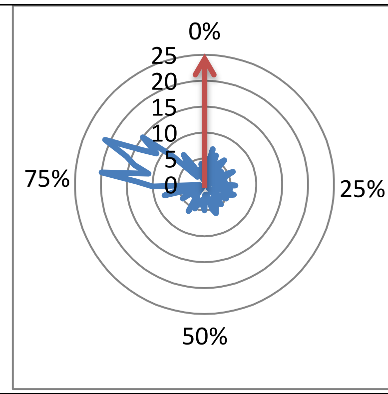
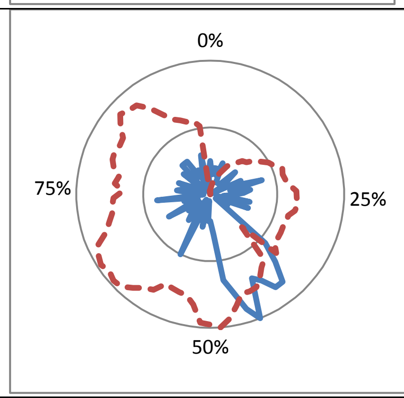
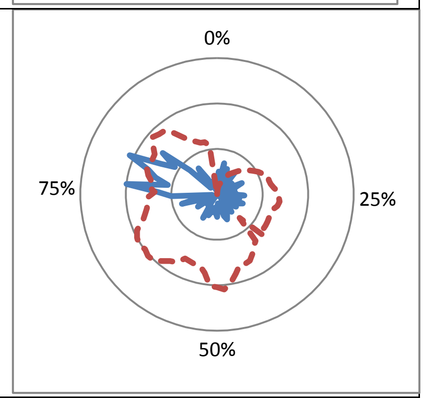
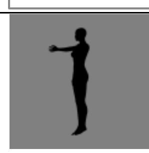
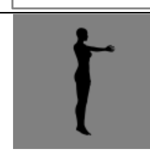
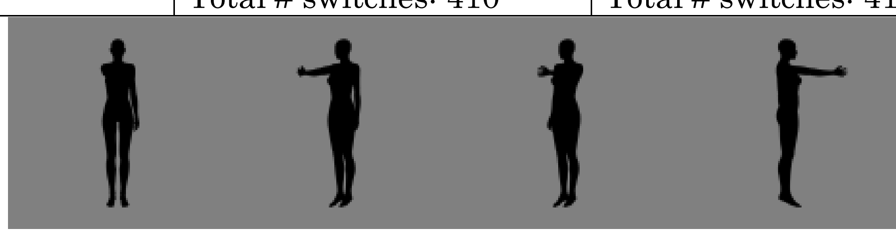
<table>
<tr><td>Figure 25a
Arm Face-Forward – Left Stop</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test: Rayleigh test for non-uniformity</td><td>
mode = 162.5
mean = 159.60
r-length = 0.304
p-value = 5.55 e-017
Total # switches: 410</td><td>
mode = 298.7
mean = 292.32
r-length = 0.256
p-value = 2.13 e-012
Total # switches: 411</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 25 a: Humanoid with one arm forward with a stopping of the rotation at each of the cardinal points

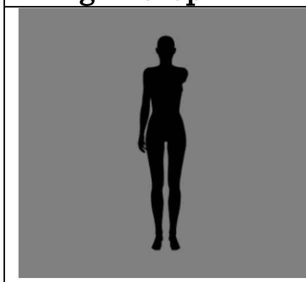
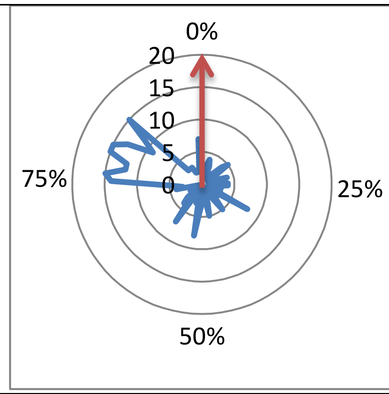
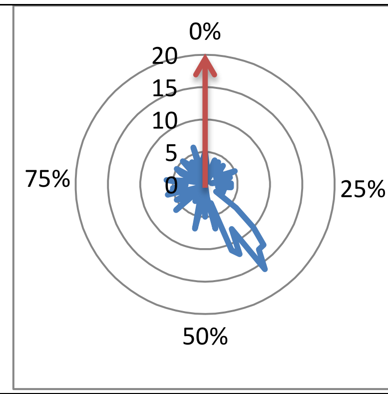
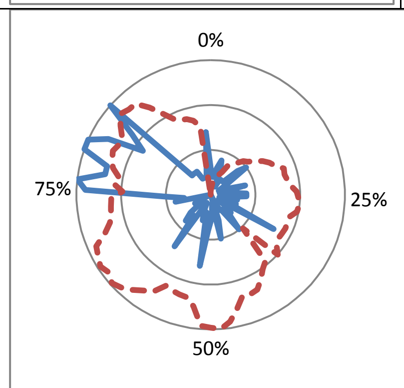
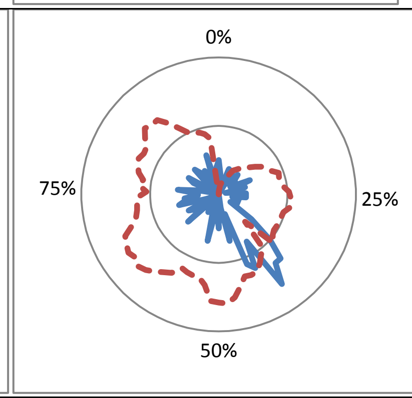
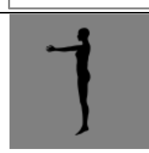
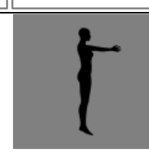
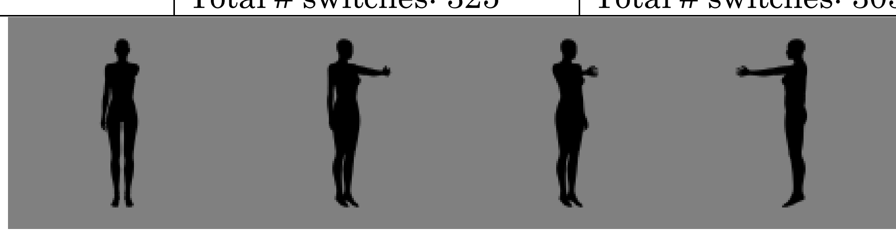
<table>
<tr><td>Figure 25b
Arm Face-Forward
– Right Stop</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test:
Rayleigh test for non-uniformity</td><td>
mode = 281.1
mean = 294.75
r-length = 0.304
p-value = 2.00 e-014
Total # switches: 325</td><td>
mode = 149.3
mean = 154.09
r-length = 0.231
p-value = 8.96 e-008
Total # switches: 305</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 25 b: Humanoid with one arm forward with a stopping of the rotation at each of the cardinal points

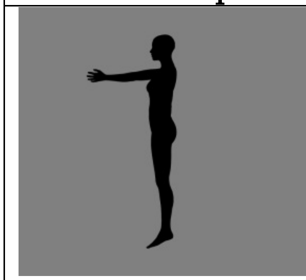
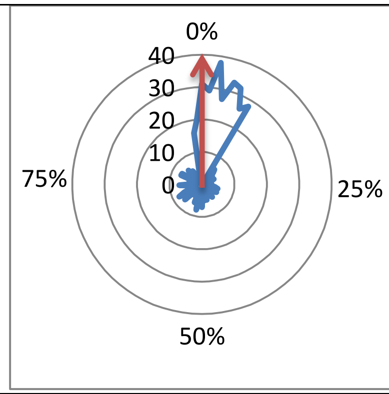
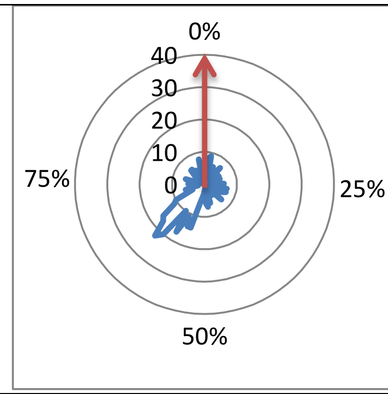
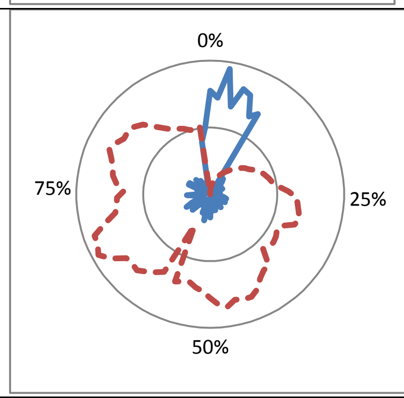
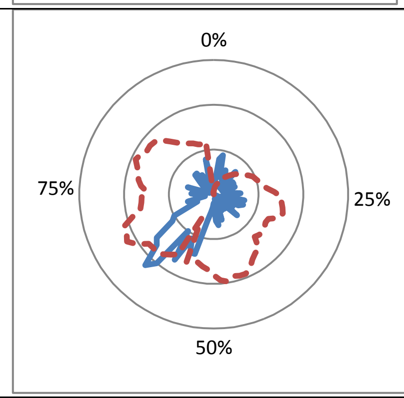
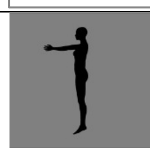
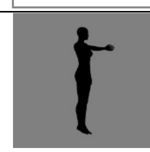
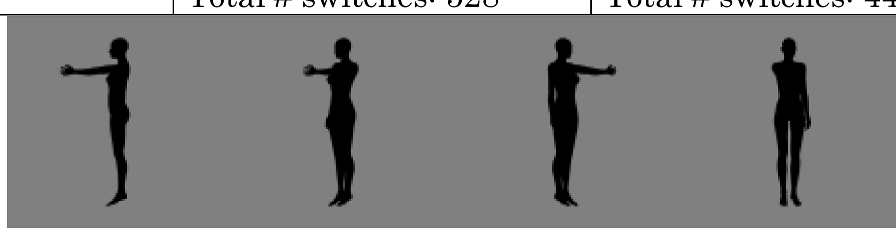
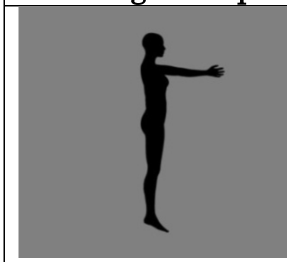
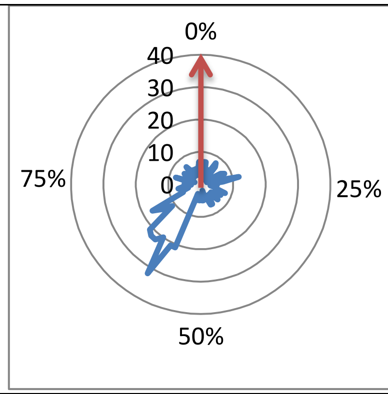
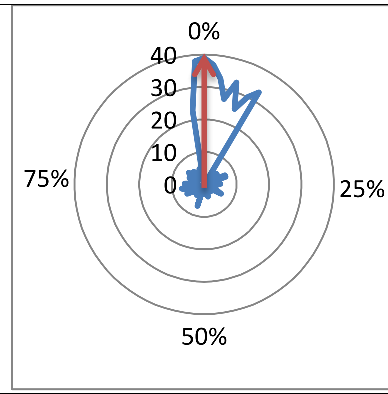
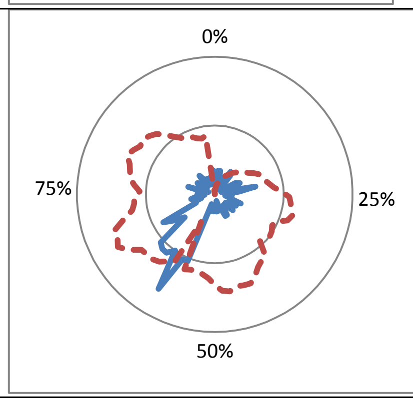
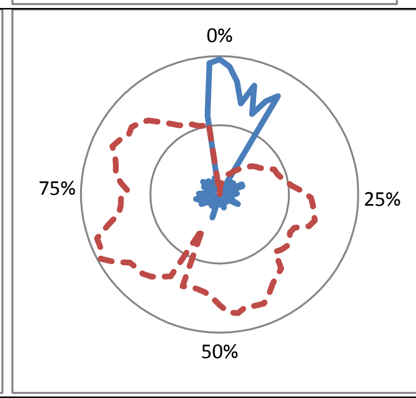
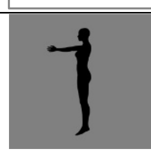
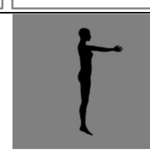
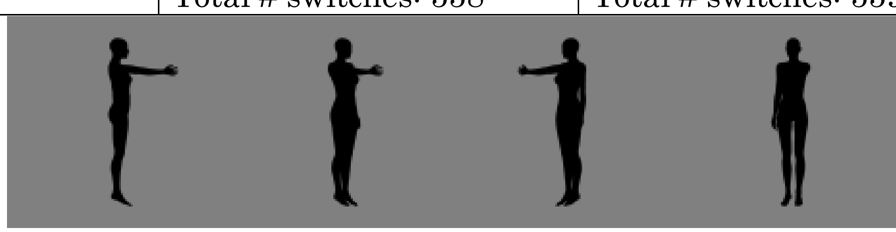
Figure 25c Arm Forward – Face Left Stop	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode = 13.6 mean = 12.56 r-length = 0.417 p-value = 4.54 e-042 Total # switches: 528	 mode = 228.6 mean = 219.57 r-length = 0.244 p-value = 1.47 e-012 Total # switches: 445
 Degrees: 0 | 90 | 180 | 270		

Figure 25 c: Humanoid with one arm forward with a stopping of the rotation at each of the cardinal points

Figure 25d Arm Forward – Face Right Stop	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode = 215.5 mean = 221.71 r-length = 0.256 p-value = 1.20 e-016 Total # switches: 558	 mode = 4.8 mean = 11.3 r-length = 0.530 p-value = 1.02 e-073 Total # switches: 559



Degrees: 0 | 90 | 180 | 270

Figure 25 d: Humanoid with one arm forward with a stopping of the rotation at each of the cardinal points

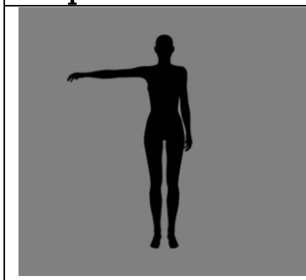
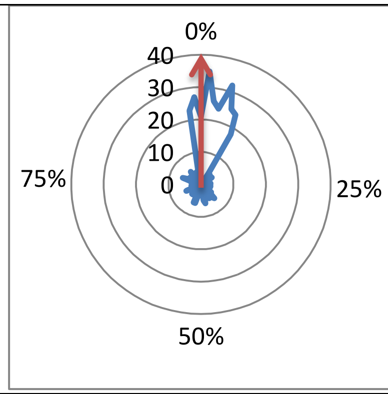
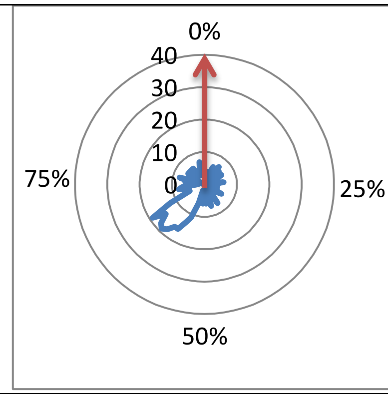
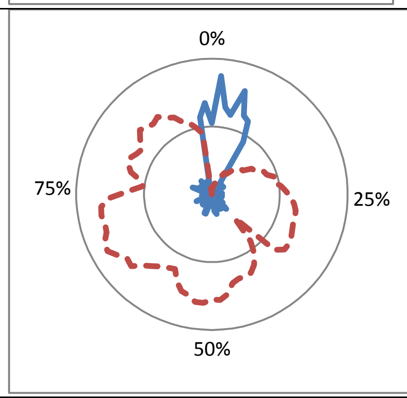
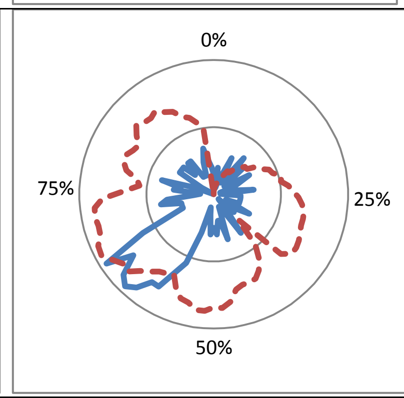
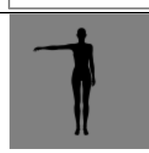
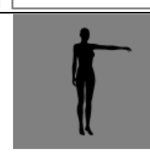
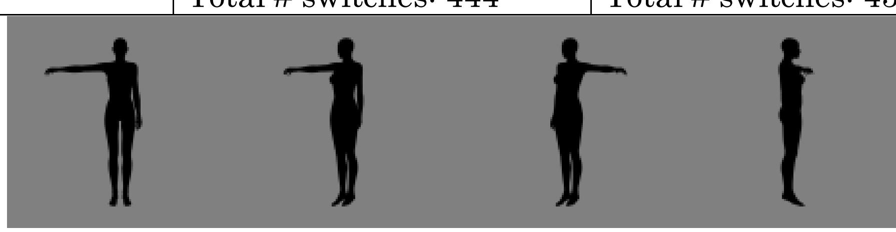
<table>
<tr><td>Figure 26a
Arm Out – Left
Stop</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test:
Rayleigh test for non-uniformity</td><td>
mode = 8.8
mean = 14.25
r-length = 0.473
p-value = 1.84 e-046
Total # switches: 444</td><td>
mode = 228.4
mean = 232.53
r-length = 0.304
p-value = 3.37 e-019
Total # switches: 450</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 26 a: Humanoid with one arm outstretched to the side with a stopping of the rotation at each of the cardinal points

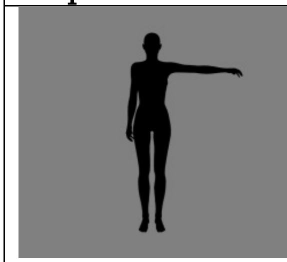
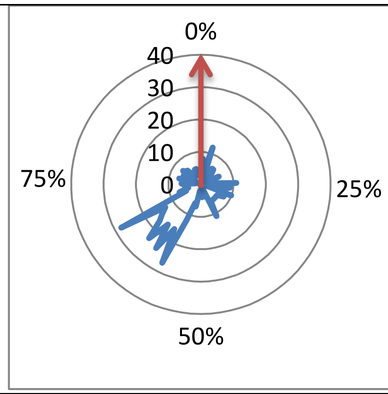
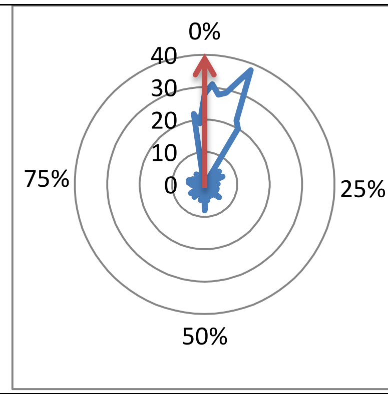
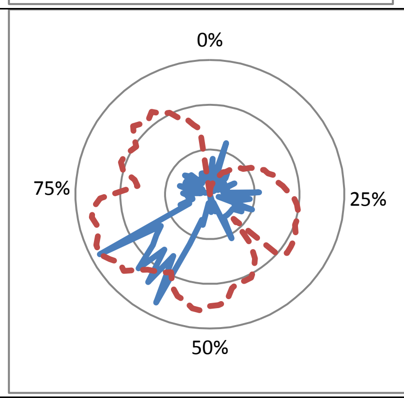
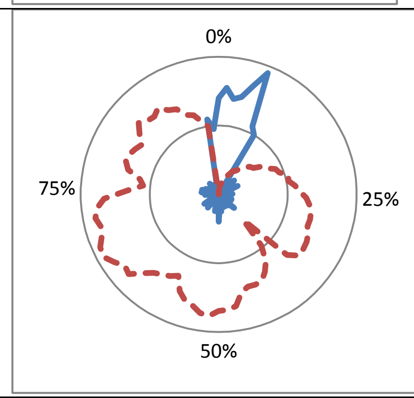
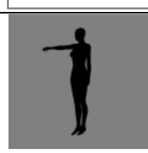
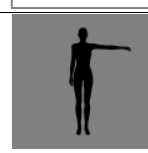
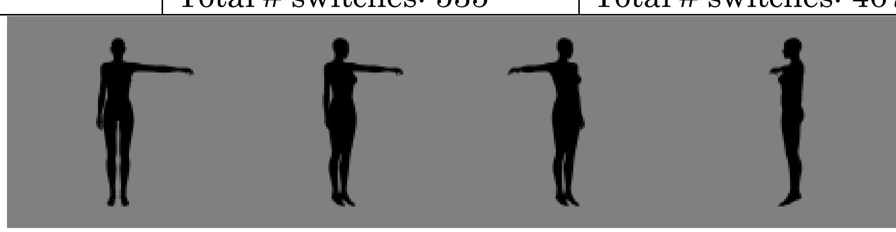
<table>
<tr><td>Figure 26b
Arm Out – Right
Stop</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test:
Rayleigh test for non-uniformity</td><td>
mode = 246.0
mean = 219.29
r-length = 0.268
p-value = 4.76 e-017
Total # switches: 535</td><td>
mode = 26.4
mean = 19.99
r-length = 0.490
p-value = 8.44 e-054
Total # switches: 467</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 26 b: Humanoid with one arm outstretched to the side with a stopping of the rotation at each of the cardinal points

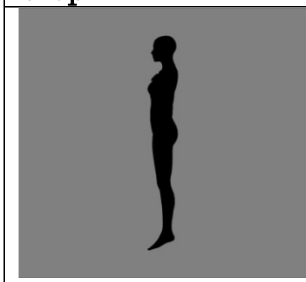
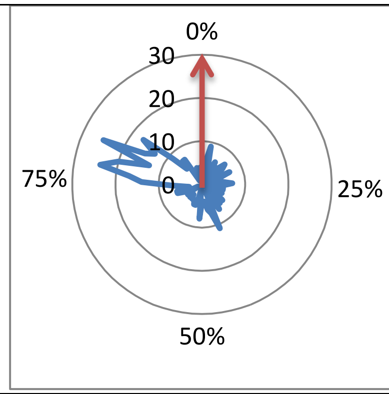
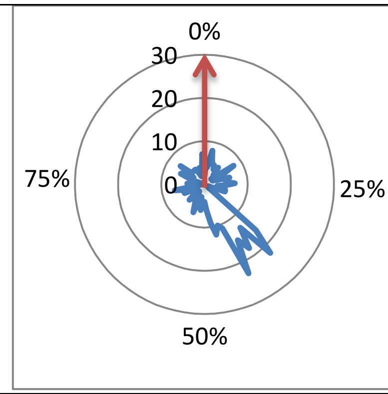
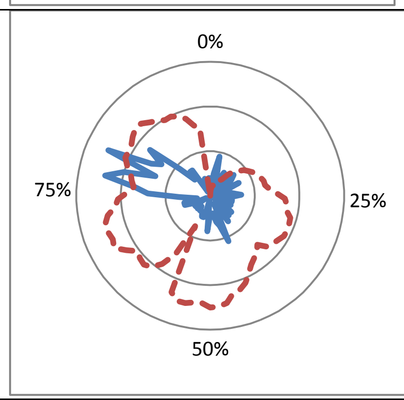
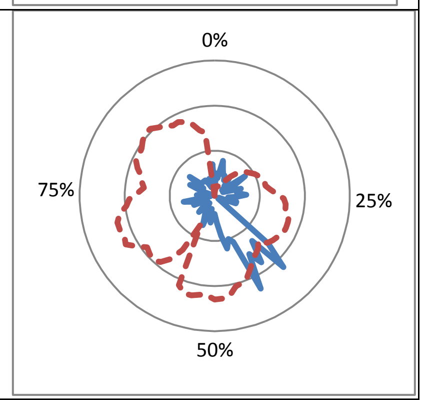
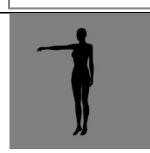
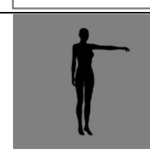
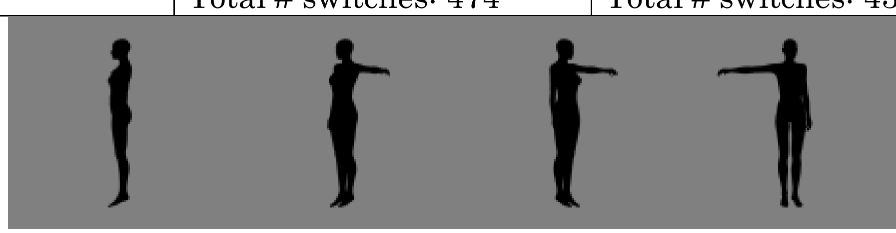
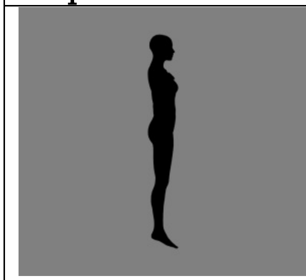
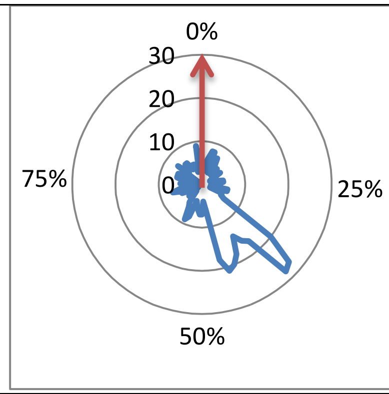
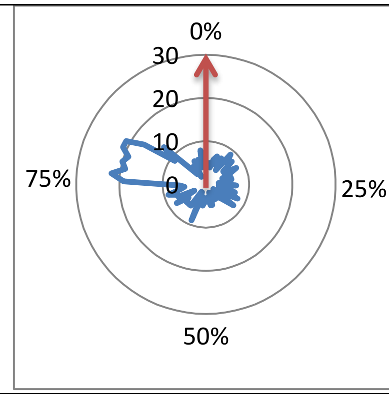
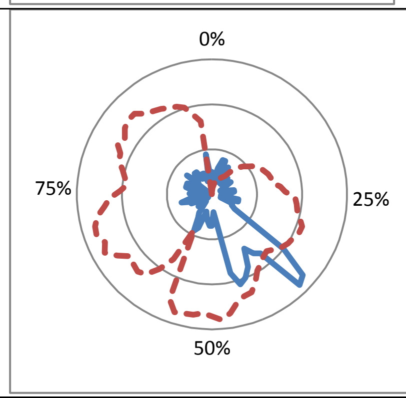
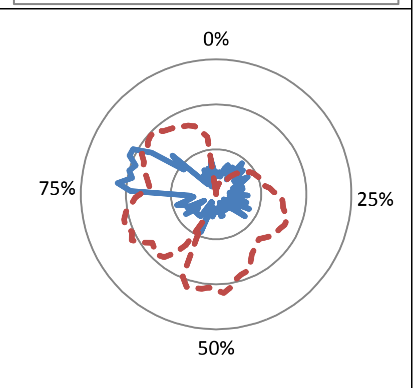
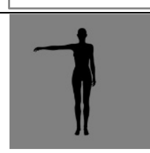
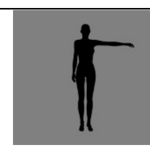
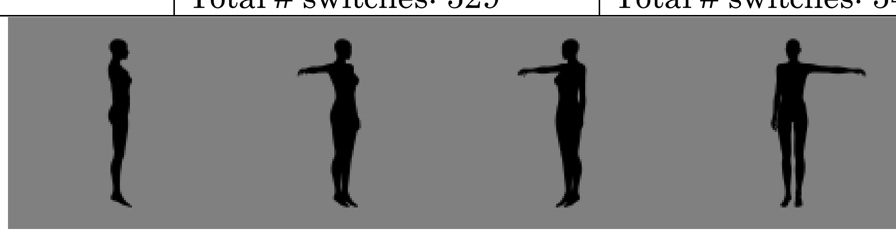
<table>
<tr><td>Figure 26c
Arm Out – Left
Stop</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td>Statistical Test:
Rayleigh test for non-uniformity</td><td>
mode = 298.9
mean = 290.95
r-length = 0.244
p-value = 5.31 e-013
Total # switches: 474</td><td>
mode = 158.4
mean = 144.88
r-length = 0.244
p-value = 1.18 e-011
Total # switches: 435</td></tr>
<tr><td colspan="3">
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 26 c: Humanoid with one arm outstretched to the side with a stopping of the rotation at each of the cardinal points

Figure 26d Arm Out – Left Stop	Anti-Clockwise Reversal	Clockwise Reversal
		
Frame-by-Frame analysis of linear image motion		
Statistical Test: Rayleigh test for non-uniformity	 mode = 140.8 mean = 152.14 r-length = 0.280 p-value = 1.31 e-019 Total # switches: 529	 mode = 281.4 mean = 301.3 r-length = 0.256 p-value = 3.67 e-016 Total # switches: 547



Degrees: 0 | 90 | 180 | 270

Figure 26 d: Humanoid with one arm outstretched to the side with a stopping of the rotation at each of the cardinal points

Experiment Set 4 – Stimulus Complexity

Although it is clear that high-level biases can affect our perception of multistable stimuli, it is impossible to have a complete picture without a better understanding of low-level biases that may exist. To explore possible low-level effects on the KDE flat triangles were used in Experiment Set 1. This supported the data from the humanoid and robot stimuli in suggesting that directional switches were more likely in poses which did not require a depth reversal and where the linear image motion was at its lowest. But high level biases experienced by subjects seemed quite strong, to the point where any lower-level effects, such as those due to linear image motion, may have been lost. To study the possibility of low level biases more directly, a set of stimuli was created using simple geometric shapes intended to minimize any high-level biases. In particular manipulation of image complexity seemed like a good starting point with which to explore low-level factors in perceptual rotation reversals.

The manipulations made to image complexity here deal with the addition or removal of connecting lines (paired and unpaired), the removal of high information-value corners, and the total number of objects on screen. Each of these changes alters the way in which reinterpretation of the image must occur, but does so without introducing high-level biases that seem to exist in humanoid figures. Instead each change alters something about the nature of the 3D object and the amount of information processing that must occur during a perceptual switch.

Experiment 4 - Methods

Stimuli

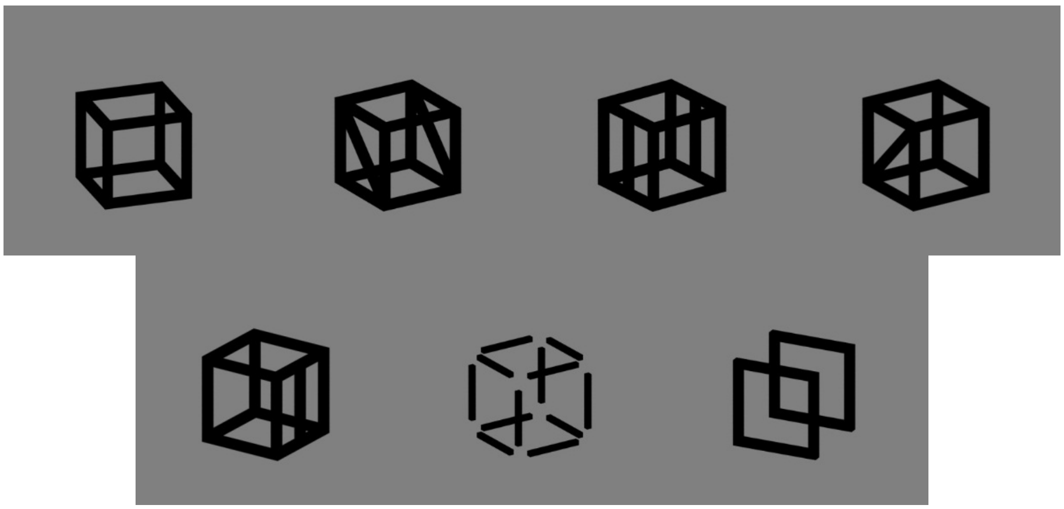


Figure 27: Stimuli used in experiment 4

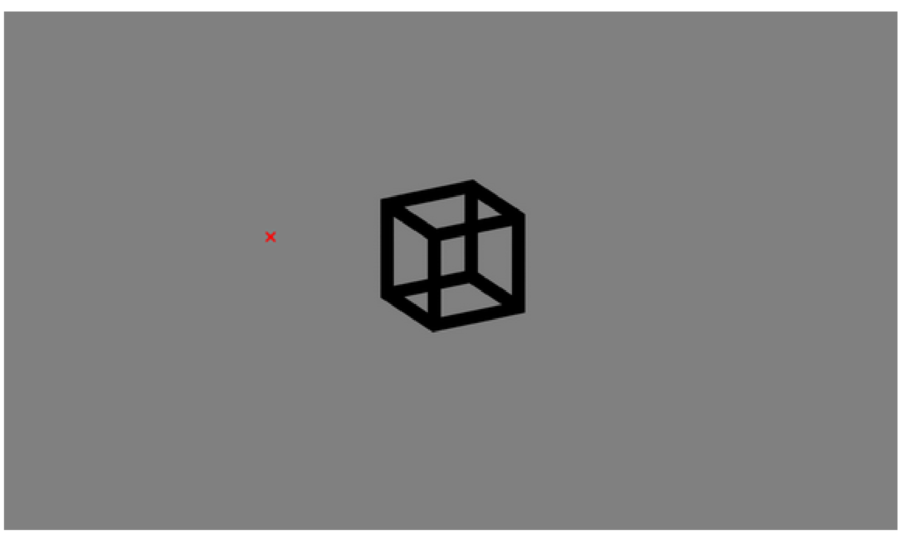


Figure 28: Sample stimulus from experiment 4

As in Experiment Set 1, DAZ Studio 3 was used to create the stimuli.

The stimulus configurations above were chosen because they are all similar enough to a standard Necker cube to provide for a reasonable comparison and yet different enough to allow for potentially large differences in how they are treated by the visual system.

1) Figure 29 is a standard Necker cube which can serve as a basic control condition.
2) Figure 30 is a standard Necker cube with two additional connection edges that go diagonally across opposite faces. A perceptual switch of this object requires that the direction of the angled edges reverse.
3) Figure 31 is a standard Necker cube with two additional connection edges that go vertically across opposite faces. This is meant to serve as a comparison to Figure 30 as they both have the same number of additional lines, but a perceptual switch of this object does not require reassigning the direction the added edges.
4) Figure 32 is also meant as a comparison to figure 30 as it is identical, but with only a single additional edge. A perceptual switch of this object would require a shifting of the location of the additional edge from front to back.
5) Figure 33 is meant as a comparison to figure 31 as it is identical, but with only a single additional edge. A perceptual switch of this object

would require a shifting of the location of the additional edge from front to back

6) Figure 34 is designed to test how the visual system will respond when high-information portions of the object have been eliminated. The lack of corners does not prevent subjects from perceiving this object as a cube, but the visual system may not be able to treat unconnected lines in the same way that it treats a solid object.
7) Figure 35 is designed to test how the visual system deals with two separate objects. In some ways this is simpler than a standard Necker cube, as it has four fewer lines, and in other ways it could be considered more complex since it is perceived as two separate objects.

Each of the seven stimuli consisted of 82 views (frames) representing a full revolution of a 3D object. This resulted in an angular rotation between successive frames of 4.39 degrees. The frames were presented at a rate of 10 frames per second which resulted in one full revolution of the object every 8.2 seconds. These stimuli were intentionally made to rotate slower than previous stimuli due to the fact that every 90 degrees of rotation was a possible position for direction reversal.

To avoid physical distortions due to perspective the all of the images for this experiment were presented in an isometric projection. Unlike the previous experiments it was impossible to use a frontal view because the cubes need to be seen at an angle (from above or below) otherwise they just look like 2-dimensional boxes. Thus they were in an isometric view with an angle of elevation of 35.264 degrees.

The background color of the stimuli was a mean grey. The 3D silhouettes were a pure black.

Subjects were seated 26 inches from the monitor. At this distance the monitor distended a viewing angle of 31.2 degrees horizontally and 14.9 degrees vertically. The images distended between 5.1 to 7.3 degrees of horizontal visual angle depending on the frame, and 6.7 to 7.7 degrees of vertical visual angle depending on the frame.

Subjects

23 undergraduate students at UCSD were presented with a rotating silhouette that utilized the Kinetic Depth Effect to create an ambiguity in the direction of rotation. Subjects were compensated for their time with class credit for their current courses. Two of the subjects' data were not included in the final analysis. These subjects were removed for not following the experimental protocol.

Instructions

Subjects were instructed to indicate the initial direction of motion of the stimulus (clockwise or anti-clockwise as viewed from above) when it first appeared with a keypress [<] or [>]. Thus [>] indicated anti-clockwise rotation as viewed from above, and [<] indicated clockwise rotation as viewed from above. They were told to then press the [<] or [>] key to indicate the new direction of rotation whenever a directional change in motion occurred. They also had the option of pressing the [SPACEBAR] if the image started to do something that could not be described as rotating.

Image Presentation

Stimulus presentation order was randomized for each subject. During the experiment each stimulus was presented in the center of the screen for the 120 second duration of each trial block. During a single block there were three viewing conditions that each last 120 seconds; free-viewing, left-fixation, and right-fixation. The subject would first be allowed to freely view the image for 120 seconds. After that viewing period the screen would temporarily blank out for 1 second as a fixation cross would randomly appear to the left or right of the screen (9 degrees off center). The subjects had been previously instructed to foveate on the fixation cross. One second after the screen blanked out the ambiguous figure would reappear for 120 seconds. Finally the screen would once again blank out for 1 second as the fixation cross moved to the final viewing position. One second after the screen blanked out the ambiguous figure would reappear for 120 seconds. After each block of Center/Left/Right viewing the subject was given an option to briefly take a break and continue the experiment when they were ready by pressing any key.

The complete experiment consisted of three viewing conditions of each of the seven stimuli blocks. Stimulus block presentation order was randomized for each subject and data was collected for the full duration of each block, but not during the break periods.

Data Collection

The data collected was the current frame being presented when the computer registered the subjects' response indicating they noticed a change. Thus there was a slight lag between the time the subjects noticed a change, decided to press a button, and completed the button press. There was likely a reaction time delay of about 400-500ms before the button was pressed (Kornmeier & Bach, 2012). This translates into a delay of 4-5 frames of the stimulus. Thus the peaks in the data presented below occur slightly *after* the subjects experienced a directional switch.

To confirm whether or not subjects were responding slowly due to a reaction time lag a control experiment was conducted which presented subjects with a single KDE image that was presented at two different speeds. The results of this experiment, shown later, confirm that there seems to be a consistent response-time lag which can at least partially account for subject response peaks at non-cardinal points of the rotation. There is also likely a small delay from image processing.

The total run time for each subject was 42 minutes (6 minutes per stimulus) for the seven stimuli presented. This allowed for 154 full rotations of each stimulus with a total of 462 full rotations across all stimuli.

At the end of the experiment subjects were asked to debrief the experimenter. This consisted of verbal descriptions of what they saw and questioning as to specific positions where the images seemed to change direction.

As usual, the stimuli must always undergo a mirror depth reversal at any switch in rotation direction. The two possible perceived orientations of a Necker cube will be referred to as 'viewed from above', when the uppermost vertices in the projected image are seen as the back edge of the top face, and 'viewed from below', when the uppermost vertices in the projected image are seen as the front edge of the top face. As the cube rotates, each top face vertex follows the same elliptical trajectory in the image, and the upper and lower pair of vertices change places as the edges they define move from front to back. In these experiments the vertices moved in an anti-clockwise direction. For anticlockwise rotation of the vertices in the image, the perceived 3D rotation is necessarily anti-clockwise from above if the cube is in fact perceived as viewed from above (upper vertices in back), and necessarily clockwise if the cube is perceived as viewed from below. It has previously been shown, that when viewing ambiguous stimuli, the visual system has a preference for perceiving the stimulus as being viewed from above (Troje & McAdam, 2010; this is not unexpected since this view is the only one that arises naturally when intervening faces of the cube are opaque. A viewing from above bias is necessarily accompanied by a bias to anti-clockwise direction of rotation with these stimuli, although if the image sequence were played backwards, the association between the perceived orientation of the cube and its perceived direction of 3D rotation would be reversed.

Experiment 4 - Results & Conclusions

Analysis of Peak Positions for Rotation Reversals

For all of the stimuli in this experiment the data indicating the positions at which directional switches were most likely to occur are largely inconclusive. Some of the stimuli, such as figure 33 (Necker Cube with a single additional vertical line) begin to suggest possible preferred switching positions. But the positions of these possible preference points are inconsistent across all of these stimuli and there is significantly more noise in the data. Compared to the quality of the data collected from previous experiments it would be unwise to suggest that these stimuli do indeed have preferred direction reversal positions.

One possible reason for this noisy data could come from the stimulus itself. Necker Cubes are one of the most prototypical multistable illusions and, when viewed, subjects do not need a rotating KDE silhouette to experience a perceptual change. Thus it could easily be the case that the noise in the data is due to a high frequency of perceptual reversals of the Necker Cube that are unrelated to the position within its rotation. In fact, the orientations in which a standard Necker Cube is most characteristically multistable are those in which the subject can easily see both the front and back of the cube, not the orientations where the front and back occlude each other, producing an image reduced to 2 vertically displaced rectangles (Figure 35), and any impression of depth is weak. A similar suggestion was made by Ullman (1979) when he stated that in cases where information about the structure of a 3D object was available from static cues, those cues may begin to affect how the 3D structure is perceived. The results of the previous experiments may be diagnostic of when rotation reversals will occur in a typical KDE stimulus. But in this case the promiscuous depth reversals induced by a Necker Cube may be confounding the data, thus making any systematic variation of rotation reversals with position difficult to discern.

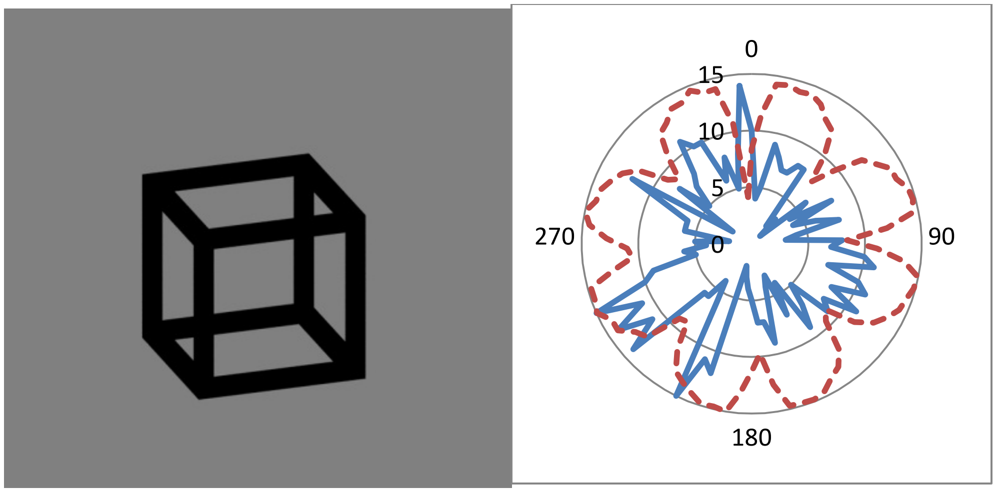


Figure 29: Basic Necker Cube (Total Reversals: 597 in 43.9 revolutions)

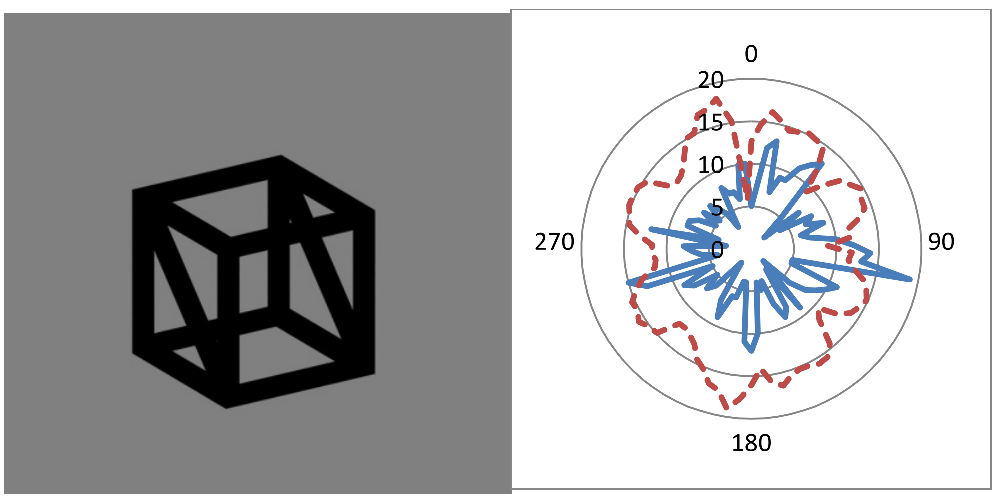


Figure 30: A Necker Cube with two additional diagonal connections (Total Reversals: 621 in 43.9 revolutions)

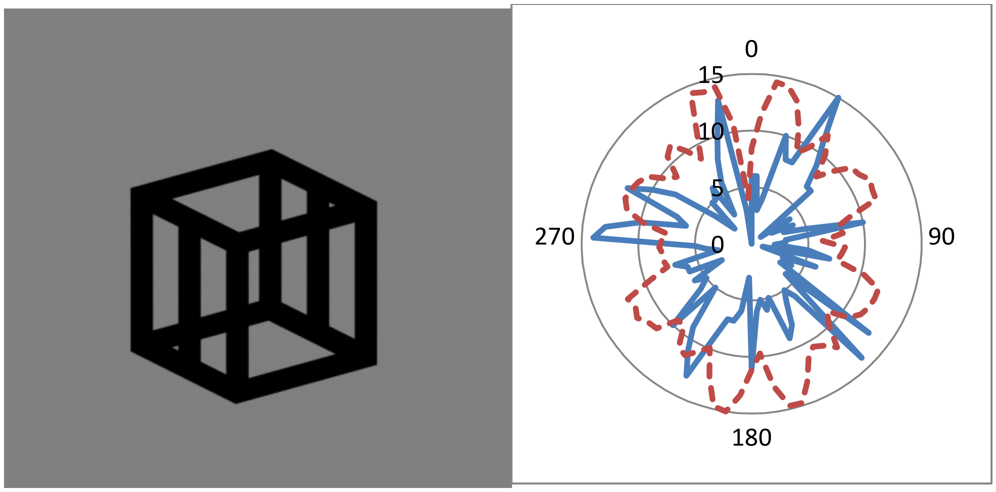


Figure 31: A Necker Cube with two additional vertical connections (Total Reversals: 534 in 43.9 revolutions)

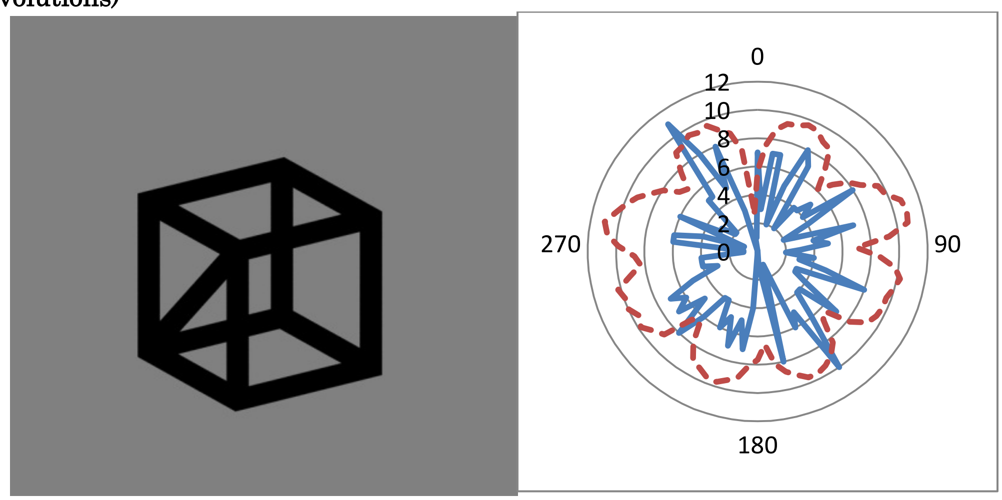


Figure 32: A Necker cube with a single additional diagonal connection (Total Reversals: 372 in 43.9 revolutions) [the asymmetrical line was occluded at 0 and 180]

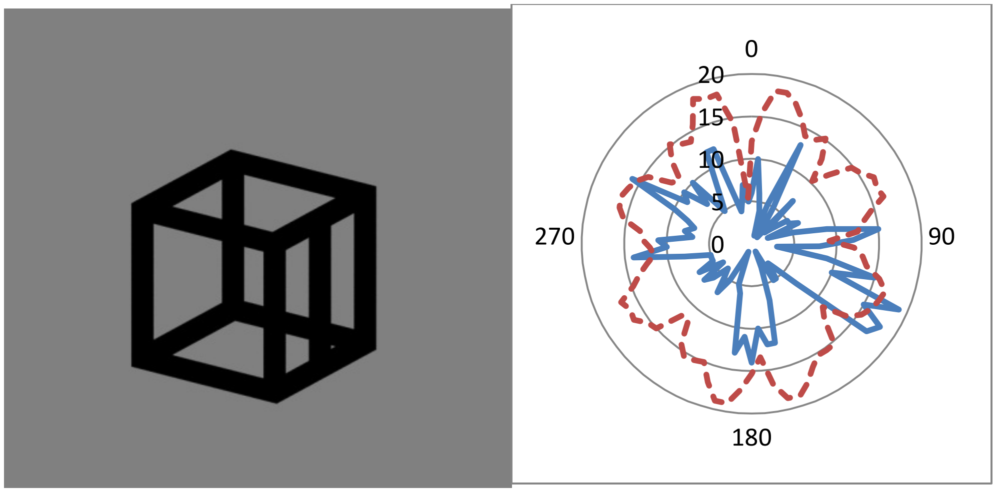


Figure 33: A Necker cube with a single additional vertical connection (Total Reversals: 618 in 43.9 revolutions) [the asymmetrical line was occluded at 90 and 270]

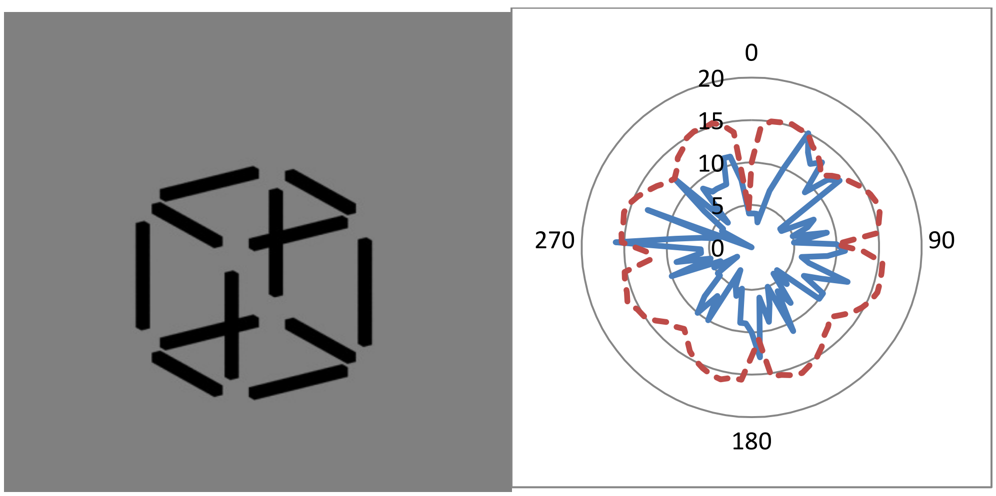


Figure 34: A Necker cube with missing corner connections (Total Reversals: 624 in 43.9 revolutions)

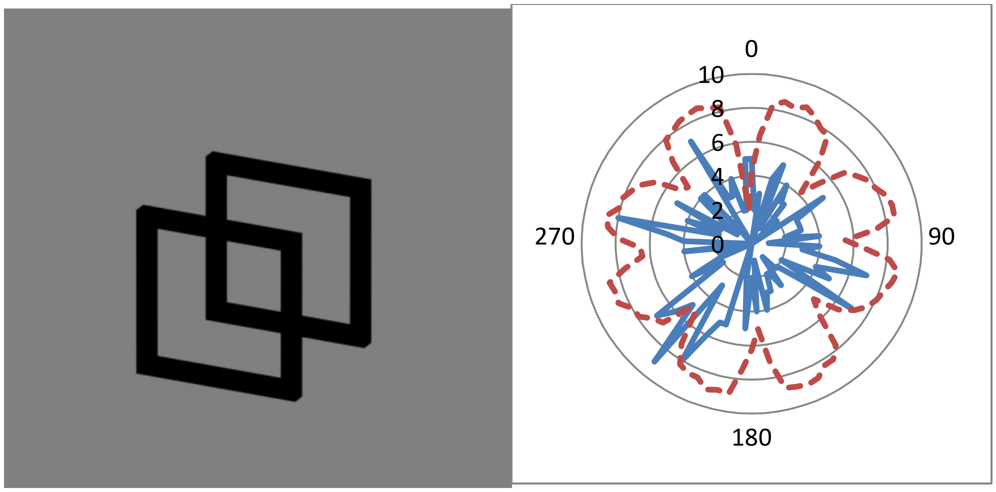


Figure 35: Two squares, which if linked, would form a Necker cube (Total Reversals: 270 in 43.9 revolutions)

Analysis of Total Number of Reversals

Although it might be impossible to determine if position plays a role in rotation reversals of these stimuli it is still possible to conduct other analyses.

By analyzing the total number of direction reversals across all stimuli it is possible to see that two stimuli stand out as having far fewer total reversals (Figure 32 & Figure 35).

Table 6: Comparison of reversals across stimuli

Stimulus Figure	Fig. 29 Basic Necker	Fig. 30 Two Diagonals	Fig. 31 Two Verticals	Fig. 32 One Diagonal	Fig. 33 One Vertical	Fig. 34 No Corners	Fig. 35 Two Squares
# Reversals	**597**	**621**	**534**	**372**	**618**	**624**	**270**

The case of Figure 35 (Two Squares):

This configuration of the Necker cube was chosen to see if there would be a major change in the way that perceptual reversals occurred when there were two separate objects rather than one.

Subjects reported far fewer reversals compared to the other stimuli (more than 50% less than the average across all other stimuli). This could be due to the fact that there are two objects that the visual system must update when a reversal occurs. If each object is treated as separate and unique by the visual system, then twice as much neural processing may be needed for a perceptual switch to occur. Additionally there was no clear pattern of reversals occurring at any of the key points within the rotation of the stimulus indicating that the poses which portrayed a edge-view or a frontal view of each square was unimportant. Ultimately this stimulus is inherently different from the rest and the reduced number of directional reversals may be attributable to that difference.

The case of Figure 32 (One Diagonal) is less clear. If the addition of a diagonal, rather than vertical, line were an important factor then it might be expected to switch similarly to the cube that had two diagonal lines added. On the other hand if an asymmetry in the stimulus were the relevant factor then it would be expected to switch in a similar fashion to the cube with a single vertical line added. Unfortunately neither of these comparisons conforms to predictions.

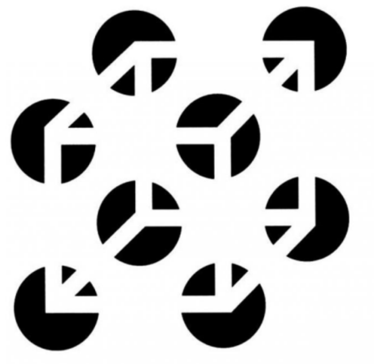


Figure 36: Kanizsa Necker Cube

Finally, it worth examining the Necker Cube with no corners (Figure 34). This cube is of particular interest because corners contain a wealth of information about the orientation of a cube. It is in the corners where the angles be between the connection lines are most visible and in fact a Necker Cube can be defined solely by its corners as was done in the Kanizsa Necker Cube shown in Figure 36 (Kanizsa, 1955). One might expect that fewer reversals would occur for this stimulus given the lack of corner information, but it switches just as readily as the rest of modified Necker Cubes despite the lack of corners. This suggests that the extra information contained in the corners is not necessary for a directional switch to occur.

Viewing From Above Bias

Table 7: Percentage of extra time spent in the anti-clockwise state

Stimulus	% Anti-Clockwise Preference
Basic Cube	14.2
Two Diagonals	19.4
Two Verticals	16.5
One Diagonal	0
One Vertical	0.5
Unconnected	16.2
Two Squares	11.0
Total	**11.1**

As noted earlier, previous research has suggested that a 'viewing from above' bias may alter perception of ambiguous figures (Sundareswara & Schrater, 2008; Troje & McAdam, 2010). The Necker Cubes in this experiment were oriented isometrically such that their two percepts were both accessible to the viewer. The nature of the stimuli was such that rotation direction and the percept of viewing from above or below were tied together and would switch together (in this case a viewing from above percept would correspond to an anti-clockwise rotation). Clearly therefore a viewing from above bias would lead to a preference for anticlockwise rotation.

An analysis of the raw data did in fact show that subjects spent a majority of the time viewing the stimuli in an anti-clockwise, viewing from above, state. The data presented in Table 7 show the anti-clockwise preference for subjects broken down by stimulus. Overall subjects perceived the stimuli in the anti-clockwise state 11.1% more of the time than in the clockwise state.

Interestingly however, neither of the two asymmetric stimuli (Figure 32- One Diagonal, Figure 33 - One Vertical) exhibited this viewing from above bias. The lack of bias is not associated with an unusual overall rate of switching in the case of Figure 33; though Figure 32 does switch far less than most of the other stimuli but this might be expected to enhance rather than diminish a directional bias, so the reduced directional bias for these stimuli does not have an obvious basis..

Hemispheric Difference

Table 8: Hemispheric difference results

	# LVF Switches	# RVF Switches	% Left Visual Field Switches
Basic Cube	207	243	54.0%
Two Diagonals	243	276	53.2%
Two Verticals	198	237	54.5%
One Diagonal	129	171	57.0%
One Vertical	219	282	56.3%
Unconnected	198	246	55.4%
Two Squares	96	90	48.4%
Total	**1290**	**1545**	**54.5%**

To confirm that the above data were not the result of just a few subjects' data biasing the results the following additional statistical analysis was done. First all of the data was analyzed on a per-subject / per-condition basis. Second any conditions in which there were not at least 15 perceptual switches were eliminated. This prevented very low-data subjects from having high percentage switch rates in a single hemi-sphere when that result could simply be due to chance. Then we computed (# of switches in RVF)/(Total # switches) to determine the percentage of perceptual switches happening in the RVF. This percentage was subtracted by 50% and then a per-subject t-test was conducted with a null hypothesis of equal switching in both hemispheres (result of $p<0.0073$). Overall this shows a minor, but consistent tendency for subjects to experience more perceptual switches when viewing these stimuli with the left hemisphere of the brain.

A similar analysis was conducted on the data from Experiment Set 1 & 2 but no significant result was found. It is unclear why only the Necker Cube-like stimuli produced this hemispheric asymmetry. Further experiments will need to be conducted to explore these results.

A final note about these data is that they displayed a hemispheric difference for all stimuli (except for Fig 35 – Two Squares). Significantly more rotational reversals occurred when the stimuli were viewed in the Right Visual Field (RVF) and thus processed in the Left Hemisphere. A similar result was first reported in 2009 (Azoulai & MacLeod, 2009), however subsequent experiments, including those presented earlier in this document had failed to replicate the result.

These previous experiments consisted of three relevant stimuli. The KDE ballerina popularized by Nobuyuki Kayahara, an isometric Necker cube, and an isometric pyramid. Subjects reported all three of these stimuli undergoing more perceptual reversals when presented in the RVF. The results for the two geometric shapes are similar to the results from this experiment, with an average of 54.5% of perceptual switches occurring in the RVF. The results from the ballerina were also biased towards extra switches in the RVF, however it was unknown at the time that this image contained numerous problems, including a non-isometric view, which strongly favored it being seen as spinning clockwise.

Over the course of the current experiments it became clear that exposure time was a factor which could strongly influence how people experience KDE stimuli. For example, in this experiment, subjects experienced over 40% more perceptual rotation reversals in the second half of the experiment (2059 reversals) as compared to the first half (1455 reversals). Additionally there were many stimuli which seemed to have a hard cap on how often they could reverse. For example most of the humanoid figures in previous experiments reversed at rates approaching twice every 360 degree rotation. These rotation reversals correlated with the peak positions at which they were likely to reverse. Therefore subjects viewing these images over an extended period of time could begin to see them reversing at every peak resulting in a hard cap of two rotation reversals per revolution regardless of which hemisphere they were using to view the stimulus.

This result is relevant in the search for a hemispheric difference because the more a subject is exposed to these KDE stimuli the closer they will be towards approaching the hard cap of two reversals per rotation. And once they near that cap any differences between the reversal rates of the left and right hemispheres become impossible to detect, as both hemispheres are reversing at the maximum rate.

The fragility of the original result may be due to two factors:

1) Most of the images presented to the subjects in the initial experiments, as well as in the current ones, were geometric shapes, and therefore not affected by the high level biases which can cause the humanoid figures to have strong peak reversal locations at which they reach the reversal rate cap.
2) The overall exposure time of the subjects to the KDE images was much shorter (only 18 minutes as compared to 42 minutes here), and so the data collected was from the early part of the learning curve, where fewer reversals were occurring overall, so any differences across hemispheres were more noticeable.

If these potential confounds could be eliminated it might be possible to rediscover if a hemispheric difference if it does truly exist as the data suggest. This was one of the factors that led to the choice of Necker Cubes in the present experiment, as it more closely resembles the majority of stimuli in which the hemispheric difference was initially found.

Why would there be a hemispheric difference and why would it present itself as an increase in directional switching in the right visual field (left hemisphere)?

This result may be related to previous findings from static images, such as hierarchical letters that show the left hemisphere has a bias to the details and the right hemisphere has a bias for the big picture, as demonstrated in brain damaged patients and using hierarchical letters (Navon, 1977; Nass & Stiles, 1996). This could also relate to the findings of Chen & He (2003), where they reported a visual field asymmetry for binocular rivalry switch rates. In these experiments subjects experienced more perceptual switches and a faster rate of rivalry for images viewed by the RVF (left hemisphere). Finally, experiments using vestibular caloric stimulation showed that activation of the left hemisphere, but not the right, can alter perceptual rivalry rate of a static Necker cube causing subjects to have difficulty switching percepts (Miller et al, 2000).

Since the right hemisphere has an advantage for global image processing it is possible that the multistable KDE stimulus is more "sticky" in the right hemisphere, where the object is processed as a whole, and thus less likely to get reinterpreted by the brain and switch directions. On the other hand left-hemisphere processing of the local image information is less constrained by top-down influences from the perceptually current global configuration and thus may be more susceptible to reinterpretation resulting in an increase in directional switches for the Left Hemisphere.

Interestingly, a similar hemispheric difference was reported using binocular rivalry (Chen & He, 2003). Their results not only show a hemispheric difference for the reversal rate of a binocularly rivalrous stimulus, but they also reported increase in switch-rate for the Left Hemisphere for right-handed subjects. However the results from Chen and He are the result of an alternation between two simple patterns and would not be accounted for by the advantage to global processing in the right hemisphere that may explain the current data presented.

Note also, that it has been shown that the perception of point-light walkers depends on both global and local signal processing (Chang & Troje, 2009). It is also possible to design stimuli made from point-light walkers which are ambiguous and can undergo a depth-reversal. If the current result is correct then it might be expected that similar data could be found for ambiguous point-light walkers which undergo fewer depth reversals when viewed with the Right Hemisphere (Left Visual Field). This would be a worthy follow-up study.

Checking for time delay response

In the above experiments as noted, all of the subject's responses to perceptual rotation reversals were slightly delayed in time from the profile or front/back positions. It was assumed that this shift is simply a response time delay. To test this hypothesis an experiment was conducted using a silhouette from a previous experiment, in two image cycles which differed only in how quickly the 3D figure spun. The expectation was that for the slower moving figure, rotation reversal would be reported at a fixed time delay relative to the critical moment in the image sequence (for instance the profile pose) and therefore at a point where the figure was less advanced past that critical point in its rotation cycle.

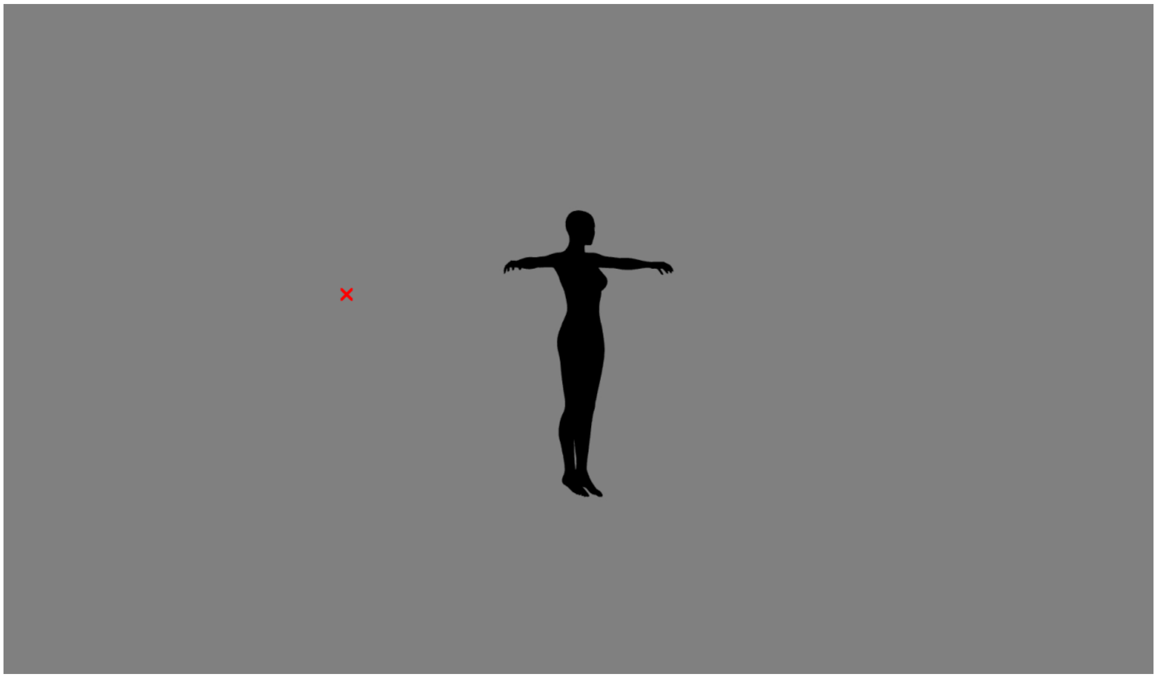


Figure 37: Sample stimulus from the Control Experiment

Control Experiment - Methods

Stimuli

As in Experiment Set 1, DAZ Studio 3 was used to create the stimuli. Both the slow and the fast condition used the same silhouette, of a humanoid with arms outstretched sideways. The KDE stimulus was created using 163 separate image frames to better pinpoint exactly when subjects reported direction reversals.

In the slow condition the silhouette rotated at a rate of 7.5 frames per second resulting in a rotation velocity of 21.73 sec/revolution.

In the fast condition the silhouette rotated five times faster, at a rate of 37.5 frames per second, resulting a rotation velocity of 4.35 sec/revolution.

The images were presented in a frontal orthographic projection.

The background color of the stimuli was a mean grey. The 3D silhouettes were a pure black.

Subjects were seated 26 inches from the monitor. At this distance the monitor subtended a viewing angle of 31.2 degrees horizontally and 14.9 degrees vertically. The images subtended between 2.2 to 8.8 degrees of horizontal visual angle depending on the frame, and 8.8 degrees of vertical visual angle.

Subjects

10 subjects participated in this study. They consisted of graduate and undergraduate students working at vision laboratories at UCSD.

Instructions

Subjects were instructed to indicate the initial direction of motion of the stimulus (clockwise or anti-clockwise as viewed from above) when it first appeared with a keypress [<] or [>]. Thus [>] indicated anti-clockwise rotation as viewed from above, and [<] indicated clockwise rotation as viewed from above. They were told to then press the [<] or [>] key to indicate the new direction of rotation whenever a directional change in motion occurred. They also had the option of pressing the [SPACEBAR] if the image started to do something that could not be described as rotating.

Image Presentation

Stimulus presentation order was randomized for each subject with either the slow or fast block coming first. During the experiment each stimulus was presented in the center of the screen for the duration of each trial block. The images were presented to the subject in exactly the same method as the previous experiments. Each block contained three viewing conditions (center, left, & right). During each block the subject would first be shown the 'center' condition and allowed to freely view the image for 120 seconds while they indicated changes in direction. After that viewing period the screen would temporarily blank out for 1 second as a fixation cross would randomly appear to the left or right of the screen (9 degrees off center). The subjects had been previously instructed to foveate on the fixation cross. One second after the screen blanked out the ambiguous figure would reappear at the center of the screen for the 120 seconds. Finally the screen would blank again and the fixation cross would appear in the opposite position. One second after the screen blanked out the ambiguous figure would reappear at the center of the screen for the 120 seconds. After each block of

Center/Left/Right viewing the subject was given an option to briefly take a break and continue the experiment when they were ready.

Data Collection

The data collected was the current frame being presented when the computer registered the subjects' response indicating they noticed a change. This experiment was testing for a slight lag between the time the subjects noticed a change, decided to press a button, and completed the button press.

The total run time for each subject was 12 minutes (6 minutes per stimulus).

One fast block with center, left, and right conditions of 120 seconds each. And one slow block with center, left, and right conditions of 120 seconds each.

At the end of the experiment subjects were asked to report any observations of interest and were asked at what specific positions the images seemed to change direction.

Control Experiment - Results & Conclusions

A visual inspection of the data clearly shows different means for the fast and slow rotating images, they are still relatively close. To confirm that this difference was significant and that the data sets were in fact different a Kuiper two-sample test was run to compare the data from the Left-Slow condition to the Left- Fast conditions, as well as a second comparison of the Right-Slow and Right-Fast conditions.

In both comparison cases the Kuiper test gave its maximal result of $p < 0.001$ indicating that there was in fact a difference between the Slow and Fast conditions. Thus it can be safely concluded that the delay in the results of the previous experiments are at least partially accounted for by a response time delay.

Since the polar plots of reversal points are dominated by a single lobe shortly after the profile position was reached, the direction of the mean resultant vector provides an appropriate estimate of the point during the rotation cycle at which rotation reversals are reported. The reversal reports came on average 29.5 degrees after the profile view in the slow rotation condition, and 43.9 degrees in the fast rotation condition. Subtracting the mean angular direction of the fast moving figure at the moment of reported reversal from that of the slow moving figure yields a difference of 14.4 degrees. This difference is the additional amount of angular rotation that the faster figure underwent before a directional switch was reported.

Table 9: Comparison of total number of reversals

	Fast Stimulus	Slow Stimulus
Total number of reversals:	**755**	**303**

Another very interesting result can be found by looking at the total number of reversals of the stimuli and comparing how many "opportunities" the stimuli had to switch in their peak locations.

Since one stimulus was moving much faster, it arrived in the preferred profile position for direction reversals five times as often as the slower moving stimulus. Yet looking at the difference in the total number of reversals, the fast condition only had a 250% increase over the slow condition, despite having 500% as many "chances" to undergo a directional switch in the preferred position.

This difference suggests that increasing the rate of revolution can increase the how often KDE images undergo direction reversals, but that the increase is not linear. Three possible explanations exist for this behavior:

1) As the rate of rotation increases, the stimulus reaches its peak rotation reversal positions more often. However the increased rotation velocity requires that when a direction reversal occurs, the perceptually implied change in angular momentum is correspondingly increased. In the slow condition the visual system needs to instantaneously see the object reverse its angular velocity by 33.14 degrees/sec. While in the fast condition a rotation reversal requires an angular velocity change of 165.52 degrees/sec. Moreover, if we imagine that a limited range of poses conducive to

reversal provide a ‘window of opportunity’ for reversals, the duration of this window is shorter the higher the velocity.

2) As the rotation velocity increases there is a corresponding increase in the linear image motion across the screen. In the case of this stimulus the preferred point of direction reversal, in profile position, corresponds to the point of maximal average contour velocity across the screen (owing to the extended arms). Experiment 3 showed that linear image velocity has at least a small effect on direction reversals. It is possible that increased linear image velocity, for the fast stimulus in this experiment, acts to inhibit perceptual direction reversals.

3) Finally consideration must be made for an upper limit to how often a perceptual direction reversal can occur in the visual system. If perceptual direction reversals are a result of adaptation and inhibition (Kang & Blake, 2010), then there may be a biologically limited time course for how often directional switches can occur. Additionally from an evolutionary perspective it can be argued that there is an advantage to the visual system taking some time to reassess a particular scene before presenting your consciousness with an alternative explanation. However if the scene is constantly

> in flux that benefit would be quickly lost. Thus there is likely to be an upper limit to how frequently any type of perceptual reversal can occur in the visual system.

These data can be used to determine whether or not the cardinal points of the revolution are the actual positions where the subjects experienced perceptual rotation reversals. To do this the following formula is constructed to calculate the R_d (Response Delay) for each condition:

$$F_r = C_a + (R_d * R)$$

F_r = Mean Angle of Reported Reversal

C_a = Relevant Cardinal Angle

R_d = Response Delay

R = Rotation Speed

If the cardinal positions are the actual locations where directional reversals are experienced then the frame at which the reversal is reported will be offset from the cardinal position frame by a number of frames directly proportional to the framerate, and the Response Delay will be the same for both the slow and fast conditions.

However the data show that this is not the case at all. Using the formula above, $R_{d(FAST)}$ = 481 ms and $R_{d(SLOW)}$ = 1763 ms; surprisingly long as they are, especially in the slow condition, these values are simply the mean delays between the profile view and the majority of reversal reports. The long delay in the slow condition makes it clear that in fact subjects are not experiencing direction reversals at the cardinal points at all.

The substantial implied integration time is not implausible. At the cardinal points in the rotation the image of the object has not yet done anything to indicate a perceptual direction reversal. For example if it simply stopped in the cardinal position no perceptual direction reversal would be necessary as there would be no image motion at all. It is only *after* the object continues to rotate out of the cardinal position that the image of the object can provide a reason to make a revised estimate of the direction it is spinning.

Some research has shown that a delay of at least 390 ms would be expected before a subject can respond to a perceived change of an ambiguous stimulus (Kornmeier & Bach, 2012). As suggested above, the average response delay calculated above for the faster moving stimulus (481 ms) may include both the actual response delay and a delay correlated with how much the object needed to rotate before the visual system might initiate a perceptual direction reversal.

It is possible to estimate the amount of rotation required by, taking the response delay calculated (481 ms) and subtracting the expected response delay (390 ms) based on the work of Kornmeier & Bach. This yields a difference of 91 ms. Multiplying this by the angular velocity yields of the fast object gives an estimate of 7.53 degrees of angular rotation required before the visual system might initiate a perceptual rotation reversal.

However this is not the only way to come up with an expected amount of rotation required for a perceptual switch to occur. Consider the following:

If it is assumed that the response delay from the moment of reversal is the same for the two conditions, we may substitute its value for Rd in the above equation, and solve for Rd on the assumption that the critical angle C_a is not the profile view but some slightly earlier or later view, but is again identical in both fast and slow conditions. This yields

Rd (nonvisual report delay): 161 msec

Critical angle at which reversal is experienced: 26.5 degrees past profile

Visual processing time intervening between profile view and critical angle: 317msec in fast condition, 1.58 sec in slow condition

On this analysis only a very short component (162 msec) of the delay is the stimulus-independent reaction time. This delay between the moment of reversal and the moment of report, while shorter than typical reaction time measures, is not unreasonably short given that it represents only the output stages of neural and motor processing, whereas reaction times to external stimuli also include the time for sensory stimuli to enter consciousness.

The analog of the sensory component of a typical reaction time is the second, rotation-speed dependent component that accounts for the bulk of the report delay: the time intervening between the profile view and the critical angle at which the perceptual rotation reversal is experienced.

These examples are just first attempts at determining how much visual angle would be required for a rotation reversal, but it is likely that there are more factors involved. For example slower moving objects might have direction reversals that are harder to detect and thus have a longer response delay. There could also be an interaction between a critical amount of visual angle past the cardinal points and a critical amount of time for which the visual system has witnessed the object beyond the cardinal point. Still it is a good proof of concept of how one could attempt to determine the actual angle which triggers perceptual rotation reversals.

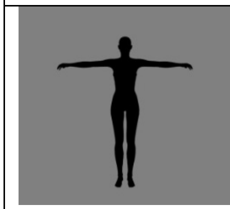
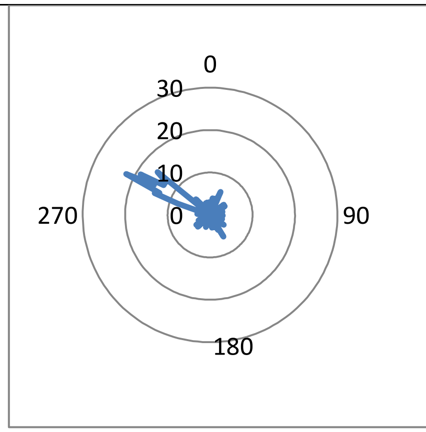
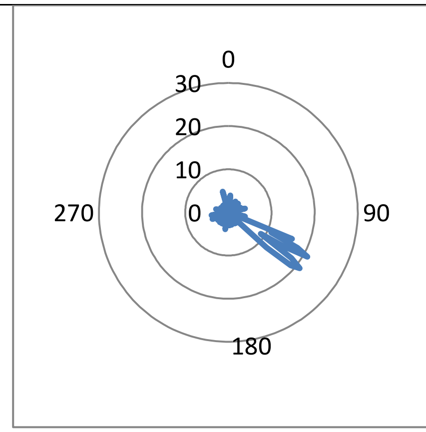
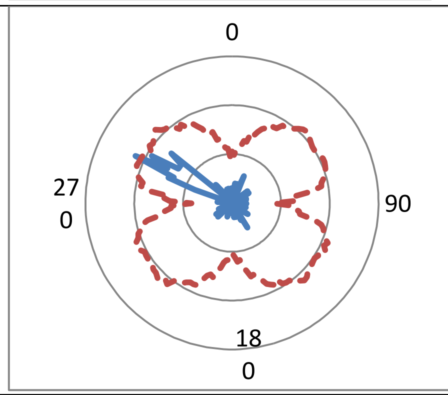
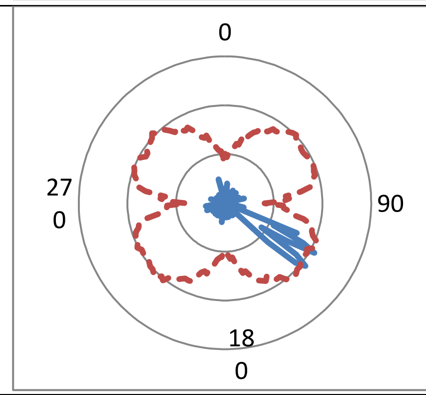
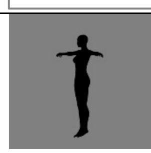
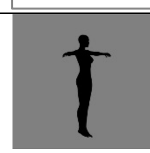
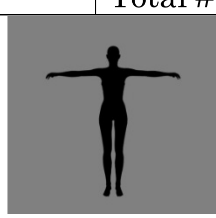
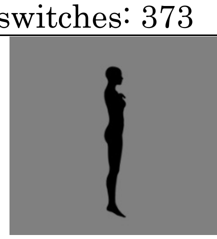
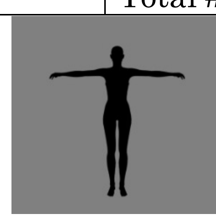
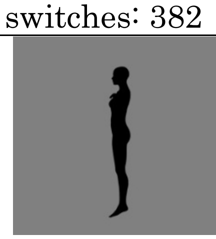
<table>
<tr><td>Figure 38a - Fast</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td></td><td></td></tr>
<tr><td></td><td>
mode = 298.3
mean = 311.85
p-value = 6.93 e-020
Total # switches: 373</td><td>
mode = 121.5
mean = 127.66
p-value = 6.85 e-031
Total # switches: 382</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 38 a: Results from the fast stimulus in the control experiment

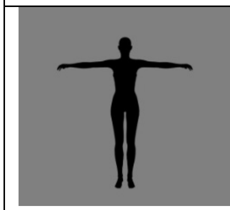
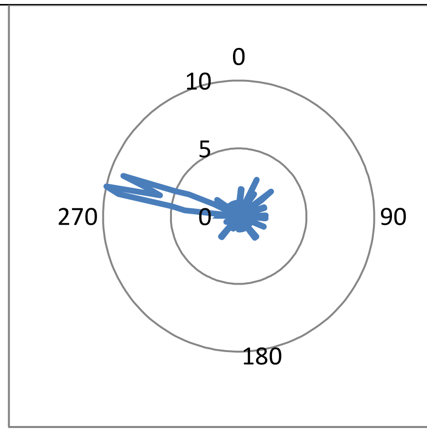
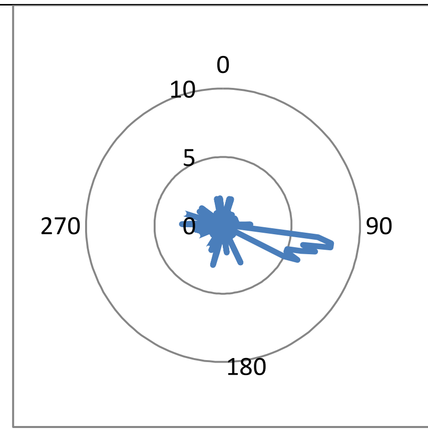
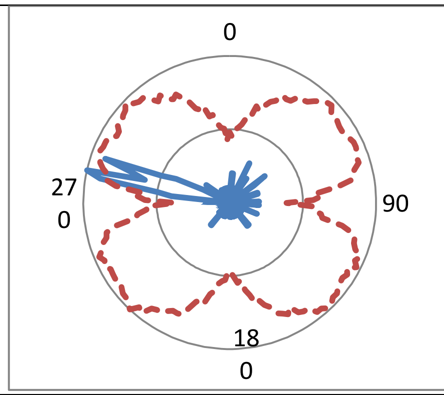
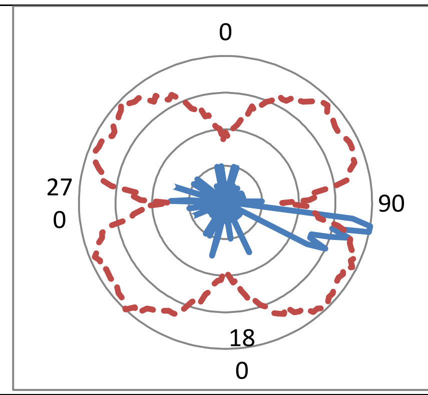
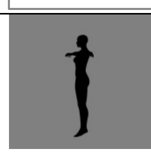
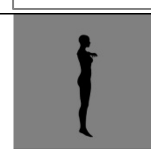
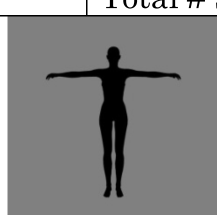
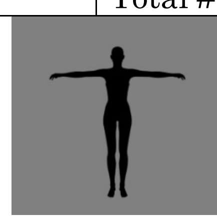
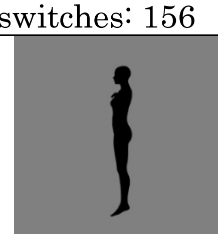
<table>
<tr><td>Figure 38b-Slow</td><td>Anti-Clockwise Reversal</td><td>Clockwise Reversal</td></tr>
<tr><td></td><td>
</td><td>
</td></tr>
<tr><td>Frame-by-Frame analysis of linear image motion</td><td>
</td><td>
</td></tr>
<tr><td></td><td>
mode = 285.1
mean = 301.61
p-value = 1.19 e-007
Total # switches: 147</td><td>
mode = 101.6
mean = 117.43
p-value = 4.29 e-010
Total # switches: 156</td></tr>
<tr><td colspan="3">   
Degrees: 0 | 90 | 180 | 270</td></tr>
</table>

Figure 38 b: Results from the slow stimulus in the control experiment

Frame by Frame Difference Analysis

It is important to take a more analytical look at why people might be experiencing rotation reversals at specific points within the stimuli. To this end a frame-by-frame analysis was conducted on each of the KDE stimuli to analyze at how much the image of the object changed from one frame to the next. If all the rotation reversals coincide with points of minimum image change then this could indicate a simple low-level phenomenon. On the other hand reversals occurring out of sync with minimum image change may have a more high-level explanation. Furthermore it should be noted that the image-change from frame to frame correlates directly with the total linear image speed. This is due to the fact that it is the contour of the silhouettes that defines the image in each frame. The frame-by-frame difference analysis is therefore measuring the motion of image contour over time.

Frame by Frame Difference Analysis - Method

The steps to generate the frame-by-frame difference analysis of each grayscale image-series are as follows:

1) The mean gray of the image as a whole was calculated.
2) This value was used to divide the image into pure black and white, white being all values above the mean and black all values below.

3) Next a simple XOR logical function was used to compare one image-frame to the next and find the difference between frames with larger values indicating a greater change from one frame to the next.
4) Finally the frame-by-frame difference was plotted together with the directional switch data for each stimulus.

<u>Frame by Frame Difference Analysis - Results & Conclusions</u>

The frame by frame analysis results have been included in the relevant sections of the text above. The linear image velocity shown in the graphs is at a scale that is proportional to the other relevant data so that comparisons can be made and it is represented by a dashed red line.

It should be noted however that this analysis revealed an artifact in the stimuli where the linear image motion will suddenly drop rather than vary sinusoidally.

These drops are caused by two properties of the stimuli

1) Some of these drops in linear image motion are due to the fact that during the revolution of a KDE stimulus there are some image frames where the silhouette occludes itself. This generally occurs in the profile position and is a result of the thick body of the stimulus obscuring any view of the thinner arm. This results in two or more sequential image frames which are almost identical and have minimal

linear image motion from one frame to the next. This *does* represent a very brief stalling of the stimulus' linear motion, however the 3D rotation is consistent and it is merely the image that stalls.

2) The second cause of these drops in linear motion is due to a minimal difference between image frames at the cardinal positions (front/back or profile). In these cases the frames used to present the stimulus will sometimes pass through the cardinal points from one frame to the next without a full view of the object in its cardinal position. This is especially true in the profile positions because the image frames before and after the profile view are almost identical to each other.

The abnormalities in linear image motion discussed here are clearly present in the stimuli presented. However the data show that overall linear image motion has only a minor effect on the how subjects experience rotation reversals throughout these experiments. Therefore while it is important to note the artifacts, they do not seem to be affecting the experimental results.

Model of Perception of Rotation Reversals

The results of the experiments shown here suggest that there are multiple factors determining at which points within a revolution that rotation reversals are most likely to be experienced. The key factors seem to be (in order of importance):

1) **3D Object Reinterpretation**: Every time a perceptual rotation reversal occurs there is a mirror reversal in depth of the KDE silhouette. Certain points within the rotation require a much smaller depth reinterpretation of the 3D image. The data show that the larger the reinterpretation required the less likely a rotation reversal will occur.
2) **Forward Facing Bias**: There is a clear bias for seeing the KDE stimuli facing towards the subject (for humanoid figures). This results in an increased likelihood of rotation reversals occurring when the stimuli are in the profile position and the continuation of rotation in the current direction would result in an away-facing figure.
3) **Linear Image Motion**: This seems to be more of a factor images which have less high level bias. It is still unclear exactly how much linear image motion affects perceptual rotation reversals but at a minimum it may be facilitating reversal at locations where other factors already would induce one.

4) **3D Object Angular Velocity**: Stimuli which move more quickly through their revolution (higher angular velocity), will undergo fewer direction reversals per rotation than an identical, but more slowly rotating stimulus. This is likely due to 1) a neurological limit on how often a perceptual direction reversal can occur and 2) an increase in the total angular velocity change between rotation states which would therefore require a larger reinterpretation of angular velocity from one perceptual state to the next as compared to the more slowly rotating stimulus.
5) **Time**: The longer the amount of time that has passed without a perceptual rotation reversal the more likely it is that a rotation reversal will occur on the next frame.
6) **Experience with KDE Stimuli**: As people gain experience viewing KDE stimuli they become more likely to experience a rotation reversal in any given frame. This seems to be a factor both in terms of a specific stimulus and KDE stimuli as a whole, indicating some kind of perceptual generalization.
7) **Image size**: The data suggest two main effects: (1) there is a greater probability for a subject to see a smaller stimulus facing away from

them, and (2) larger stimuli are slightly more likely to result in a rotation reversal.

These observations suggest that it is possible to construct a formula which would predict the likelihood of a perceptual rotation reversal occurring, for any given frame. While this dissertation does not empirically test the effectiveness of this formula, I will describe here the functions that would be included. The final formula is a weighted combination of these functions. Unless otherwise noted all functions depend on the current frame (position) of the figure.

OR = A quantity representing the amount of object reinterpretation required if a rotation reversal occurs in a given frame. This will vary sinusoidally for any given KDE stimulus and will reach zero in frames where no object reinterpretation needs to occur. As this variable increases a rotation reversal becomes **less** likely.

FB = Forward facing bias. This varies sinusoidally with the orientation of any humanoid figure, with its maximum on the frames where the figure is in a profile position and where continued rotation in the current direction will cause the figure to begin to face away from the viewer. As this variable increases a rotation reversal becomes **more** likely.

OM = 3D object motion. The higher the angular velocity of a stimulus the less likely that a perceptual direction reversal will occur at any given moment. For most experiments this is a constant. As this variable increases a rotation reversal becomes **less** likely.

IM = Linear image motion. This is how many pixels change between this frame and the last. Low rates of change from frame-to-frame slightly increase the likelihood of a rotation reversal. As this variable increases a rotation reversal becomes **less** likely.

T = Time since last rotation reversal. This is independent of the current frame.

F(T) = This is the function, with time as an input, that modulates rotation reversals as a function of when the last reversal occurred. This function starts as a large negative quantity (for T=0) that inhibits rotation reversals. As time passes, and the value of T increases, this function becomes positive and increases. This is then passed through a sigmoid function. The steepness of the sigmoid is determined by experience ($XP_s + XP_t$)). As this variable increases a rotation reversal becomes **more** likely.

XP_s = Stimulus Experience is a sigmoidal function representing the total amount of experience the observer has with a particular KDE stimulus. As this variable increases a rotation reversal becomes **more** likely.

XP_t = Total Experience is a sigmoidal function representing the sum total amount of experience the observer has with all KDE stimuli. As this variable increases a rotation reversal becomes **more** likely.

Frame = current frame of the KDE movie

a,b,c,d,e = Scaling coefficients

Likelihood (frame) =

$$a*FB(frame) + b*(XP_t+XP_s)*F(T) - c*OR(frame) - d*IM(frame) - e*OM$$

The combination of these functions is not guaranteed to lie between zero and one, as would be required of a probability, nor would we want the function to reach one anyway, since no single frame ever causes a direction reversal every time it is seen. This could be solved by passing the sum through a transfer function such as a sigmoid or even a simple step function. The maximum and minimum asymptotes of these transfer functions would be scaled according to subject, according to the frame which caused the most reversals, and the average reversal rate across all frames, respectively.

Conclusion

Contrary to much of the previous research on multistable perception, the current work clearly shows that we can make very strong assertions about when and why perceptual rotation reversals will occur with the Kinetic Depth Effect.

The following factors have been clearly identified as affecting perceptual rotation reversals:

1) **Depth reinterpretation** – In the course of a revolution of a KDE stimulus the image frames will vary in how much reinterpretation of the 3D object in depth would be required in the case of a rotation reversal. Those image frames with where depth reinterpretation required is minimal are more likely to be a position where a rotation reversal occurs. This is the strongest of the parameters discovered so far which can modulate perceptual direction reversals. It is may also be considered a lower level parameter since it is the shape of the stimulus that ultimate matters but not the content.

   In the case of the Ames window discussed earlier, the visual system chooses to map the 2D image on the retina in such a way that perspective is maintained. This results in a completely unrealistic and physically impossible percept and demonstrates that the visual system has a strong preference to maintain perspective over other visual cues.

Here the visual system produces rotation reversals that result in an instant change from facing-toward to facing-away. This may mean that there is a preference for minimizing the amount of depth reinterpretation required or perhaps minimizing the amount of mass transferred through space.

2) **Forward Facing Bias** – After depth reinterpretation this seems to be the second most powerful parameter determining the position at which a perceptual direction reversal is most likely to occur. This preference is so strong, that after prolonged viewing of KDE stimuli, subjects would sometimes be a state of never seeing the stimulus face away from them. This is quite clearly a high-level bias, and taken together with depth reinterpretation it shows that both high and low-level factors play a role in determining the position at which perceptual rotation reversals are most likely to be seen.

3) **Viewing From Above Bias** - When KDE stimuli are presented in an isometric parallel view rather than a frontal view those stimuli are much more likely to persist in the rotational state which correlates with viewing from above. This seems to be a general bias of the visual system and it is unsurprising to find that it can impact the viewing experience of KDE stimuli.

4) **Linear Image Motion** – This was shown to be only a minor factor in determining where a perceptual rotation reversal would occur. This is

surprising because many KDE silhouettes are primarily defined by the linear image motion produced by their rotation. In fact even a static contour of a silhouette provides minimal information about the original object. Much more information is required for the visual system to be able to accurately transform that a silhouette image into a 3D percept, and that information is contained primarily within the linear image motion of the contour. Thus it is fairly surprising that modulating the linear image motion doesn't produce a larger effect.

Ultimately this data shows that directional reversals of a multistable KDE stimulus can be predicted with great accuracy. This not only shows that perceptual reversals need not be random, but it also will allow future vision research to use the paradigm of KDE rotation reversals as a window into the inner workings of the human visual system.

References

Alais, D. & Blake, R. (2005). Binocular rivalry and perceptual ambiguity. Cambridge, MA: MIT Press.

Ames, A. (1952) The Ames Demonstrations in Perception, New York, Hafner Publishing.

Azoulai S. & MacLeod D. (2009), *Hemispheric differences in the kinetic depth effect*, Poster Presented at the Optical Society Annual Meeting, Seattle, Washington.

Bennett, B. M., Hoffman, D. D., Nicola, J. E., & Prakash, C. (1989). Structure from two orthographic views of rigid motion. *The Journal of the Optical Society of America*, *A*(6), 1052-1069.

Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*(10). http://www.jstatsoft.org/v31/i10

Blake, R. (1989). A neural theory of binocular rivalry. *Paychological Review 96*(1), 145-167.

Blake, R. (2001). A Primer on Binocular Rivalry, Including Current Controversies. *Brain and Mind*, *2*, 5–38.

Breese, B.B. (1909). Binocular rivalry. *Psychological Review, 16*(6), 410-415. doi:10.1037/h0075805

Breese, B.B. (1899). "On inhibition". *Psychological Monographs 3*(1), i–65. doi/10.1037/h0092990

Brown, K.T. (1955) Rate of apparent change in a dynamic ambiguous figure as a function of observation time, American Journal of Psychology.68, 358–371.

Carter, O., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception*, *32*(3), 295–305.

Carter, O., Pettigrew, J. D., Hasler, F. , Wallis, G. M., Liu, G. B., Hell, D., & Vollenweider, F. X. (2005). Modulating the rate and rhythmicity of perceptual rivalry alternations with the mixes 5-HT2a and 5-HT1a agonist psilocybin. *Neuropsychopharmacology, 30*, 1154–1162.

Chang, D. H. F., & Troje, N. F. (2009). Characterizing global and local mechanisms in biological motion perception. *Journal of Vision, 9*(5):8, 1–10). doi:10.1167/9.5.8.

Chen, X., & He, S. (2003). Temporal characteristics of binocular rivalry: visual field asymmetries. *Vision Research*, *43*(21), 2207–2212. doi:10.1016/S0042-6989(03)00359-6

Einhäuser, W., Martin, K. A.C., & König, P. (2004). Are switches in perception of the Necker cube related to eye position? *The European Journal of Neuroscience*, *20*(10), 2811–2818. doi:10.1111/j.1460-9568.2004.03722.x

Fahle, M. (1982). Binocular rivalry: suppression depends on orientation and spatial frequency. *Vision Research 22,* 787-800. doi:10.1016/0042-6989(82)90010-4

Glen, J.S. (1940) Ocular movements in reversibility of perspective. *Journal of General Psychology 23,* 243-281.

Gregory, R. L. (1973). The confounded eye. *Illusion in nature and art*, pp. 49-96.

Hildreth, E. & Ullman, S. (1982). *Measurement of visual motion*, Retrieved from Massachusetts Institute of Technology Artificial Intelligence Laboratory Web site: http://dspace.mit.edu/handle/1721.1/45554

Huang, T. S. & Lee, C. H. (1989). Motion and structure from orthographic projections. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(5), 536-540.

Kang, M.S. (2009). Size matters: a study of binocular rivalry dynamics. *Journal of Vision, 9,* (1):17, 1-11. doi:10.1167/9.1.17

Kang, M.S. & Blake, R. (2010). "What causes alternations in dominance during binocular rivalry?" *Attention, Perception, & Psychophysics, 72*(1), 179-186.

Kanizsa, G (1955), "Margini quasi-percettivi in campi con stimolazione omogenea.", *Rivista di Psicologia* **49** (1): 7–30

Kleffner, D.A. & Ramachandran, V.S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*(1), 18-36.

Knill, D. C., Kersten, D. & Yuille, A. (1996). A Bayesian formulation of visual perception. In D. C. Knill & W. Richards (Eds.), *Perception as Bayesian Inference.* Cambridge University Press.

Knill, D. C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Research, 43*(7), 831-854.

Kornmeier, J., & Bach, M. (2012). Ambiguous figures - what happens in the brain when perception changes but not the stimulus. *Frontiers in human neuroscience*, *6*(March), 51. doi:10.3389/fnhum.2012.00051

Lehky, S.R. (1988). An astable multivibrator model of binocular rivalry. *Perception 17,* 215-228.

Leopold, D.A. & Logothetis, N.K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature 379,* 549-553.

Logothetis, N.K., Leopold, D.A. & Sheinberg, D.L. (1996). What is rivalling during Binocular rivalry? *Nature 380,* 621-624.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, *200*(1140), 269–94.

Marr, D. (1982). Vision: a computational investigation into the human representation and processing of visual information. New York: W. H. Freeman.

McClelland, J.L. & Rumelhart, D.E. (1988). Explorations in parallel distributed processing: A handbook of models, programs, and exercises. Boston, MA: MIT Press.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision* 4(7), 539–551. doi:10.1167/4.7.2

Miller, S., Hansell, N., Ngo, T., Liu, G., Pettigrew, J., Martin, N., & Wright, M. (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proceedings of the National Academy of Sciences*, *107*, 2664–2668.

Miller, S., Liu, G. B., Ngo, T., Hooper, G., Riek, S., Carson, R., & Pettigrew, J. (2000). Interhemispheric switching mediates perceptual rivalry. *Current biology : CB, 10*(7), 383–92.

Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *Journal of Neurophysiology, 98*(3), 1125-1139. doi:10.1152/jn.00116.2007

Nakayama, K. (1982). Motion parallax sensitivity and space perception. In A. Hein & M. Jeannerod (Eds.), *Spatially Coordinated Behavior,* (pp. 223-242). Academic Press.

Nass R & Stiles J (1996): Complication of the perinatum: Congential focal lesions. In Frank Y (Ed), Pediatric Behavioral Neurology. Boca Raton, FL: CRC Press, pp.55-64.

Navon, D. (1977). Forest before the trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–383.

Necker, L.A. (1832). Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid. *London and Edinburgh Philosophical Magazine and Journal of Science, 1*(5), 329–337.

Pastukhov, A., Vonau, V., & Braun, J. (2012). Believable change : Bistable reversals are governed by physical plausibility. *Journal of vision*, *12*(1), 1–16. doi:10.1167/12.1.17.

Pettigrew, J. D., & Miller, S. M. (1998). A “sticky” interhemispheric switch in bipolar disorder? *Proceedings. Biological sciences / The Royal Society*, *265*(1411), 2141–2148. doi:10.1098/rspb.1998.0551

Richards, W., Koenderink, J.J., & Hoffman, D.D. (1985). *Inferring 3D Shapes from 2D Codons.* Retrieved from Massachusetts Institute of Technology Artificial Intelligence Laboratory Web site: ftp://publications.ai.mit.edu/ai-publications/pdf/AIM-840.pdf

Ross H E, Plug, C., 1998. The history of size constancy and size illusions. In Walsh, V. & Kulikowski, J. (Eds) *Perceptual constancy: Why things look as they do.* Cambridge: Cambridge University Press, 499–528.

Schouten, B., Troje, N. F., & Verfaillie, K. (2011). The facing bias in biological motion perception: structure, kinematics, and body parts. *Attention,* *perception & psychophysics*, *73*(1), 130–143. doi:10.3758/s13414-010-0018-1

Shannon, R. W., Patrick, C. J., Jiang, Y., Bernat, E., & He, S. (2011). Genes contribute to the switching dynamics of bistable perception. *Journal of vision*, *11*(3), 1–7. doi:10.1167/11.3.8.Introduction

Sinha, P, & Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature*, *384*(5), 460–463.

Sundareswara R, Schrater P. "Perceptual multistability predicted by search model for Bayesian decisions" Journal of Vision. 2008;8:1–19. doi: 10.1167/8.5.12

Tong, F. (2001) Competing theories of binocular rivalry: a possible resolution. *Brain and Mind 2*(1)*,* 55-83.

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in cognitive sciences*, *10*(11), 502–11. doi:10.1016/j.tics.2006.09.003

Troje, N. F., & McAdam, M. (2010). The viewing-from-above bias and the silhouette illusion. *i-Perception*, *1*(3), 143–148. doi:10.1068/i0408

Ullman, S. (1979). The interpretation of visual motion. Cambridge, MA: MIT Press.

Ullman, S. (1984-A). Rigidity and misperceived motion. *Perception*, *13*(2), 219–20.

Ullman, S. (1984-B). Maximizing rigidity: the incremental recovery of 3-D structure from rigid and non-rigid motion. *Perception*, *13*(3), 255–74.

Ullman, S & Yuille, A. (1987). Rigidity and smoothness of motion. *Volume 989 of AI memo.* Defense Technical Information Center. Retrieved from www.dtic.mil/dtic/tr/fulltext/u2/a261212.pdf

Vanrie, J., Dekeyser, M., & Verfaillie, K. (2004). Bistability and biasing effects in the perception of ambiguous point-light walkers. *Perception*, *33*(5), 547–560. doi:10.1068/p5004

Wallach, H., & O'Connell, D. N. (1953). Kinetic depth effect. *Journal of Experimental Psychology*, *45*(4), 205-217.

Wheatstone, Charles (1838). Contributions to the physiology of vision.—Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, 128,* 371-394.

Zar J.H. (1999). *Biostatistical Analysis. (*4th ed.). Upper Saddle River, NJ: Prentice Hall.