MUG SHOTS: SYSTEMATIC BIASES IN THE PERCEPTION OF FACIAL ORIENTATION WITHIN PICTORIAL SPACES

MAXWELL ESSER

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

MAY, 2023

© MAXWELL ESSER, 2023

Abstract

Pictures are 2-D projections of a 3-D world, so pictorial spaces behave differently than the 3-D visual spaces we inhabit. For instance, the angular orientation of a face pictured in half-profile view is systematically overestimated by the human observer – a 35° view is estimated to be approximately 45°. What is the cause for this perceptual orientation bias? We tested three different hypotheses. (1) The phenomenon is specific to pictorial projections due to the twofoldness of the medium and does not occur in 3-D space. (2) It can be explained with the depth compression expected when the vantage point of the observer is closer to the picture than the point of projection. (3) The visual system uses a shape prior that does not match the elliptical horizontal cross section of a typical head. Our results support the third hypothesis, and this effect can be mitigated through adding geometric information through structure-from-motion.

Acknowledgements

The completion of this project amidst the COVID-19 pandemic was made possible through the contributions and support of my colleagues in the Biomotion Lab at York University, as well as my supervisor Dr. Niko Troje. Though the constantly changing circumstances and restrictions of the pandemic forced the focus of this thesis to shift multiple times, the support and guidance I received from Dr. Troje enabled me to refocus multiple experiments around one unique phenomenon – orientation bias when viewing faces in pictorial space. I have immense gratitude to Dr. Troje for welcoming me to the Biomotion Lab and providing consistent and reliable support to me despite the global circumstances. Dr. Troje assisted in the conception of each of the experiments contained within this thesis, provided me with relevant literature recommendations, wrote the first draft of the abstract, and provided written feedback on the thesis paper itself. Additionally, Dr. Troje presented the findings of this research project at the May 2022 Vision Science Society (VSS) conference. Dr. Anne Thaler is also deserving of special thanks for her contributions to these experiments. Dr. Thaler provided hands-on guidance and feedback throughout the process of constructing and implementing each of the experiments, informing the coding and statistical analyses of each, as well as running participants for the third experiment in my absence. Lastly, I want to acknowledge Dr. Adam Bebko for his help debugging the first experiment, which was designed in Unity3D. My other colleagues in the Biomotion Lab provided emotional support, relevant literature, and reviewed the current manuscript. My success, and the success of this project, were facilitated by the support and contributions of my colleagues and supervisor. It has been an absolute privilege to be a part of the Biomotion Lab and work closely with each member throughout the completion of my MSc. Thank you all.

Table of Contents

List of Tables

List of Figures

Introduction

Experiment I

Experiment II

Introduction

As the years progress, people spend increasingly more time interacting with images through screens. People engage with their favorite characters in movies, TV shows, and video games. Video displays have long been used for educational purposes, such as training and simulations for prospective doctors and pilots; they may even one day be used to replace certain jobs. Psychophysical studies frequently make use of pictures and videos to help explain real-world phenomena. In recent history, particularly during the COVID-19 pandemic, teleconferencing through Zoom has become adopted ubiquitously as a primary means of communication. Be it for school, conferences, professional group meetings, or simply to catch-up between friends, video communication has become an integral part of people's normal social interactions.

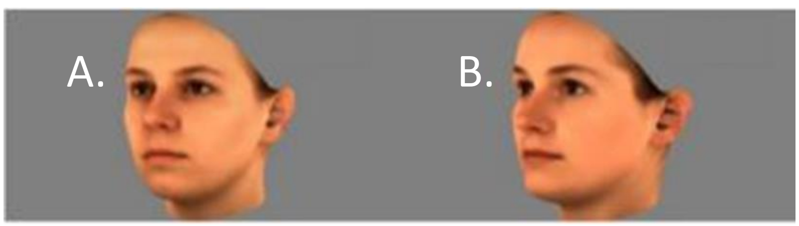

***Figure 1.1**: This image depicts the same face oriented 35˚ (A) and 45˚ (B) left of a frontal-facing view. When asked which head is oriented 45˚ left of a frontal view (or exactly halfway between a frontal and a profile view), people tend to incorrectly choose head A. Head B is similarly overestimated to be approximately 55˚ left of a frontal view.*

However, this type of interaction is an ill-fitted substitute for the advantages of in-person communication. This is in part due to inherent differences between two different visual mediums, and their respective abilities to accurately portray spatial relationships between depicted elements. Images, such as those typical of picture or video displays presented through screens, are two-dimensional (2-D) projections of the three-dimensional (3-D) world. They offer a view into a space that observers can readily understand even though they are not part of it. We call such spaces "pictorial spaces". Pictorial spaces stand in contrast to the spaces we see in front of us when we open our eyes in the real world, which we refer to as "visual spaces”. One such inherent difference between visual and pictorial spaces is that only in visual spaces does the observer occupy a well-defined location with respect to other objects contained in that space. Since an observer viewing a scene in pictorial space cannot occupy said space, their ability to discern the relative locations and orientations of objects in the observed pictorial space, with respect to themselves, might be compromised. This makes sense intuitively and has been previously documented (Goldstein, 1987; Hecht et al., 1999; Troje et al., 2018).

The studies conducted in this thesis are aimed at explaining one specific and prominent phenomenon which is characteristic of pictorial space that is demonstrative of the directionality of depicted objects being compromised. This phenomenon is deemed “orientation bias”. Throughout this thesis, when referring to a pictured object’s “orientation,” this is defined as the direction the pictured object appears to point when extended out of a picture (Cutting, 1988; Goldstein, 1988). When viewing faces presented in pictorial spaces, such as those viewed during video conference calls, observers tend to systematically overestimate the angle of the pictorial face’s orientation around its vertical axis when viewed from a fronto-parallel perspective, biasing their estimates away from the surface normal of the image plane. This orientation bias effect is maximized when presented stimuli have a 45˚ orientation, in which the average estimate is approximately 55˚, a deviation of approximately 10˚ away from the surface normal (Goldstein, 1987; Troje et al., 2018; Fig. 1.1). We often assume that pictorial spaces act as complete proxies for visual ones, but the orientation bias phenomenon suggests that this is not the case. This assumption is made, for example, when we present a visual scene to an observer in a psychophysical study, or when we communicate with another person by means of video-conferencing technology. In both of these cases, having the perception of orientation being skewed diminishes the accuracy of the proxy. Furthermore, maintaining the orientation of pictured persons is particularly important for social interactions, as it facilitates accurate perception of nonverbal cues, such as head nods, gestures, and gaze direction. In fact, “head orientation contributes 68.9% on average to the overall gaze direction, and focus of attention estimation based on head orientation alone can get an average accuracy of 88.7% in a meeting application scenario with four participants” (Steifelhagen & Zhu, 2002). Therefore, understanding the factors which contribute to the orientation bias phenomenon, and thus better understanding the key factors which differentiate how people perceive visual and pictorial spaces, can help inform future vision research as well as motivate technological advancements which could be used to improve the teleconferencing experience.

Factors which may Contribute to the Orientation Bias Phenomenon

Depth Compression

To explore the potential factors which contribute to orientation bias, we must first examine some of the fundamental differences between visual perception in visual and pictorial spaces. Depth compression is one such important factor to consider when analyzing the orientation bias phenomenon, as it directly affects angular estimations in pictorial spaces by deforming depicted objects (Hecht et al., 1999; Koenderink & van Doorn, 2008; Fig. 1.2). Compressing an object in depth changes the spatial relationship between elements in the scene, which causes the slant of depicted objects to appear more oblique, thus potentially causing an overestimation of observed angles (Fig. 1.2). Depth compression contributing to difficulty discriminating relative orientation is also recorded by Reichel et al. (1995). They

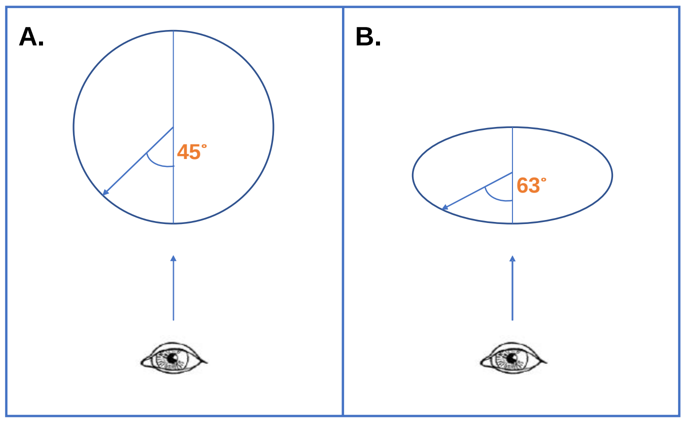

Figure 1.2*:* This image depicts the effects of compression of subjective space on apparent angular orientation. Two heads (represented as circles) from a top-down view. **A** shows the unaltered head viewed in the real world, in which the orientation relative to the frontal is a clear 45˚. **B** shows the same image compressed in depth by 50%. Doing this changes the spatial relationship between elements in the scene, causing the angle to appear to be more obtuse.

performed an experiment which employed 2.5-D sketches to produce compelling perceptions of 3-D surfaces despite being pictorial representations and inquired to the perceived orientations of the depicted surface normal. Their results indicated observers' imprecision when it comes to judging relative orientations of 2.5-D surfaces with Weber fractions ranging from 40%-140%. Though this imprecision does not equate to the inaccuracy we have labeled the orientation bias effect, it does support the notion that people often have difficulty estimating the local surface depth of 3-D objects depicted in pictorial spaces. There are two potential mechanisms which could cause perceived depth compression that will be explored in this thesis. The first is the twofoldness of the pictorial medium, a characteristic of pictures being 2-D objects which simultaneously represent 3-D spaces. The second is the vantage point of the observer relative to the pictorial plane. Changing the vantage point from the center of projection can alter the perceived pictorial array. To properly analyze the orientation bias phenomenon, both potential sources of depth compression will be examined.

Twofoldness

Pictorial spaces necessarily exist on a physical pictorial plane which itself exists in visual space. The duality between the flatness of the pictorial medium and the depth of the pictorial depiction has been defined by Wollheim (1998) as "twofoldness" – the simultaneous representation of a picture's surface and the representation of objects in the picture – and is likely a key mechanism in the differential perception of objects contained in visual and pictorial spaces. The twofoldness of pictures could create a cue conflict, as the depth cues of the pictorial medium in visual space – such as motion parallax and stereopsis – indicate that you are looking at a flat object, while simultaneously the depth cues of the pictorial depiction – such as the shading, linear perspective and object occlusion – give some indication of a depicted object's 3-D dimensions. The result of such a cue conflict could be cue integration, resulting in the perception of a depth that is somewhere between the depicted dimensions based on the pictorial array and the flatness of the picture plane. This mechanism could be thought of as twofoldness-induced depth compression, and such a compression could potentially bias orientation estimates of depicted objects.

Nanay (2011), a philosopher, suggests that the idea of twofoldness is supported by the respective functions of the dorsal and ventral neural pathways responsible for visual processing. Specifically, the dorsal pathway is involved with processing visual information which mediates the required sensorimotor transformations to interact with perceived objects, while the ventral pathway processes visual information to form an accurate perceptual representation of a target object based on its visual properties (Goodale & Milner, 1992). The independent functioning of these two distinct neural pathways is made evident through patients suffering from optic ataxia and visual agnosia. Optic ataxia, which is caused by lesions in the dorsal visual stream in the posterior parietal cortex, is characterized by patients' inability to use vision to physically interact with objects, despite being able to perceive identifiable qualities of those objects. Contrarily, visual agnosia, which is caused by lesions in the ventral visual processing stream in the lateral occipital complex, is characterized by patients' inability to perceive identifiable object properties despite not interferring with their ability to use vision to inform the sensorimotor transformations required to physically interact with those objects (Goodale & Westwood, 2004). Thus, Nanay (2011) theorizes that when viewing objects in pictorial space, the dorsal visual stream, which would normally govern processing of the information necessary to interact with those depicted objects, instead processes the information necessary to interact with the physical pictorial plane which inhabits the same visual space as we do. If this were the case, the dorsal and ventral visual streams might process conflicting information regarding a depicted object's physical dimensional properties, and this conflicting information could be integrated.

Twofoldness is also exemplified by humans' ability to compensate for the deformations of depicted objects which result from viewing pictures from oblique angles (Vishwanath et al., 2005; Nanay, 2011). Perceptual access to the orientation of the pictorial plane has constraining effects on the perceived rotation of depicted objects (Goldstein, 1987). This is likely associated with the scaling and foreshortening of depicted objects which results from viewing pictures from variable distances or oblique vantage points respectively. These geometric changes alter the perceived orientation of these objects, however, they do not affect the perceived spatial layout of the pictorial scene (Cutting, 1988; Goldstein, 1988; Vishwanath et al., 2005). Cutting (1988) suggests that the reason these geometric transformations affect the perceived orientation of pictorially depicted objects but not their perceived spatial layout or dimensions is likely due to the source of the respective percepts. The perception of a pictured object's orientation stems directly from the visual information provided by the optics of the picture as presented to the eye. Thus, the affine transformed orientation of depicted objects is interpretted directly from this visual information, and not derived from other sources, such as the orientation of the picture plane. Alternatively, the perceived spatial layout and dimensions of depicted objects can be derived from surface slant as well as the optical changes which accompany viewing pictures from extreme slants; these derived percepts remain relatively constant in the face of changing visual information (Cutting, 1988; Goldstein, 1988). Since the perception of orientation is not derived from sources other than the direct optical information presented to our eyes, the human visual system may be incapable of compensating for the integration of depth cues characteristic of twofoldness-induced depth compression when determining the orientation pictured objects.

Vantage Point

Another potential source of depth compression that is necessary to consider in the context of this thesis is the vantage point of the observer. Multiple researchers, such as Farber and Rosinski (1978), Hecht et al. (1999), and Vishwanath et al. (2005) report results that are demonstrative of depth compression resulting from changing the viewing location of the observer. The relationship between depth compression and vantage point has been specifically attributed to the viewer's displacement from a picture's center of projection (CoP) along the normal of the pictorial plane (Farber & Rosinski, 1978;

Vishwanath et al., 2005; Fig. 1.3). Depth compression theoretically results when an observer views a picture from a vantage point that is closer to the picture plane than the picture's CoP. This distortion would therefore be caused by the observer adopting an 'incorrect' vantage point which differs from the picture's CoP, and thus could be considered vantage-point-induced depth compression. As stated previously, depth compression is important to consider when analyzing the orientation bias phenomenon because it directly affects angular estimations in pictorial spaces by deforming depicted objects (Hecht et al., 1999; Koenderink & van Doorn, 2008; Fig. 1.2). Compressing an object in depth changes the spatial relationship between elements in the scene, which causes the slant of depicted objects to appear more oblique, thus potentially causing an overestimation of observed angles (Fig. 1.2).

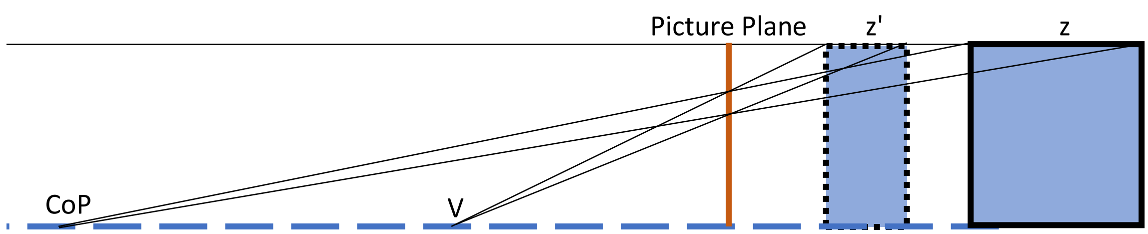

Figure 1.3: This diagram demonstrates how moving the vantage point (V) closer to the picture plane than the picture's center of projection (CoP) results in the perceived depth (z') being compressed relative to the original depth (z). This phenomenon results from geometry given that the optical arrays viewed from the vantage point must cross through the same points on the picture plane as those captured from the picture's CoP.

Occluding Contour Shift

Varying the viewing distance between observers and stimulus objects could also affect perceptions of the object's orientation in both visual and pictorial spaces. Specifically, changing the distance between the two can drastically shift the occluding contour of the objects, and this could potentially redefine the range of values considered when making an orientation estimate. This effect can be demonstrated simply using geometry (Fig. 1.4). When considering the orientation bias of faces,

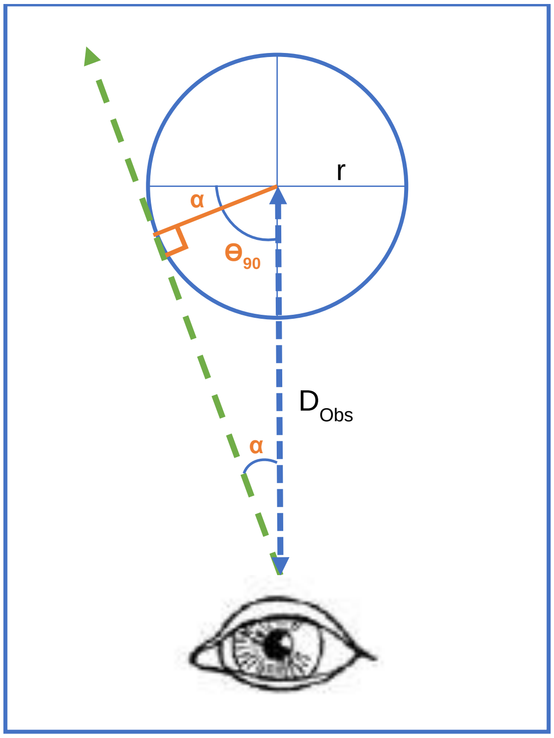

Figure 1.4: This diagram demonstrates the distance effect hypothesized under the assumption that the profile is defined by the occluding contour of the head. The green tangent line represents the observer's gaze, which intersects the edge of the circle (the head) defining the occluding contour. Thus, an observer may interpret a profile view when a head is oriented to Θ_{90}. This would cause the profile view to appear smaller than it really is ($\Theta_{90} < 90^\circ$), which could lead to misestimations of angles between frontal and profile. As distance (D_{Obs}) increases, the size of α decreases and Θ_{90} approaches 90˚. Thus, angular estimations between 0˚- 90˚ should become more accurate when distance increases.

let us approximate the shape of a head as cylindrical. In this scenario it is possible that observers may incorrectly assign a 90˚ designation to a face that only appears to be oriented such that it is viewed in profile. This could result from the location of the face's occluding contour – perhaps observers interpret a profile view as when the observed face is oriented perpendicularly to their line of sight (Troje et al., 2018; Fig. 1.4). If this were the case, then decreasing the distance between the observer and the stimulus would result in the angle at which the head is perceived to be in profile decreasing too. In the context of Fig. 1.4, as α increases, Θ_{90} decreases. This perceived profile would redefine the range between 0˚- 90˚ orientations, so it would necessarily affect the accuracy of angular estimates of any points in between the frontal and profile view, such as the perceived midpoint – the 45˚ orientation. If Θ_{90} is erroneously interpreted as being 90˚, then the relationship between the predicted angle (ϴ') and the presented angle (ϴ) can be modelled:

Eq.1:

$$\theta' = \theta \cdot \frac{90}{90 - \alpha} = \theta \cdot \frac{90}{90 - \arcsin\frac{r}{D_{Obs}}}$$

This equation demonstrates how the more the perceived profile differs from the true profile, the more angular orientation estimates will differ. A closer face would result in a larger angle α, thus causing more extreme overestimations of angular orientation. Alternatively, further stimuli should result in estimates which are approximately equal to the true orientation, because α approaches 0. This demonstrates that orientation bias caused by misperceptions due to an occluding contour shift would decrease as the distance between the observer and the stimulus increases.

Presence and Associated Depth Cues

Another factor to consider when examining the differences between visual perception in visual and pictorial spaces is the experience of presence. As the word ‘presence’ has multiple meanings attributed to it, we must first define it rigorously in the context of the present thesis. In discussing the sensation of presence in VR, Slater (2009) reserves the term ‘place illusion’ to specifically refer to the sensory experience of physically inhabiting a space. In this sense, presence is something that is frequently experienced in visual spaces, as the observer has a sense of place within such spaces – they have an innate understanding of their location in relation to other objects within the space they are inhabiting. Contrarily, it is unusual to feel a sense of presence within pictorially represented spaces as one cannot physically occupy a location within such spaces. The sensation of presence is entirely experiential in nature, and a strong enough feeling can catalyze physiological and psychological experiences which may not have occurred without it (Powers & Emmelkamp, 2008). Presence is also associated with emotional responses, training efficacy, and situational awareness (Satava & Jones, 1998; Powers & Emmelkamp, 2008). In visual spaces, the sensation of presence facilitates the distillation of visual sensory data into practical information that is used to help navigate physical environments and social situations. The visual system processes multiple different depth cues, such as motion parallax, stereopsis, linear perspective, texture gradients, shading, and occlusion, to provide the observer a

Figure 1.5: This image shows Uncle Sam's famous recruitment poster which, due to the lack of these depth cues, is perceived as pointing at the observer, regardless of their location in relation to the poster. Alternatively, if one were to imagine a statue of him pointing in a real, 3-dimensional scene, the new optic arrays would be subject to motion parallax and stereopsis as observers moved about to report Uncle Sam's orientation.

physical sense of place. In pictorial spaces, some of these depth cues are unavailable, most notably motion parallax and stereopsis which are critical for maintaining directionality – the ability to discern the alignment and orientation of objects in relation to their positions in space. The lack of these depth cues hinders the ability to feel a sense of presence in pictorial spaces.

Furthermore, recent literature is demonstrative of a paradigm shift which supports the notion that embodied cognition – the idea that the body or the body's interactions with the environment contribute to cognition – differentiates visual and pictorial spaces. Embodied cognition supports the idea that the sensation of presence, as well as our spatial perception, may in part be derived from the sensorimotor contingencies which are characteristic of the visuospatial modality (Nöe, 2006; O'Regan & Nöe, 2001). Motion parallax, which is defined as the interrelated movements of the retinal projections of elements in a scene that can occur when an observer moves relative to the scene, is a sensorimotor contingency that is maintained in visual space but not pictorial space. Take for instance Fig. 1.5, depicting Uncle Sam's famous recruitment poster. If you were to physically move, viewing the image from oblique angles, Uncle Sam will point at you regardless of your position. This is known as the Mona Lisa Gaze Effect, which states that pictorial objects oriented perpendicularly to the picture plane appear to rotate along with the observer as they move around the picture in visual space (Rogers et al., 2003). Another theory called the Differential Rotation Effect posits that objects oriented more perpendicularly to the picture plane appear to rotate less when viewed from different angles, and more parallel oriented objects appear to rotate more (Goldstein, 1987). Since Uncle Sam is pointing perpendicularly to the picture plane – as do Mona Lisa's eyes – the perceived view remains consistent of him pointing directly at the observer, regardless of their physical location. Since motion parallax is not a feature of pictorial spaces, moving side to side does not reveal new perspectives of the depicted elements, nor can the spatial or depth information provided by these movements be incorporated into our perception of the depicted scene. Alternatively, if an observer were to move around a statue of Uncle Sam in visual space, he would not be constantly pointing at them. Instead, the observer would be privy to new visual perspectives of the statue, gaining depth information of elements in their view, spatial relationships between parts of the statue as well as the statue relative to objects in its surroundings, and additional information on the dimensions of the statue. This is because motion parallax does not survive pictorial projection.

Koenderink and van Doorn (2008) suggest that the availability of depth cues, such as motion parallax, unique to visual spaces may be responsible for the discrepancies between orientation perception in visual and pictorial spaces. It is highly likely that stereopsis also plays a role in the differential perception of these spaces. Stereopsis is defined as the depth perception provided by means of the binocular disparity of the images in the two eyes. As such, stereopsis too is a feature of only visual space and not pictorial space, as pictorial depictions by nature are monocular representations of 3-D scenes. Depth perception, provided through binocular disparity or motion parallax, informs the creation of a more accurate 3-D reconstruction of a viewed object and could greatly improve observers' abilities to specify aspects of surface structure, such as orientation (Reichel et al., 1995).

Object Shape

The mechanism by which many of the above factors affect perceptions of orientation are intrinsically linked to the perception of a depicted object's shape. Twofoldness is a potential source of depth compression, which could cause distortions in the shape of the observed object and thus, affect its perceived orientation. In discussing twofoldness and the perceptual effect of alternate viewing angles on the depicted pictorial elements, it is also evident that the perceived orientation of an object is influenced by the geometrical changes which accompany these new viewing angles, such as scaling and foreshortening (Cutting, 1988; Goldstein, 1988; Vishwanath et al., 2005). Similarly, viewing pictorial objects from 'incorrect' vantage points could cause depth distortions through vantage-point-induced depth compression (Farber & Rosinski, 1978; Vishwanath et al., 2005). The shape of a viewed object also affects the location of its occluding contour, which could influence the perceived orientation of that object. The unique depth cues which contribute to the sensation of presence in visual spaces, namely motion parallax and stereopsis, theoretically improve the accuracy of orientation perception through enriching the understanding of the perceived object's true geometrical dimensions (Reichel et al., 1995; Koenderink & van Doorn, 2008). All these findings demonstrate that removing the geometric information of viewed objects through limiting the available depth cues, or altering their geometric information through visual compression, reduces the accuracy with which orientation estimates can be made. This may suggest that the robustness of the available geometric information of a viewed object could be imperative to the formulation of accurate orientation estimations.

The presumed geometrical properties of objects may also affect their visual perception. This idea would suggest that when viewing familiar objects – such as faces – or objects which follow certain geometric rules – such as boxes which have right angled corners and are symmetrical – observers rely on their developed geometrical prior for those objects to inform their judgements (Vishwanath et al., 2005). For example, people experience discomfort living in trapezoidal rooms because the prior of rooms being rectangular dominates the true form as indicated by the motion parallax and optic flow (Crunelle, 1996; Cornilleau-Pérès, 2002). In pictorial spaces, these expectations could allow viewers to tolerate noticeable distortions of viewed objects, such as depth compression, scaling, and foreshortening (Perkins, 1973; Busey et al., 1990; Vishwanath et al., 2005). Additionally, the viewpoint-from-above assumption, which describes the human tendency to consider objects as if viewed from above, supports the notion that depth information from an existing shape prior could be incorporated into this top-down representation and used to inform angular orientation estimates (Cornelis et al., 2016). Of course, the existence of shape priors makes it possible that having an incorrect prior can affect the accuracy of judgements which draw upon it. Therefore, the provision of accurate and abundant shape information of viewed objects should inform orientation judgements and mitigate potential biases, such as the orientation bias phenomenon.

Overview

The aim of this thesis is to explain the facial orientation bias phenomenon. The discussed literature provided multiple potential factors to consider which are likely contributors to this phenomenon. The purpose of the following experiments was to rigorously test each of these factors to determine their respective impacts, if any, on the experience of orientation bias. The first experiment tested whether orientation bias is unique to pictorial projections due to their twofoldness. Specifically, it tested whether twofoldness-induced depth compression is the primary contributor. This experiment used virtual reality to compare the magnitude of orientation bias based on whether the stimulus was presented in a simulated visual or pictorial space through manipulating the availability of motion parallax and stereopsis. This experiment also examined the individual effect of both aforementioned depth cues on the perception of orientation, as well as the effect of changing the perceived distance between an observer and the stimulus object in both types of spaces. The second experiment tested whether vantage-point-induced depth compression is responsible for the experienced orientation bias in pictures. It also examined whether retinal image size, which is a potential source of geometric information and detail, plays a role in the magnitude of this phenomenon. The final experiment was motivated by the difference in the size of the bias in the previous two experiments. It was hypothesized that the method of adjustment employed in the first experiment allowed for participants to perceive shape-from-motion, and this mitigated the size of the bias compared to the second experiment. This experiment tested whether humans adopt an inaccurate shape prior for human heads in the absence of sufficient geometric information, such as when presented in pictures. This experiment tested the impact of object shape by using an object with an existing cylindrical shape prior, a coffee mug, and warping its dimensions such that its width-depth ratio would match that of a typical human head. Additional geometric information was also added through motion, which could in theory demonstrate whether providing additional structure through motion could mitigate the strength of orientation bias.

Experiment I: Visual Condition as a Predictor for Orientation Bias

Motivation and Hypotheses

We expect that orientation bias might exist solely in pictorial spaces and not visual ones due to the conflicting depth cues which may cause twofoldness-induced depth compression. The approach taken by the current experiment is meant to replicate Troje et al.'s (2018) results, which demonstrated that the visual condition – whether a face was presented to observers in visual or pictorial space – had a significant effect on the perceived orientation. Specifically, mean 45˚ estimations of faces were 41.15˚ in visual space and 38.63˚ in pictorial space, and these means were significantly different with $F(1, 38) = 67.44$, $p < .001$ (Troje et al., 2018). However, though these results highlighted a significant difference caused by the viewing condition, it failed to expand upon the mechanism for such a difference. The current experiment attempts to address this by dividing the two visual conditions used by Troje et al. (2018) by their component depth cues – motion parallax and stereopsis – which differentiate them. This would create four unique visual conditions: Window (both motion parallax and stereopsis enabled), Motion Parallax Only (stereopsis is disabled), Stereopsis Only (motion parallax is disabled), and Picture (both motion parallax and stereopsis are disabled). The addition of these intermediate conditions could provide a more comprehensive understanding of the perceptual mechanisms which differentiate orientation perception in visual and pictorial spaces and potentially elucidate ways to mitigate the orientation bias phenomenon. The current experiment will test three hypotheses:

H1.1: The orientation bias phenomenon stems from differences in the visual condition – whether the scene is presented in 3-D visual space or 2-D pictorial space – and is only expected in pictures due to twofoldness-induced depth compression.

H1.2: Differences in orientation perception are influenced independently by the available depth cues: Motion Parallax and Stereopsis.

H1.3: The magnitude of the orientation bias should decrease as the perceived viewing distance between the observer and the stimulus object increases due to the occluding contour shift.

H1.1 is supported by the results of Troje et al. (2018), but also justified based on depth compression literature such as Hecht et al.'s (1999) finding of angles appearing more obtuse when compressed in pictorial space. H1.2 is supported by Koenderink and van Doorn (2008), Reichel et al. (1995), and Cornilleau-Pérès (2002) who demonstrated the dominance of static grid cues, which represent stereoscopic depth, over motion parallax in indicating plane tilt. Finally, H1.3 is supported by the geometry which could define orientation based on the occluding contours of faces (Fig. 1.4).

This study was conducted in virtual reality. Participants were placed in a virtual room, where they stood on a marked circle facing a frame on the wall, dubbed the "Alberti frame". The Alberti frame changed its behavior between trials, sometimes acting as a window, sometimes acting as a picture, and sometimes acting as an intermediate condition (Motion Parallax Only or Stereopsis Only). Each trial, participants were presented with a head on a pedestal facing a random direction and were instructed to adjust the head such that it was facing 45˚ away from the frontal position. Their answers indicated the angles which they perceived as 45˚ in each condition, thus allowing the assessment of the effects of each variable as well as interactions between them on the bias phenomenon we are attempting to explore.

Methods

Participants

There were 22 participants (5 male: 17 female) who were randomly recruited for this study through York University's research participant pool (URPP), whose mean age was 20.73 ± 4.32 SD, median age 19. All participants had normal or corrected-to-normal vision. Upon arrival to the lab, participants read and signed an informed consent form, detailing the study and ensuring that it was being carried out in accordance with York University's regulations. Participants received 1.5 URPP credits as compensation for their participation in the study.

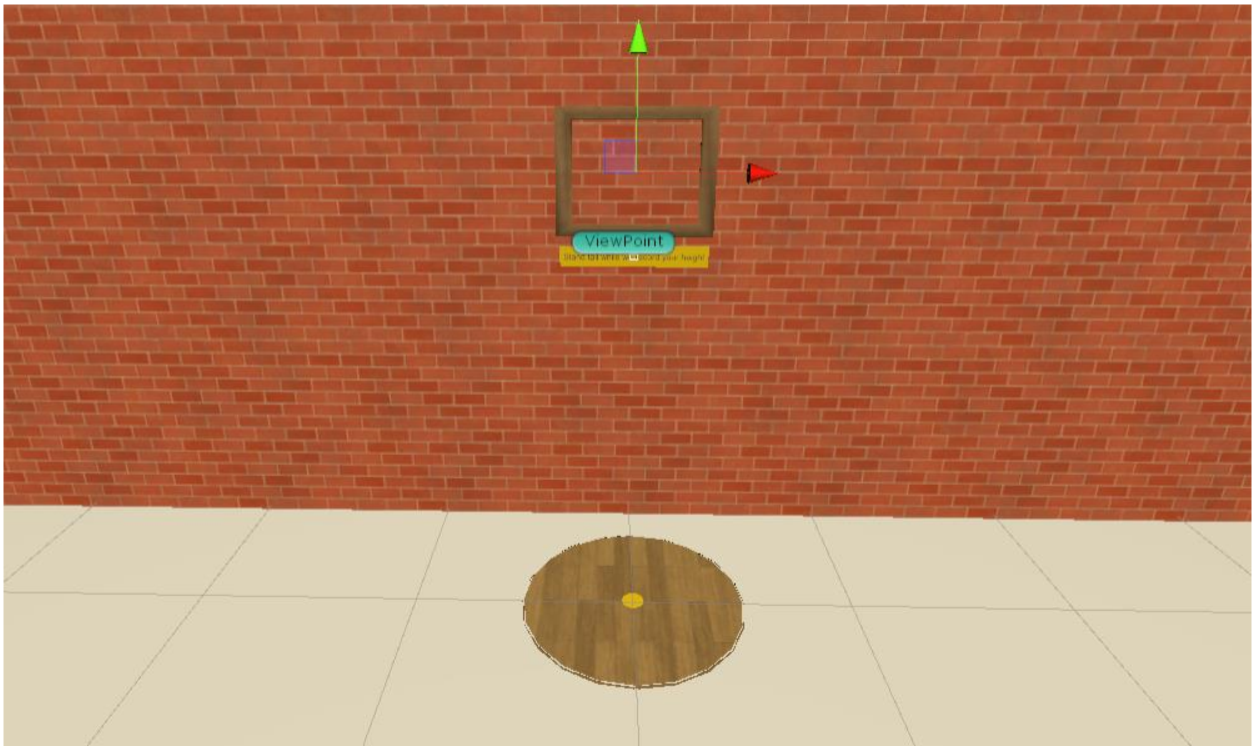

Figure 2.1*:* This image depicts the virtual space occupied during the experiment. Participants were instructed to stand on the platform and look towards the wall. By clicking a button, the Alberti frame would snap to eye-level. The 'Alberti frame' has the capacity to act as a picture or a window, as well as two intermediate conditions by allowing what is behind it to be viewed with or without motion parallax and stereopsis. Artifacts such as the 'viewpoint' bubble or the red and green arrows were not seen by participants in virtual reality as they are only visible from the Unity3D editor window.

Materials

The virtual environment the study was conducted in was generated using Unity3D, a 3-D game engine. Participants wore an HTC Vive head-mounted display (HMD) and were also given a Vive controller to control the stimuli's orientation throughout the experiment. During the experiment, the participants were instructed to use the circle-pad on the controller to turn the stimulus left or right, and once they believed that they had oriented the stimulus halfway between frontal and profile view, which should correspond to a 45˚angle, to pull the controller's trigger to record their choice and advance to the next trial. When the experiment began, the participant was placed in a virtual brick room. On the floor was a wooden platform, indicating where they should stand. 70 cm in front of them was a wall with a wooden frame – the Alberti Frame – which controlled the visual condition under which the stimulus was viewed and was set to the height of the participant upon beginning the experiment (Fig. 2.1). The stimuli were generated behind the frame.

Stimuli

The stimuli objects consisted of six different heads which were obtained from Troje et al.'s (2018) study. These heads were created using FaceGen Modeler Core 3, a software application which imports photographs of real humans at frontal and profile positions and allows the user to place points of reference on each photograph. The program then uses this data to generate a 3D model of the face that can be exported as an .fbx file, which is compatible with Unity3D. These heads were created to be life-sized and have realistic dimensions – their mean x, y, and z dimensions being 16.1 cm, 24.3 cm, and 21.8 cm respectively, which did not differ significantly from the mean head dimensions of 20 randomly selected volunteers (Troje et al., 2018). In each trial, one of these heads would be placed on a pedestal, either 1 m, 1.5 m, or 2 m away from the participant. The heads were initially rotated about their vertical axis to be presented at a randomly generated angle between 10˚ and 90˚ left or right of a frontal view. The pedestal was cylindrical to ensure that participants could not use any corners or vertices to estimate the correct angular orientation of the stimuli (Fig. 2.2). The stimuli were generated the same way in all four visual conditions, the difference being in how the Alberti frame operated, either providing both stereopsis and motion parallax, providing either depth cue in isolation, or providing neither of them, in each of the respective viewing conditions.

Figure 2.2: The stimuli presented to the observer were 6 computer generated head models. These heads were displayed on round pedestals, such that there were no directional cues attached to the shape of the pedestal. The backdrop for the heads was another brick wall, the texture of which was meant to increase the effect of motion parallax. Blue highlight and directional arrows are artifacts of viewing in the Unity3D editor window and were not visible to participants during the experiment.

Figure 2.3: Sample images of the 6 computer generated head models which were presented to the observers during the experiment.

Design

There were four visual conditions (Window, Stereopsis Only, Motion Parallax Only, and Picture), three distances (1 m, 1.5 m, and 2 m), six different heads (1-6), and two viewing directions (left and right) for a total of 144 unique trial conditions. The visual conditions were blocked, and block order was randomized within each 4-block package – 1 for each visual condition – to mitigate any potential order effects. Each block was completed four times, so a grand total of 576 trials were recorded. The study used a within-subjects design.

Procedure

Before starting the experiment, participants were fitted with an HMD, shown how to use the Vive controller to orient the stimuli, and told roughly where to stand for the experiment. They were given 5 sample trials to learn how to use the controller and become accustomed with the task of orienting the stimulus to the 45˚ goal. Participant age and gender were recorded, and the experiment began. Participants were instructed to stand on the virtual platform, face the frame, and stand tall while their height was recorded by the experimenter. The frame and the stimulus then snapped to the participant’s eye-height, such that results would be consistent between participants regardless of height. At the beginning of each trial, the orientation of the stimulus was randomized between 10˚ and

90˚ in the viewing direction to ensure that the participant must manually reorient the stimulus each trial using the method of adjustment. Furthermore, stimulus rotation was locked between 0˚ and 90˚ in the correct viewing direction, meaning it was impossible for participants to mistakenly orient the stimulus in the wrong viewing direction. Participants used the controller's circle-pad to adjust the stimulus' orientation and pulled the controller's trigger when they believed they had oriented the stimulus to the intended 45˚ angle. Participant responses were recorded in a spreadsheet, which also contained all the input information, and the difference between the 45˚ target angle and the participant's selected angle was calculated to determine the error in each trial, which could be influenced by the phenomenon of interest – namely, orientation bias. Upon experiment completion, a message was displayed on a plaque below the Alberti frame urging the participants to notify the experimenter that they had finished. They were then awarded their URPP credits for their participation and were thanked for their participation. The experiment took an average of 30-40 minutes to complete.

Data Analysis

The study was analyzed using a repeated measures design and compared the scores of participants' trials in each of the four visual conditions to assess our hypotheses. Three discrete, within-subjects, independent variables were examined: visual condition, head shape, and distance of the heads from the observer. The head shape variable was also broken down in terms of the heads' X/Z aspect ratios, which were determined by measuring using an object with known dimensions to get the width of the skull (ear-to-ear) and the length of the skull (brow-ridge-to-back); the recorded X/Z aspect ratio reflects the average between two measurements, one with and one without hair on the model head. There was one continuous dependent variable, being the recorded angle participants perceived as 45˚. Using R statistical software, a three-way repeated measures ANOVA was conducted to determine the effects of the three independent variables and whether there were any statistically significant interactions between these factors.

Results

Orientation bias can be calculated by comparing the participants' judgements in the 45˚ estimation task – the set angle – with the true 45˚ target angle they were instructed to orient the stimulus to. A set angle below 45˚ indicates an overestimation of the observed head's angular orientation, while alternatively, a set angle greater than 45˚ would represent an underestimation of the observed head's angular orientation. A three-way repeated measures ANOVA was conducted to verify whether there were significant effects or interactions resulting from the three independent variables which could have contributed to the phenomenon of orientation bias. The three independent variables were the distance (1m, 1.5m, 2m), head number (1-6), and the four visual conditions. Additionally, a one-sample, two-tailed t-test was performed to determine whether the hypothesized orientation bias existed at all.

Experiment I: Mean Orientation Bias Three-Way Repeated Measures ANOVA

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
VC	3	63	1.903	1.38E-01		2.00E-03
D	2	42	5.126	1.00E-02	*	2.00E-03
H	5	105	3.582	5.00E-03	**	4.00E-03
VC:D	6	126	2.297	3.90E-02	*	7.96E-04
VC:H	15	315	0.731	7.53E-01		7.27E-04
D:H	10	210	2.963	2.00E-03	**	1.00E-03
VC:D:H	30	630	1.032	4.20E-01		2.00E-03

Table 2.1: Analysis of variance demonstrating whether the explored factors – visual condition (VC), observer-head distance (D), or specific head observed (H) – affect the magnitude of the observers' orientation bias, by themselves or through interactions, when viewing a computer-generated head. Single, double, and triple asterisks corelate to p<.05, p<.01, and p<.001 respectively. The ges represents the generalized eta squared.

The result of the t-test indicated that there was a statistically significant orientation bias which operated as described – a systematic overestimation of angular orientation (*n* = 22, M_{Bias} = 3.7˚ [41.3˚ guessed as 45˚], *p* = .021*). This result confirms the existence of the orientation bias phenomenon, though the magnitude of the bias effect is smaller than expected.

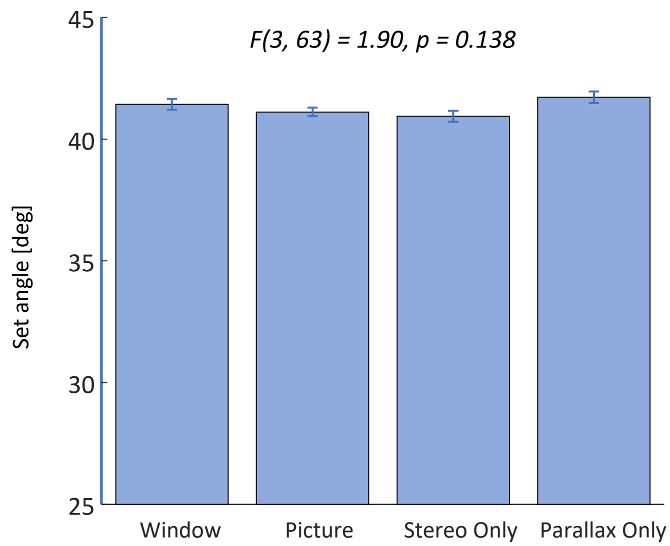

Figure 2.4*:* This bar graph depicts the set angles which were interpreted as 45˚ in each of the four visual conditions across all n = 22 participants. Error bars represent the standard error of the mean (SEM). These results do not seem to indicate any significant effects of visual condition alone.

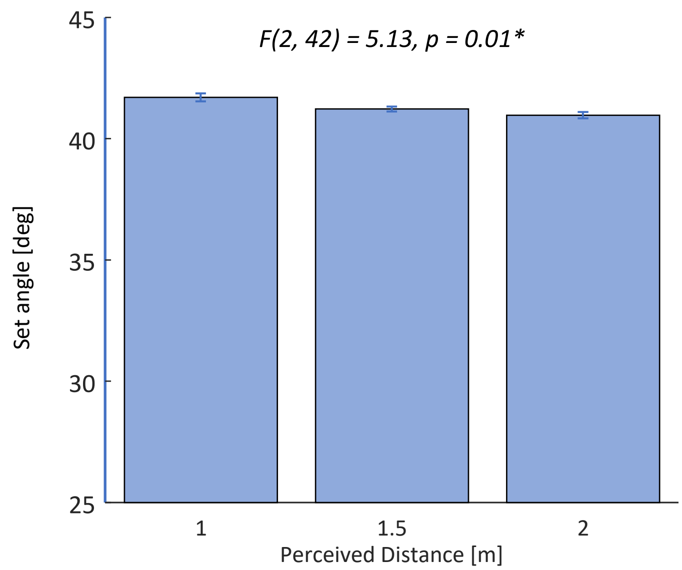

Figure 2.5*:* This bar graph depicts the set angles which were interpreted as 45˚ for each of the three distances. Error bars represent the standard error of the mean (SEM). These averages span every visual condition and head for each of the n = 22 participants. There is a slight positive trend between the perceived distance between the observer and the stimulus and the magnitude of the bias. One reason why this effect might appear smaller could be due to half of the visual conditions contributing to this average lacking stereopsis, and naturally the effects of distance should be less pronounced in those conditions which lack depth perception.

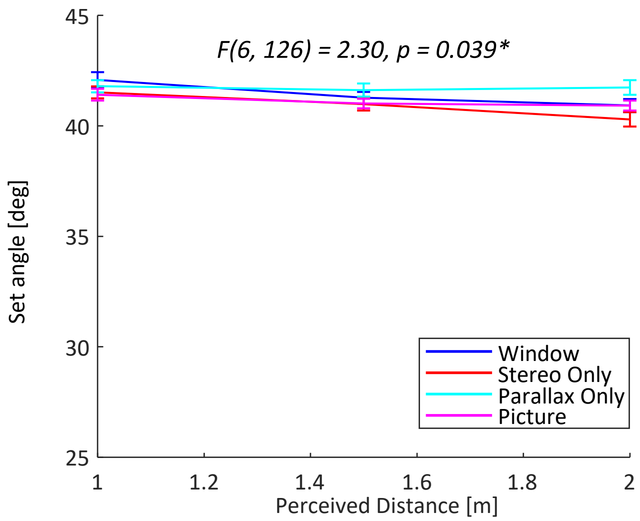

Figure 2.6*:* This graph demonstrates how the interaction between visual condition and the perceived distance between the stimulus and the observer affects the angular overestimation characteristic of orientation bias. Error bars represent the standard error of the mean (SEM). This overestimation was recorded across all n = 22 participants. The effects of distance can be observed by the negative slope in both the window and stereopsis only conditions. This distance-dependence appears to be virtually non-existent in the picture and parallax only conditions. Additionally, the two conditions with motion parallax (window and parallax only) had consistently lower observed biases than those without.

The current study indicated that there were no significant differences attributed to the visual condition alone (p = .138; Tab. 2.1). The average bias scores of participants (n = 22) in the four visual conditions (M_{Win} = 3.58˚, M_{SO} = 4.06˚, M_{PO} = 3.28˚, M_{Pic} = 3.89˚; SD_{Win} = 9.72˚, SD_{SO} = 9.65˚, SD_{PO} = 9.91˚, SD_{Pic} = 9.62˚) were all within one degree of each other (Fig. 2.4).

The results of this study also identified the effect of distance on orientation bias as statistically significant, (p = .010; Tab. 2.1; Fig. 2.5). The mean bias seemed to gradually increase with the distance (M_{1m} = 3.30˚, $M_{1.5m}$ = 3.73˚, M_{2m} = 4.08˚; SD_{1m} = 9.49˚, $SD_{1.5m}$ = 9.83˚, SD_{2m} = 9.82˚). Furthermore, there was a minor but statistically significant interaction effect between the visual condition and the distance (p = .039; Tab. 2.1; Fig. 2.6). The effect of distance is clearly pronounced in the conditions which have stereopsis. However, it seems to be less pronounced or vanish completely in the Motion Parallax Only and Picture conditions.

To isolate the effects of motion parallax and stereopsis on the orientation bias, an additional three-way repeated measures ANOVA was performed to test for significant interactions between stereopsis (On/Off), motion parallax (On/Off), and the simulated distance (1m, 1.5m, 2m). This analysis revealed that motion parallax on its own was a significant factor, and it appears that when motion parallax is enabled, the magnitude of the orientation bias is diminished (*p* = .016; Tab. 2.2; Fig. 2.6). Motion parallax does not appear to interact with distance. In contrast, even though stereopsis on its own was not significant, the interaction between stereopsis and head distance did have a statistically significant effect on the mean bias. It appears that stereopsis, when enabled, increases the magnitude of the bias with increasing distance (*p* = .007; Tab. 2.2; Fig. 2.6).

Experiment I: Three-Way Repeated Measures ANOVA of Stereopsis, Motion Parallax, and Distance on Orientation Bias

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
St	1	21	0.516	4.81E-01		2.92E-04
MP	1	21	6.917	1.60E-02	*	2.00E-03
D	1.58	33.28	5.126	1.70E-03	*	2.00E-03
St:MP	1	21	0.094	7.63E-01		1.82E-05
St:D	2	42	5.617	7.00E-03	**	7.39E-04
MP:D	2	42	0.800	4.56E-01		8.06E-05
St:MP:D	2	42	0.381	6.86E-01		5.69E-05

***Table 2.1**: Analysis of variance demonstrating whether the explored factors – stereopsis (St), motion parallax (MP), or observer-head distance (D) – affect the magnitude of the observers' orientation bias when viewing a computer-generated head and whether there are any significant interactions between these factors. Single, double, and triple asterisks corelate to *p<.05, p<.01,* and *p<.001* respectively. The ges represents the generalized eta squared.

The results reported in Tab. 2.1 also indicated that the specific shape of the head observed was significant (*p* = .005; Fig. 2.7; Fig. 2.8). Head 6 had the least bias, while heads 2 and 5 had the most on average (M_1 = 3.94˚, M_2 = 4.10˚, M_3 = 3.80˚, M_4 = 3.44˚, M_5 = 4.09˚, M_6 = 2.84˚; SD_1 = 9.22˚, SD_2 = 9.61˚, SD_3 = 9.31˚, SD_4 = 9.64˚, SD_5 = 9.98˚, SD_6 = 10.51˚). One objective factor in which the heads differed was the ratio of their x-width to their z-depth. Replotting the head shape results in terms of their aspect ratios results in an apparent linear relationship (Fig. 2.8). The specific effect of head shape is further evidenced by the vertical separation between lines depicted in Fig. 2.9, which represents the interaction between head shape and stimulus distance. However, the significance of this result appears to be due to head 5 acting as an outlier compared to the rest of the head models used, as evidenced by its considerable slope (p = .002; Tab. 2.1; Fig. 2.9).

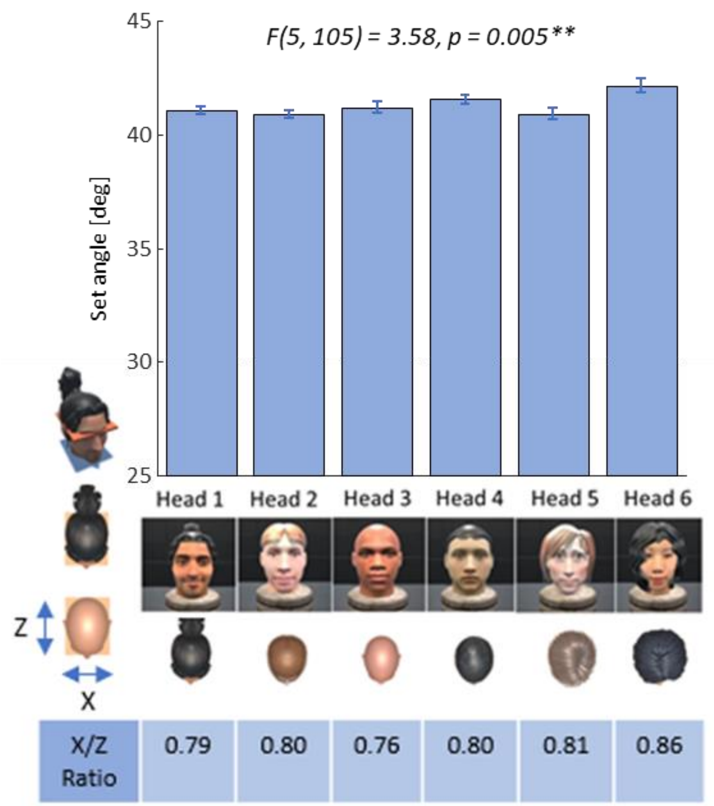

Figure 2.7: The above figure illustrates the average set angle perceived as 45˚ for each of the six different faces regardless of visual condition and viewing distance. Error bars represent the standard error of the mean (SEM). This figure shows how head 6 had the least bias (M_6 = 2.84˚ angle). The faces are arranged such that they are directly below their respective X/Z aspect ratios, which were determined by averaging their maximum width and depth with and without hair, as shown above.

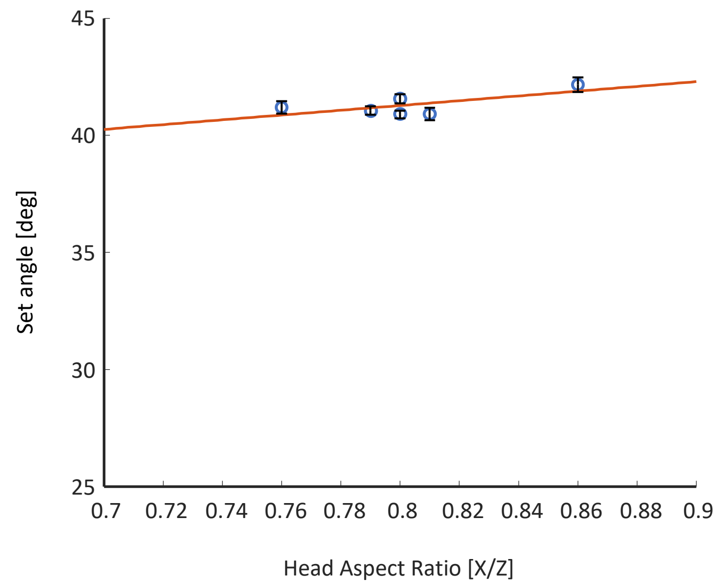

Figure 2.8: The relationship between the aspect ratio of presented stimuli and the set angle interpreted as 45˚, based on results from the current study. Error bars represent the standard error of the mean (SEM).

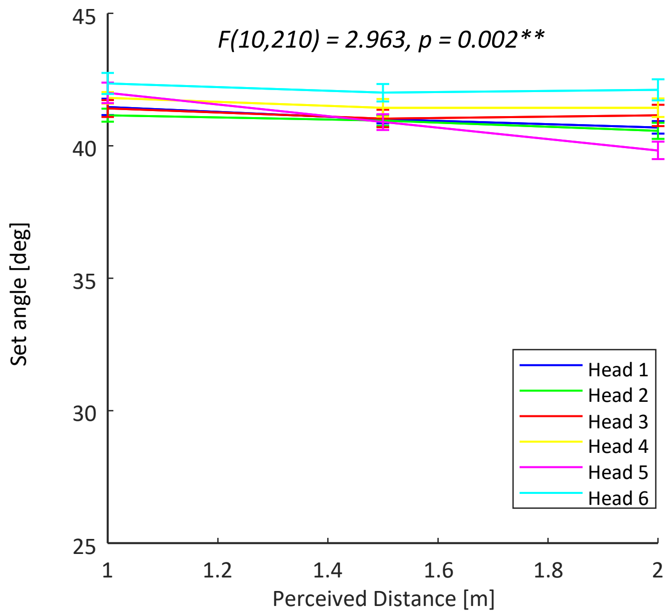

Figure 2.9*:* This figure illustrates the set angles resulting from the participants' angle estimations for each head shape at each distance, regardless of visual condition. Error bars represent the standard error of the mean (SEM). Head numbers match those in Figure 1.7. Head 5 seems to have an unusually steep slope compared to the other head shapes, indicating that distance more strongly influenced the participants' orientation bias for this head. All other heads show a very slightly negative slope, which follows the main effect of the distance, as previously reported by Troje et al. (2018). The separation of lines supports the notion that the face shapes are a significant factor to the magnitude of the orientation bias.

Discussion

This experiment followed the methodology of Troje et al. (2018) and extended upon it, adding two intermediate conditions between visual and pictorial spaces – being the Motion Parallax Only and Stereopsis Only conditions. However, some minute differences may contribute to discrepancies between the two studies' results, such as the virtual environment, the virtual camera's position, or the range of distances tested. This study's one-sample, two-tailed t-test was essential to first determine the existence of this orientation bias phenomenon. Though the results of the t-test confirmed a significant bias effect (n = 22, p = .021), the magnitude of the bias recorded (M_{Bias} = 3.7˚) is approximately half that reported by Troje et al. (2018). Additionally, this t-test considered the bias across all visual conditions, so it confirms that orientation bias exists, but not that it exists exclusively in pictorial depictions. Interestingly, the results of the three-way repeated measures ANOVA (Tab. 2.1) support that there actually was no significant main effect of the visual condition on the accuracy of observers' orientation estimates (p = .138). Thus, the core hypothesis, H1.1 – that the bias phenomenon is brought on by differences between visual conditions and should only be present in pictures due to their twofold nature – is not supported by the results of this experiment. This result does not match the anecdotal experience of facial orientation appearing skewed away from the observer to a slightly exaggerated degree (Fig. 1.1). Additionally, this result did not reach statistical significance, which contradicts the previously reported differences from Troje et al. (2018), who found quite a strong effect ($p < 0.001$) indicating that the bias is stronger in pictorial space ($M_{Pic} = 6.37˚$) than it is in visual space ($M_{Vis} = 3.85˚$).

Though there were no explicit hypotheses made regarding the main effect of the shapes of the heads used, the significance of this factor ($p = .005$; Tab. 2.1) on the magnitude of the orientation bias was expected, as it was consistent with the result reported by Troje et al. (2018), $F(5, 190) = 32.67$, $p < .001$. This shape-specific effect appears to be related to the aspect ratio between head-width and head-depth (Fig. 2.7; Fig. 2.8). As the aspect ratio increases, so does the accuracy of the 45˚ estimation as demonstrated by the bias decreasing.

In terms of the main effect of observer to stimulus distance, Troje et al. (2018) did not find any significance, $F(3, 144) = 2.20$, $p = .09$. However, the current study did find a significant effect of distance ($p = .01$; Tab. 2.1) which behaved opposite our prediction, indicating that the orientation bias increased congruently with distance. Therefore, our tertiary hypothesis H1.3 – that the shift in the occluding contour of the faces redefines the range of angles between frontal and profile, increasing bias as distance decreases – is not supported by the results of this experiment either.

Additionally, there was a significant interaction effect between the distance and the visual condition (p = .039; Tab. 2.1; Fig. 2.6). The information provided by Fig. 2.6 appears to support that stereopsis is responsible for the bias's susceptibility to distance effects, reliably increasing the magnitude of the bias with distance. One could also infer that perhaps motion parallax plays a role in decreasing the magnitude of this bias, regardless of distance. Every point in Fig. 2.6 with motion parallax has a lower bias than those conditions which lack it. The second three-way ANOVA (Tab. 2.2) which separated the contributions of motion parallax and stereopsis on the observed bias revealed a significant main effect of motion parallax (p = .016) but not of stereopsis (p = .48). However, the interaction between stereopsis and head distance did have a statistically significant effect on the mean bias (p = .007). These statistics further support Fig. 2.6 and confirm H1.2 by demonstrating how motion parallax and stereopsis influence the magnitude of the bias phenomenon differentially; stereopsis interacts with distance, increasing orientation bias as viewing distance increases, whereas motion parallax mitigates the bias to some degree, but has no interaction with distance. While these distinctions provide additional insight into the contributions of motion parallax and stereopsis, they fail to completely explain the orientation bias experienced when viewing faces in pictorial space and thus this matter requires further investigation.

Overall, our results indicate that we must reject H1.1 and H1.3, meaning that neither twofoldness-induced depth compression nor occluding contour shift are responsible for the experienced orientation bias phenomenon when viewing pictorial faces. Despite this, a small orientation bias was still observed (M_{Bias} = 3.7˚, p = .021), demonstrating that there must be another source for this bias phenomenon.

Experiment II: Depth Compression generated through Vantage Point Discrepancy as a Predictor for Orientation Bias in Pictures

Motivation and Hypotheses

Experiment I failed to confirm the twofoldness-induced depth compression hypothesis, which predicted significant differences between the visual and pictorial space conditions. It additionally failed to produce the predicted effects resulting from the shift in occluding contours as observer distance from the stimuli changes. However, there is another potential source of depth compression besides twofoldness which might contribute to orientation bias. Experiment II tested whether vantage-point-induced depth compression can produce the bias effect being examined. If this were the case, Experiment I's failure to produce a more substantial bias may have stemmed from the experiment set-up, and in particular the Alberti frame itself. As previously mentioned, the Alberti frame can act as a window, a picture, or an intermediate through enabling and disabling motion parallax and stereopsis. However, the mechanism by which this works requires the location of the camera to provide a viewpoint, and the camera used perfectly overlaps with the location of the HMD and acts as the observer's eyes. This set up may have been unable to produce the changes expected in bias because, according to Farber and Rosinski, depth compression in pictorial space is caused by 'incorrect' viewing locations in the visual space (Farber & Rosinski, 1978; Vishwanath et al., 2005).

Farber and Rosinski (1978) introduce the terms 'pictorial array' and 'environmental array' to distinguish between the array produced by the picture and that produced by the environment. When the viewing location of the observer overlaps with the center of projection (CoP) – the exact location from which the picture was taken relative to the 2-D projection – of the viewed image, the perceived pictorial array is geometrically identical with the environmental array characteristic of visual space, causing no differences to be observed (Farber & Rosinski, 1978). Experiment I was set up this way so that it could effectively examine whether twofoldness-induced depth compression contributed to orientation bias without conflating multiple potential causes of compression, as regardless of visual condition, the pictorial arrays and visual arrays would be virtually identical. Now that we have established through Experiment I that twofoldness-induced depth compression does not appear to contribute to orientation bias, we can examine vantage-point-induced depth compression. To test this, Experiment II was conducted in accordance with the depth compression model published by Farber and Rosinski (1978).

Background

Objects displayed in pictorial space are subject to certain geometric transformations based on the observer's location in the environmental space relative to the picture's CoP – however, they may only be perceived as distorted if the observer's visual system interprets the retinal image as being generated by a 3-D world. As stated above, an observer viewing a scene from the CoP would perceive a pictorial array as geometrically identical with the 3-D visual array, however, straying from this location would cause the represented pictorial space to become distorted (Farber & Rosinski, 1978; Vishwanath et al., 2005). For normal displacements, or displacements perpendicular from the surface of the picture plane, the represented space becomes compressed or expanded in depth by a compression factor *C*, which scales in relation to the magnitude and the direction of the displacement (Farber & Rosinski, 1978). *C* represents the expected factor by which the depth of a depicted scene in an image is either compressed or expanded relative to the original. According to Farber and Rosinski's reasoning, the perceived depth of a transformed picture should be the original depth of the 3-D object multiplied by *C*.

Eq.2:

$$C = \frac{D_{Obs}}{D_{CoP}} = \frac{Depth_{Pic}}{Depth_{3D}} = \frac{\tan\theta}{\tan\theta'}$$

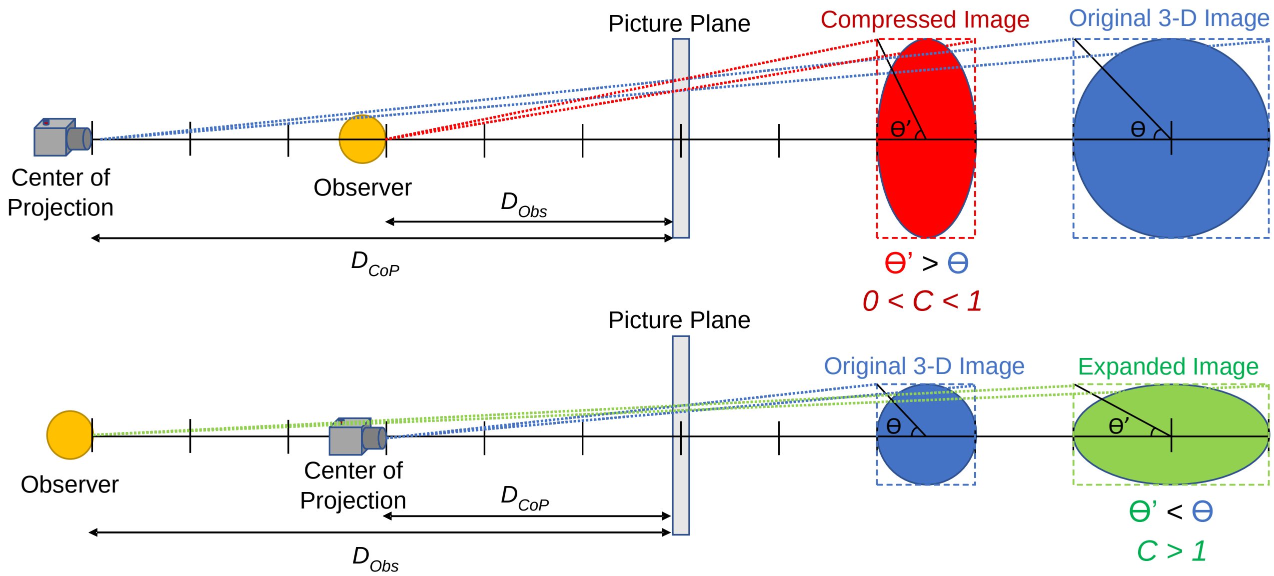

Figure 3.1: Explanatory graphic demonstrating how optical arrays crossing through the same points on the picture plane causes perceived depth compression and expansion when displaced from the CoP along the normal of the picture plane. Viewing the scene from a top-down perspective, circles are used as approximations of the heads (simplified to cylinders) to be used in the present study. In this example, when the observer stands between the 'correct' CoP and the picture plane, $0 < C < 1$ and the compressed image is shallower than the original. This compression causes surface slants within the altered image (ϴ') to appear more obtuse than in the original image (ϴ). Contrarily, when the observer is farther from the picture plane than the 'correct' CoP, $C > 1$ and the expanded image is deeper than the original. The resulting expansion causes surface slants within the altered image (ϴ') to appear more acute than they did in the original (ϴ).

If an observer were to stand between the CoP and the picture plane, it would result in $0 < C < 1$, which should create the perception of depth compression (Eq.2; Fig. 3.1). Alternatively, if an observer were to be positioned behind the CoP, it would result in a $C > 1$, which should be perceived as depth expansion (Eq.2; Fig. 3.1). Lastly, if an observer were to view an image from the CoP, it should result in a $C = 1$ and the pictorial array would be identical to the 3-D visual array, thus the perceived image depth ought to be unchanged.

Furthermore, compressing and expanding the represented space should also affect any representations which depend on spatial relationships, such as slant. Surface slants viewed in a compressed space should be biased towards the frontal plane, while slants viewed in an expanded space ought to be biased away from the frontal plane. Farber and Rosinski (1978) found the relationship between *C* and *Θ*, where *Θ* is the orientation angle of the original object and *Θ'* is the angle as perceived from the observer's location (Eq.2).

Head Image Size (m)	D_{3D} (m)	D_{CoP} (m)	D_{Obs} (m)	$D_{CoP} - D_{Obs}$ (m)	C Factor	Camera FoV (˚)	Observer FoV (˚)	FoV Deviation (˚)	Predicted Degree (˚)	Presented Degree (˚)
0.075	1.6	0.4	0.6	-0.2	1.5	10.71	7.15	3.56	45	56.31
0.075	2.4	0.6	0.6	0	1	7.15	7.15	0	45	45
0.075	3.6	0.9	0.6	0.3	0.67	4.77	7.15	2.38	45	33.69
0.15	0.8	0.4	0.6	-0.2	1.5	21.24	14.25	6.99	45	56.31
0.15	1.2	0.6	0.6	0	1	14.25	14.25	0	45	45
0.15	1.8	0.9	0.6	0.3	0. 67	9.53	14.25	4.72	45	33.69

Table 3.1*:* Table demonstrating presented head image sizes and the distances (in meters) used to generate our predictions. D_{3D} represents the distance between the CoP and the head in 3D space. D_{CoP} and D_{Obs} represent the distances between the CoP and the picture plane, and the observer and the picture plane respectively. These were the values used for stimulus generation, as they generate *C* factors which, using the current model, predict significant changes in perceived orientation. The heads (in the virtual environmental space the pictures were taken in) were approximately 0.3 m tall. The field of view (FoV) represents the angle (˚) subtended by the head as viewed from the location of the camera or the observer, respectively, and is included to demonstrate the differences generated from testing multiple image sizes, despite the prediction not changing.

Prediction Model

Using *Eq.2*, where *Θ'* is the perceived angle of the stimulus by the observer and *Θ* is the presented, or "true", orientation of the stimulus, we could solve for how a participant might perceive specific presented angles at any chosen compression factor. For the present study, we are interested in the participants' ability to perceive an orientation of 45˚ to the left of the frontal view, so by isolating for *Θ*, we can determine which "true" presented orientation would cause a perception of 45˚ (Tab. 3.1), or any other angle at a given *C*.

The works of Farber and Rosinski (1978) suggest that the orientation bias phenomenon could be generated by the spatial relationship between the observer and the CoP relative to the picture plane,

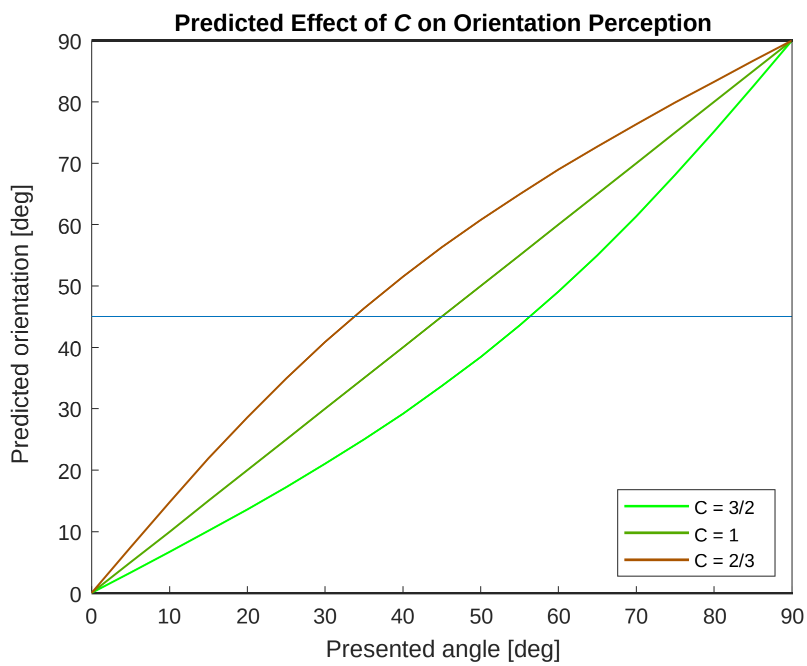

Figure 3.2*:* This image was generated using a 0.3 m head size, a 0.15 m head image size, and a 0.6 m viewing distance from the screen. The lines correspond with a *C* = 3/2, 1, and 2/3 which should induce perceived expansion, a geometrically isomorphic scene, and compression respectively. Using these *C* values as well as Eq.1, these lines demonstrate the deviation we might expect if our model is correct, leading to an angle of 45˚ being perceived when the angle presented is ~56˚, 45˚, and 34˚ respectively. This model would operate the same way regardless of image size if correct. The blue line represents our target of the participants' perception, being that of a 45˚ orientation.

thus resulting in the perception of image depth compression or expansion. Using this model, we can make the following hypotheses:

H2.1: Depth compression and expansion in pictorial spaces are resultant of disparities between the vantage point of the observer and the CoP of the image. This disparity results in perceived depth compression (or expansion) by a factor *C*. Therefore, we should see predictable changes in the perceived orientation based on *C*, explaining the orientation bias phenomenon. Specifically, angular orientation should appear more obtuse for $0 < C < 1$ and more acute for $C > 1$.

H2.2: Changing the size of the presented image, and thus the retinal image size, without changing the compression factor *C* should have no effect on the magnitude of the orientation bias.

Specifically, H2.1 predicts that if the CoP is further from the image plane than the observer, this would theoretically result in the perceived depth being compressed. This could cause angles to appear more obtuse and thus be overestimated. If it is more often the case that the CoP is further than the observer, this would explain depth compression, rather than expansion, is often observed in the context of picture perception. Additionally, H2.1 predicts that if the CoP is closer to the picture plane than the observer, it should result in perceived depth expansion, which might cause angles to appear more acute leading to a systematic underestimation. Lastly, in the case when the CoP and observer occupy the same position relative to the picture plane, as they did throughout Experiment I, there should be no observable bias from the true presented angle. H2.2 is supported by the information contained in Tab. 3.1, which shows that changing the image size should not affect the perceived orientation, only the *C* factor should.

The current study was conducted online remotely using Pavlovia (https://pavlovia.org/), an online platform designed for posting and conducting behavioural and psychophysical studies. Participants were instructed to seat themselves at arms-length (approximately 60 cm) in front of their computer monitors. They were then prompted to resize an image of a credit card to its correct dimensions; this was done to create a scale to be used such that the presented sizes of the experimental stimuli would be consistent regardless of the size of the participant’s monitor. Participants were instructed to use the “left” and “right” arrow keys on the computer’s keyboard to indicate whether they perceived the displayed face as either looking left or right of 45˚ left of the participant. This data was used to generate psychometric curves whose 50% response thresholds (point of subject equivalence, or PSEs) correspond with the angle the participant perceived to be 45˚ and whose just-noticeable differences (JNDs) indicate sensitivity. These PSE and JND values were used to evaluate the effects of the independent variables on their perceptions of the stimulus' orientation.

Methods

Participants

20 participants (16 male: 4 female) were recruited randomly for this study through social networks and all participated voluntarily. The participants' mean age was 27.05 ± 8.34 SD, with median age 25. 16 participants were right-handed while 4 were left-handed. All participants had normal or corrected-to-normal vision. This study was approved by the Office of Research Ethics at York University and was carried out in accordance with its regulations.

Stimuli

The face stimuli presented to observers during the experiment were created using MATLAB and consisted of twelve different faces (6 male, 6 female). To generate each of the twelve face stimuli, 5 different heads were randomly selected from the MPI face database, which consists of over 200 head models (Troje & Vetter, 1996). These heads were then assigned random weights and combined to generate each of the final face stimuli used in the experiment. These face stimuli were created such that the models had realistic human head dimensions within the virtual space they were rendered in. The images of each of the face stimuli were saved using custom MATLAB code. The manipulations which differentiated the images were: the expected *C* factor given the observers' vantage point, the display size of the generated image, and the viewpoint around its vertical axis from which the face was rendered. The expected *C* factor was controlled indirectly by manipulating the distance of the virtual camera's CoP relative to the stimulus object (D_{3D}). By displaying the captured images at preset sizes on the monitor (the picture plane), changing D_{3D} also directly controls D_{CoP}, the distance between the CoP and the picture plane. Since the distance between the observer and the picture plane (D_{Obs}) was relatively constant, manipulating D_{CoP} affects the ratio between D_{CoP} and D_{Obs}, which can be denoted as the expected *C* factor (Eq.2). D_{3D} was manipulated such that the *C* factors of the generated images were one of three values consistently, regardless of the displayed image size; this was the main manipulation applied. Figure 3.3 demonstrates the visual effect of altering the *C* factor through such manipulations, however, larger values than those tested in the actual experiment are shown to emphasize this effect. Assuming that the observer is a constant distance away from the screen, the expected effect of such a manipulation is perceived compression or expansion of the represented pictorial space.

Figure 3.1: These images are examples of the same 3-D head being distorted by being rendered at varying distances, thus altering the *C* factor, and theoretically inducing differences in perceived depth. All three heads are depicted from a 20˚ viewing angle. The effects are subtle, but it can be noticed that the expanded face (left) is a little narrower and elongated while the compressed face (right) is a bit wider and flatter looking. Note: the above images serve as an example and were not used in the experiment. The *C* values used in these images are twice as powerful to emphasize the visual effect.

Design

In total, there were three *C* factors (*C* = 3/2, *C* = 1, and *C* = 2/3), two displayed image sizes (7.5 cm tall and 15 cm tall), twelve different faces (6M and 6F), and 11 different orientations (20˚-70˚ left of a full-frontal view, in steps of 5˚). This created 792 unique images which were displayed in a random order to mitigate order effects. The current study was developed using PsychoPy (v2021.1.3), a software designed for developing psychological experiments. The integration of PsychoPy with Pavlovia – a platform designed for behavioral science researchers to run, share, and explore experiments online – allowed the study to be run remotely. Custom MATLAB code was used to analyze the participants' data upon completion of the experiment. The study used a within-subjects design.

Procedure

Prior to beginning the experiment trials, participants read and signed a virtual informed consent form, detailing the study and ensuring that it was being carried out in accordance with York University's regulations. Upon signing, they were prompted to enter demographic information, such as their age, gender, and handedness. After this, they were instructed to place a credit card on the screen and rescale an image of a credit card to fit it; this step was incorporated to ensure that the stimuli were displayed at consistent sizes across participants despite being displayed on monitors with variable sizes. Participants were also instructed to position themselves at arm's length from the monitor while performing the task, which we approximated to be 60 centimeters on average for our participant demographic. This is important as the distance away from the monitor should affect the compression of the image. After all these steps were complete, participants were given instructions for the task, which was to use the "left" and "right" arrow keys to indicate whether they perceived the displayed face as either looking left or right of 45˚ left of the participant.

After reading the instructions participants were able to begin the trials. In each trial, one of the face stimuli was displayed on the participant's computer monitor in front of a plain grey background (Fig. 3.3 for reference). Once the experiment had started, they were given 50 unrecorded, randomly-selected warm-up trials – though they were not made aware of this – to learn the task, which also ensured that the response times of the recorded trials would be reliable. Following this, they were presented with the 792 actual trials, which were used for analysis. After each response, a 0.3 second gap was inserted during which a blank screen was shown before the next face appeared. This was done to prevent the faces from blending or creating illusory motion. The faces were displayed until the participant responded as to whether they thought the stimulus was facing left or right of the target orientation (45˚). Upon completion of the experiment, participants were given a debrief as well as experimenter contact information and were thanked for their participation. The experiment took an average of 20-30 minutes to complete.

Data Analysis

Using custom MATLAB code, a psychometric function was fitted for each of the three *C* factor levels ($C = 3/2$, $C = 1$, and $C = 2/3$) and for both image sizes (7.5cm-tall and 15cm-tall). These fits were done for each individual participant and were used to calculate their PSE – indicative of the angle perceived as 45˚ – as well as their JND, the size of which was indicative of their sensitivity.

Using R statistical software and the participants PSEs and JNDs, the results were analyzed by a repeated measures analysis of variance (ANOVA). This analysis compared the scores of participants' trials in each of the conditions to assess our hypotheses and verified whether there were any statistically significant interactions between our independent variables. Two discrete, within-subjects, independent variables were examined: the *C* factor ($C = 3/2$, $C = 1$, and $C = 2/3$) and the presented head image size (7.5cm-tall and 15cm-tall). There were also two continuous dependent variables, being the participants' PSEs as well as their JND scores. Additionally, a one-sample, two-tailed t-test was performed to determine whether the pictorial orientation bias phenomenon existed and was significant.

Results

The result of the t-test indicated that a statistically significant orientation bias inward exists when viewing images in pictorial spaces ($p < .0002$***). On average, participants ($n = 20$) perceived an orientation of 45˚ at a presented angle of $M = 37.4˚$ ($SD = 7.33˚$, $SEM = 0.37˚$). This result is consistent in magnitude with the previous findings of Troje et al. (2018) and supports the strength of the orientation bias phenomenon, though it does not explain it.

The ANOVA on the PSEs indicated that the explanatory variables explored in the current study – image size and the vantage-point-induced compression factor *C* – do impact angular estimation and are statistically significant. However, there does not seem to be any interaction between these two variables (Tab. 3.2). The ANOVA on the sensitivities, as measured by the just-noticeable difference (JND), yielded no statistically significant relationships with the explored variables, as all $p > .05$ (Tab. 3.3). The effects of vantage-point-induced compression and head-size on participant angular estimations become apparent when viewing the resulting psychometric functions (Fig. 3.4) and have also been extracted to independent bar graphs (Fig. 3.5; Fig. 3.6) for clarity. In the tested range of sizes and compressions, it appears that the effect of *C* on perceived bias is smaller than our model predicted and in the opposite direction. Furthermore, our model failed to predict the differences between the displayed head-sizes, which had a larger and more statistically significant effect on orientation bias than *C*.

Experiment II: PSE Repeated Measures ANOVA

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
IS	1	19	27.592	4.53E-05	***	0.019
C	2	38	4.954	1.20E-02	*	0.003
IS:C	2	38	0.452	6.40E-01		0.000285

Table 3.2*:* Analysis of variance demonstrating whether the explored factors – head image size (IS) and expected compression (*C*) – affect the magnitude of the observers' orientation bias, as measured by their PSEs, when viewing a pictorial head. Single, double, and triple asterisks corelate to *p<.05, p<.01,* and *p<.001* respectively. The ges represents the generalized eta squared.

Experiment II: JND Repeated Measures ANOVA

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
IS	1	19	4.183	5.50E-02		0.009
C	2	38	0.607	5.50E-01		0.002
IS:C	2	38	1.886	1.66E-01		0.008

***Table 3.2**:* Analysis of variance demonstrating whether the explored factors – head image size (IS) and expected compression (*C*) – affect the sensitivity of the observers to changes in presented orientations when viewing a pictorial head. Sensitivity was measured by the just-noticeable difference (JND) of the psychometric functions, measured by averaging the x-distance (in degrees) between the 50% PSE and the 25% and 75% response thresholds. Single, double, and triple asterisks corelate to *p<.05, p<.01,* and *p<.001* respectively. The ges represents the generalized eta squared.

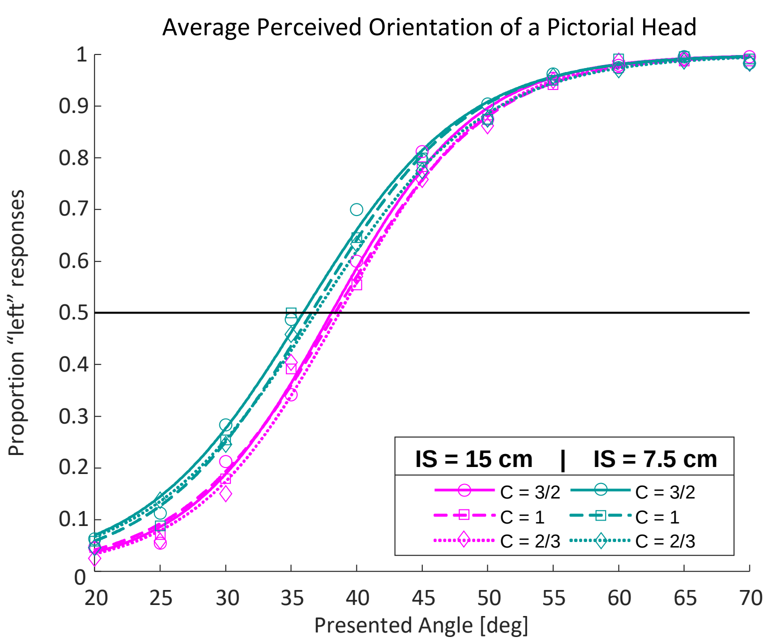

***Figure 3.2**:* Psychometric functions averaged across all n = 20 participants for each condition (image-size = 15 cm or 7.5 cm tall, *C* = 3/2, 1, or 2/3). The black line represents the 50% response threshold, the crossing of which is indicative of the angle which was perceived as 45˚ in each condition.

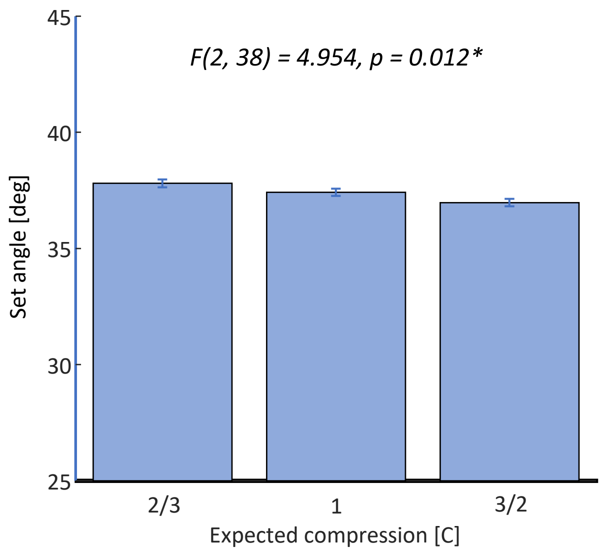

Figure 3.5*:* The average angle which was perceived as 45˚ across all n = 20 participants at each *C* value. Means were derived from 50% response thresholds of the fitted psychometric curves. Means for each *C* value were 37.80˚, 37.42˚, and 36.98˚ for *C* = 2/3, 1, and 3/2 respectively. Error bars represent the standard error of the mean (SEM).

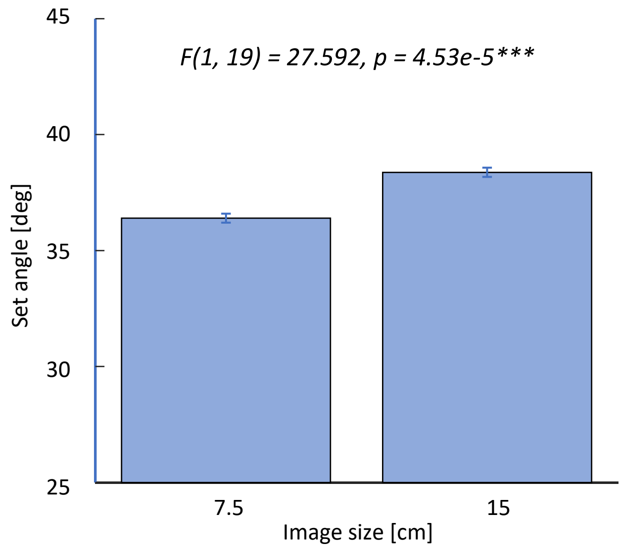

Figure 3.6*:* The average angle which was perceived as 45˚ across all n = 20 participants at each presented image size (cm). Means were derived from 50% response thresholds of the fitted psychometric curves. The mean for the 7.5 cm image size was 36.40˚, while the mean for the 15 cm image size was 38.37˚. Error bars represent the standard error of the mean (SEM).

Discussion

The results of the two-tailed t-test strongly support the existence, significance, and strength of the orientation bias phenomenon being explored. However, the predictions of the vantage-point-induced depth compression model used in this experiment fail to account for this phenomenon. The model predicted that the *C* factors which resulted from vantage point displacements would elicit a large effect on participant 45˚ estimations, approximately 10˚ in magnitude in both directions relative to the 45˚ target, with compression being expected at *C* = 2/3 and expansion being expected at *C* = 3/2. The results however indicate quite clearly that despite there being a significant effect of the *C* factor, the effect of *C* is extremely small and only covers a range of less than 1˚. Furthermore, the PSEs demonstrate that this effect is in the opposite direction of the prediction (Fig. 3.5). The model utilized in the current study predicts that expansion – theoretically caused when the vantage point is farther than the CoP (*C* > 1) – would lead to a greater predicted angle, and compression – theoretically caused when the vantage point is closer than the CoP (0 < *C* < 1) – would lead to a lesser one. However, the results fail to demonstrate this (Fig. 3.2; Fig. 3.5). Regardless of *C*, depth expansion was not observed in any of our results. These results do not support H2.1, but rather demonstrate that it is unlikely that vantage-point-induced depth compression is a primary causal factor for the experienced orientation bias phenomenon, despite its effect being statistically significant.

The statistically significant effects of presented image size were similarly unpredicted and unaccounted for by our model. Furthermore, image size had a stronger influence on the magnitude of the orientation bias than the *C* factor. However, similarly to the effect of *C*, the effect of image size was small, only causing distortions of angular perception of approximately 2˚ in magnitude (Fig. 3.6). These results make it clear that changing the retinal image size reliably biases angular estimates, which contradicts H2.2. Since the size of the deviations caused by the independent variables explored in the present study were much smaller than the magnitude of the bias confirmed by the 2-tailed t-test, it's unlikely that either are the primary cause of the orientation bias phenomenon.

Experiment III: Stimulus X/Z Aspect Ratio and Presence of Animation as Predictors for Orientation Bias in Pictures

Motivation and Hypotheses

Despite the results of Experiment II being weak in magnitude and contradicting our hypotheses, the overall observed orientation bias was again observed, but twice as powerful as the bias from Experiment I. The bias being present and significant in both experiments demonstrates its prevalence, but the difference in the strength of the phenomenon is unexplained. One important difference between the two experiments was that we used the method of adjustment in Experiment I and a method of constant stimuli in Experiment II. Participants actively rotated heads in real-time during the first experiment but were only shown static images of faces in the second and were forced to make a choice. Perhaps the rotation that participants observed while actively setting a head into the 45˚ position in Experiment I provided them with information about the heads' shapes, and the static images used in Experiment II were unable to provide the same information. Furthermore, Experiment I used full computer-generated heads, while Experiment II used facial masks. If shape information is a key factor to making angular estimations, as the Experiment I and Troje et al. (2018) results suggest, then removing such information could affect the accuracy of these estimations. In this case, perhaps geometric information was lost through having incomplete stimulus objects, such as the face masks used as Experiment II's stimuli, or the removal of motion which could help indicate an object's true dimensions. Perhaps when lacking sufficient geometric information, observers and their visual systems incorrectly approximate the head shape as cylindrical. This incorrect shape prior could affect visual perception of spatial layout and angular orientation as demonstrated by Crunelle's (1996) trapezoidal rooms. If the structure-from-motion provided by the manual adjustment of stimuli in Experiment I increased the quality of the object shape information available, it could have mitigated the bias created by the incorrect shape prior, explaining why the bias was larger in Experiment II. Based on the differential results of Experiments I and II and the above literature concerning the potential effects of shape on orientation perception, we posit that:

H3.1: The orientation bias phenomenon stems from the visual system adopting a cylindrical shape prior for human heads when additional shape information is absent, which does not match their typical elliptical horizontal cross section.

H3.2: Providing additional shape information should mitigate the strength of the bias by informing the geometric prior to be more accurate.

H3.1 is evidenced by the shape-dependent results from Experiment I which demonstrated a trend of angular orientation estimates becoming increasingly accurate as the heads' horizontal cross section became more circular. H3.2 would explain the difference in the magnitude of the recorded orientation bias in Experiments I and II, as denoted by the t-tests performed in each.

To test these hypotheses, we attempted to replicate the orientation bias with a non-facial object which was also cylindrical, but whose horizontal cross section could be slightly warped to match natural human X/Z aspect ratios without arousing too much suspicion from participants. The selected proxy objects were coffee mugs, as they have a cylindrical shape prior, have a handle which defines orientation, and can adopt different shapes while still being accepted as normal mugs. Our second hypothesis, H3.2, is not specific to human faces. Rather, if correct it should apply to other bilaterally symmetric objects as well. Therefore, we can use pictorially rendered mugs as proxies for pictorial faces despite primarily being interested in the facial orientation bias. By systematically altering the mugs' shape between circular and elliptical horizontal cross sections, we can assess whether an inaccurate shape prior affects the accuracy of angular judgements. Additionally, by manipulating variables which increase or decrease the amount of shape information available to the participants, we can determine whether increasing the amount of shape information available to the participant informs and improves their ability to make accurate angular estimations. The two independent variables which provide additional shape information being explored are the presence or absence of animation and the elevation of the viewpoint. Animation was selected as it explicitly differentiates Experiments I and II, and elevation is supported by the viewpoint-from-above assumption as a higher vantage point provides additional depth information of the viewed object.

Based on the results of Experiments I and II as well as the hypotheses stated above, it is expected that there will be main effects of the objects' X/Z aspect ratios, the presentation of the object as static or dynamic, and the elevation of the observers' viewpoint. The effect of the X/Z aspect ratio on the orientation bias was already reported by Troje et al. (2018) and replicated in Experiment I. However, if shape information is critical in making accurate assessments of orientation, then varying the shape is necessary to explore the other factors of interest. Presenting an object dynamically instead of statically, by displaying the object rotating for example, could provide some extra object shape information through structure-from-motion which can inform angular judgements, thus mitigating the bias effect. Similarly, elevated viewpoints provide more accurate shape information, allowing hints towards the true X/Z dimensions to be visible. Naturally, it would be easiest to discern orientation from a birds-eye-view and probably most challenging from a frontal view, as there is comparatively less depth information available. This supports the idea that increasing the elevation of the viewpoint should diminish the magnitude of the bias. Additionally, it is expected that there will be a significant interaction effect between the X/Z aspect ratio and the presence of animation, as this interaction would clearly demonstrate that the extra shape information provided by animation mitigates the bias caused by assuming an incorrect prior.

This study was conducted in the Biomotion Lab at York University using PsychoPy on a desktop computer. Participants placed their heads in a chinrest positioned exactly 60 cm away from the monitor.

They were then instructed to use the “left” and “right” arrow keys on the computer’s keyboard to indicate whether they perceived the displayed mugs’ handles as either pointing left or right of 45˚ left of the participant. This data was used to generate psychometric curves whose 50% response thresholds (PSEs) correspond with the angle the participant perceived to be 45˚ and whose just-noticeable differences (JNDs) indicate sensitivity. These PSE and JND values were used to evaluate the effects of the independent variables on participants’ perceptions of the stimuli’s orientation.

Methods

Participants

There were 23 participants recruited for this study through York University’s URPP portal, of whom 17 met the inclusion criteria (6 male: 9 female: 2 non-binary). The mean age of those 17 participants was 19.9 ± 2.8 SD, median age 19. All participants had normal or corrected-to-normal vision. Upon arrival to the lab, participants read and signed an informed consent form, detailing the study and ensuring that it was being carried out in accordance with York University’s regulations. Participants received 1.5 URPP credits as compensation for their participation in the study.

Stimuli

The experimental stimuli consisted of computer-graphically generated pictures of coffee mugs. The X/Z aspect ratios of these mugs were manipulated such that they could represent a perfect cylinder (1/1), an average adult human head (1/1.3) and its inverse (1.3/1) (Fig. 4.1 A). Mugs were used as proxies for human heads because warping the aspect ratio of facial stimuli may be too obvious and thus alert the participant to the manipulation. A mug is permissible to the viewer when it comes in any shape, even if it does not precisely match their expectation. Furthermore, like noses on faces, mugs have handles which give the participant the necessary information to make an orientation estimate.

The stimuli consisted of both static pictures and short videos of these mugs, which similarly to Experiment II, were captured at various orientations (20˚-70˚ left of a full-frontal view, in steps of 5˚). In the videos, the mug would pivot 20˚ to either side in a sinusoidal fashion (Fig. 4.1 C) before resettling on its initial starting position. The short animation took 2 seconds, and so to match the exposure, the pictures were displayed for the same amount of time; only after those 2 seconds could the participants respond. All mugs were given an identical polka-dot texture which helped elucidate the mugs' shape as well as their motion. Both the pictures and videos were rendered using a virtual camera, which captured orthographic projections of the stimuli at three different elevated viewing points (Fig. 4.1 B). All objects and images used in this study were created using Unity3D.

Design

The experiment was developed in and run using PsychoPy (v2021.1.3). The independent variables were the stimulus-type (static or dynamic images), the stimulus' X/Z aspect ratio (0.77, 1.00, 1.30), and the viewing angle (0˚, 15˚, 30˚ camera elevation). There were also 3 unique mug shapes – each with the same height : width : depth ratio – as well as 2 average mug diameters (10 cm and 7.5 cm) to increase the variability between trials. Each permutation was captured in 11 different orientations (20˚-70˚ left of a full-frontal view, in steps of 5˚). Each trial consisted of a single stimulus being shown, and there were a grand total of 1188 trials. The current study was blocked based on stimulus-type and was counterbalanced, such that half of the participants saw pictures first and the other half saw the videos first. The trial order within these blocks was randomized, thus mitigating any potential order effects. The study used a within-subjects design.

Procedure

Participants were instructed to use a chinrest positioned 60 cm away from the computer screen. They were then instructed to use the arrow-keys on their keyboard to indicate whether they believed

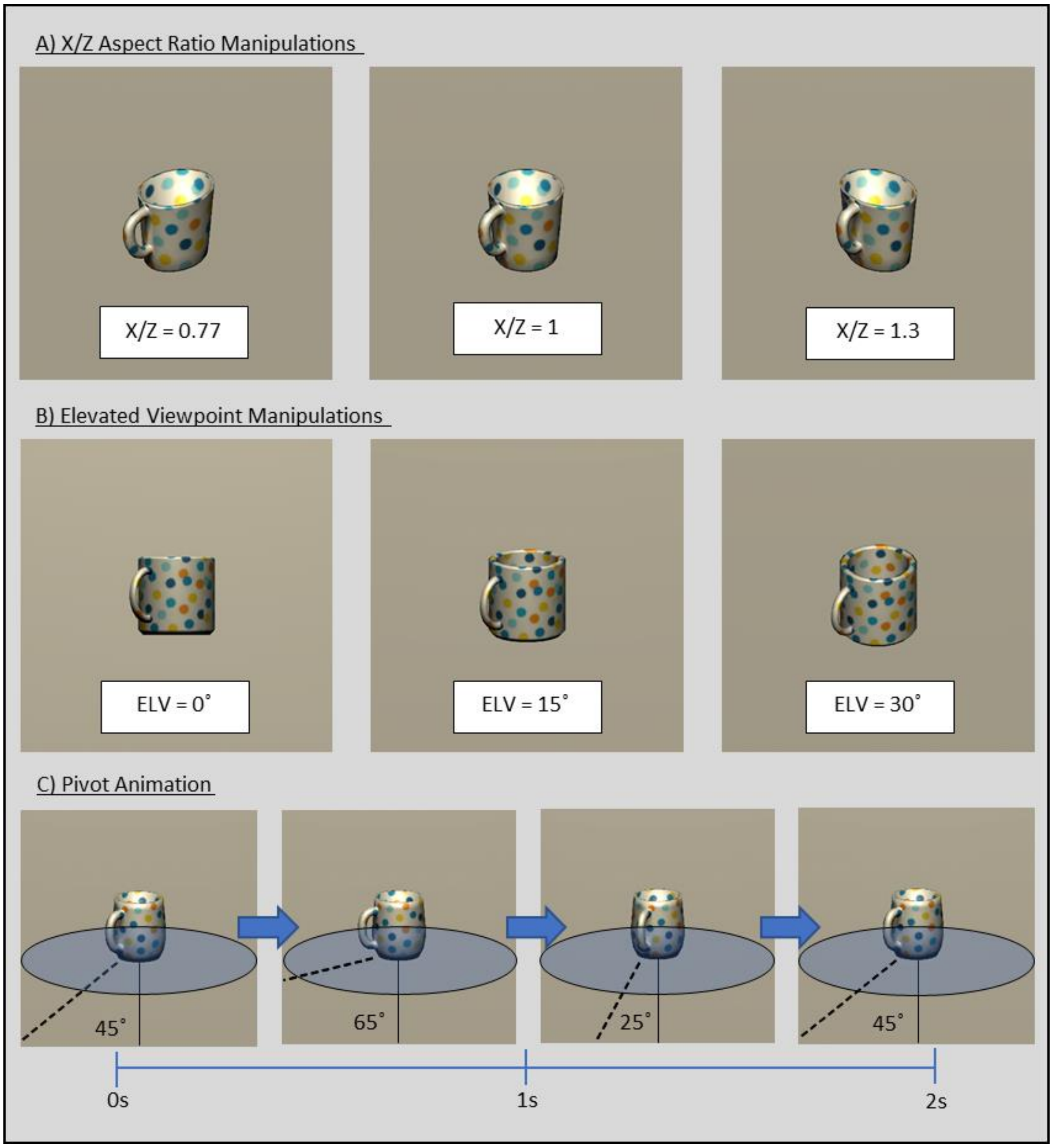

***Figure 4.1**:* **A)** These three images depict the same mug whose X (width) and Z (depth) dimensions have been warped such that the aspect ratios (left-to-right) match a natural human head-shape, a perfect cylinder, and the inverse of the head aspect ratio. All three are oriented to the 45˚ position and presented from a 30˚ elevation to emphasize the differences in shape. **B)** These three images demonstrate the different perspectives the stimuli were presented from, by varying the camera elevation. All three depict the same mug oriented to the 45˚ position. **C)** These images are meant to illustrate the sinusoidal pivot undergone by the stimuli in the videos. This animation takes 2 seconds, begins with the mug directed at its origin, then rotates 20˚ left and right of that origin, and finally re-settles at its origin.

the stimulus on the screen was rotated to the left or right of a specific target angle, being 45˚. In each trial, the mug stimulus was presented in front of a tan background. The stimulus was displayed for 2-

seconds before participants were able to respond and remained visible until a response was entered. After this, a 0.3 second gap was inserted during which a blank screen was shown before the next mug was shown. This was done to prevent the appearance of illusory motion or blending of the mugs. Participants were given 25 unrecorded, randomly selected warm-up trials in each block, such that the response times of the recorded trials would be more reliable. Participants were unaware that the warm-up trials were not recorded. There was a short break between the first and second block, allowing the participants to rest. After completing the experiment, participants were compensated with their credits. The study took on average 90 minutes to complete.

Data Analysis

Using custom MATLAB code, the angular response data provided by the participants was used to generate fitted psychometric curves, whose 50% response thresholds (PSEs) correspond with the angle which was perceived as 45˚. The just-noticeable differences (JNDs) of the generated curves were used as a measure of sensitivity. These fits were generated for each participant. R statistical software was then used to run a repeated measures ANOVA on both the PSE and JND scores of the participants to assess our hypotheses. Three discrete, within-subjects, independent variables were examined: stimulus-type (static, dynamic), elevated viewpoint (0˚, 15˚, 30˚), and X/Z aspect ratio (0.77, 1.00, 1.30). The two continuous dependent variables being the participants' PSE and JND scores. Additionally, a one-sample, two-tailed t-test was performed to determine whether there was a base pictorial orientation bias.

Results

The repeated measures ANOVA of the participants' PSE scores supports our hypotheses. The X/Z aspect ratio, stimulus-type, and viewpoint elevation each significantly influenced the perceived angle. Furthermore, there was a significant interaction between the stimulus-type and the X/Z aspect ratio of the stimuli (Tab. 4.1). These results indicate that these variables directly impact the reported orientation bias phenomenon. Additionally, the repeated measures ANOVA of the participants' JND scores revealed statistically significant effects of viewpoint elevation and the mugs' X/Z aspect ratios. It also revealed a significant interaction effect between these two variables (Tab. 4.2). These results indicate that these variables affect the sensitivity with which observers can reliably discern orientation changes as well as their confidence in making object orientation estimates. The result of the t-test indicated that a significant orientation bias inward exists when viewing images in pictorial spaces ($p < 6.3E\text{-}06$***). On average, participants (n = 17) perceived an orientation of 45˚ at a presented angle of M = 35.8˚ (SD = 5.75˚, SEM = 1.39˚). This result indicates that despite finding multiple significant results, there still appears to be a base offset for orientation bias which is unexplained by the factors explored in the current study.

Experiment III: PSE Repeated Measures ANOVA

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
ST	1	16	6.041	2.60E-02	*	0.037
ELV	1.3	20.84	6.337	1.40E-02	*	0.029
XZAR	2	32	55.698	3.78E-11	***	0.161
ST:ELV	1.35	21.54	1.819	1.92E-01		0.003
ST:XZAR	2	32	9.543	5.62E-04	***	0.011
ELV:XZAR	4	64	2.201	7.90E-02		0.006
ST:ELV:XZAR	4	64	0.509	7.29E-01		0.001

Table 4.1: Analysis of variance demonstrating whether the explored factors affect the magnitude of the observers' orientation bias, as measured by their PSEs, when viewing a mug on a computer screen. The factors examined were the presence or absence of animation – represented as the stimulus-type (ST), the elevation of the viewpoint (ELV), and the mugs' X/Z aspect ratio (XZAR). Single, double, and triple asterisks corelate to p<.05, p<.01, and p<.001 respectively. The ges represents the generalized eta squared.

Experiment III: JND Repeated Measures ANOVA

Source of Variation	*DFn*	*DFd*	*F*	*p*	*p<.05*	*ges*
ST	1	16	3.782	7.00E-02		4.10E-02
ELV	2	32	13.191	6.64E-05	***	2.10E-02
XZAR	2	32	6.166	5.00E-03	**	1.30E-02
ST:ELV	2	32	0.373	6.91E-01		7.31E-04
ST:XZAR	1.3	21.34	0.027	9.25E-01		7.13E-05
ELV:XZAR	4	64	13.941	3.11E-08	***	4.40E-02
ST:ELV:XZAR	4	64	1.15	3.41E-01		4.00E-03

Table 4.2: Analysis of variance demonstrating whether the explored factors affect the sensitivity of the observers to orientation changes, as measured by their JNDs, when viewing a mug on a computer screen. The factors examined were the presence or absence of animation – represented as the stimulus-type (ST), the elevation of the viewpoint (ELV), and the mugs' X/Z aspect ratio (XZAR). Single, double, and triple asterisks corelate to p<.05, p<.01, and p<.001 respectively. The ges represents the generalized eta squared.

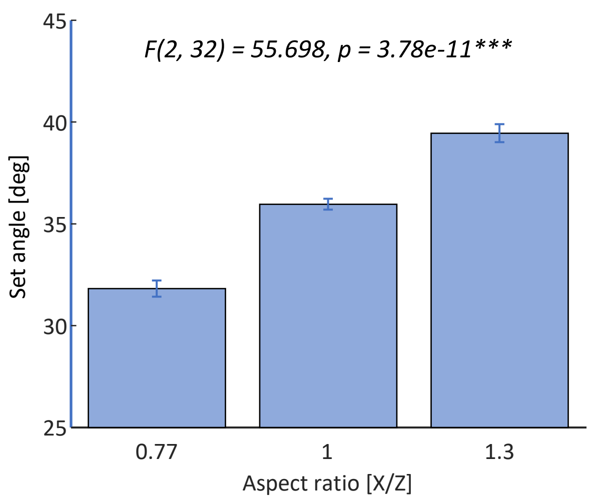

Figure 4.2: The relationship between the X/Z aspect ratio of mugs and their presented angle which was perceived as 45˚. Error bars represent the standard error of the mean (SEM).

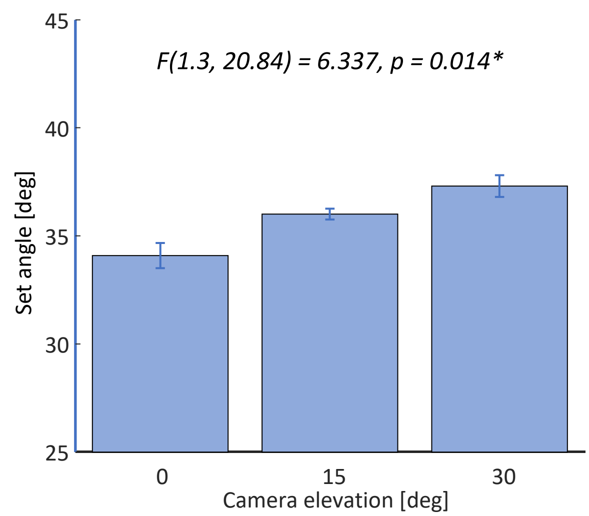

Figure 4.3: The relationship between the elevation of the camera and the presented angle which was perceived as 45˚. Error bars represent the standard error of the mean (SEM).

The significant effects described in Tab. 4.1 and Tab. 4.2 have been plotted to clearly illustrate their overall impact on perceived orientation. Altering the stimulus' X/Z aspect ratio from its expected cylindrical shape reliably shifted its perceived orientation, with narrower shapes biasing perception toward the normal of the pictorial plane and wider shapes biasing perception away from that normal (Fig. 4.2). This result appears to be linear in nature and was expected based on the results from Troje et al. (2018) as well as the replicated shape effect observed in Experiment I.

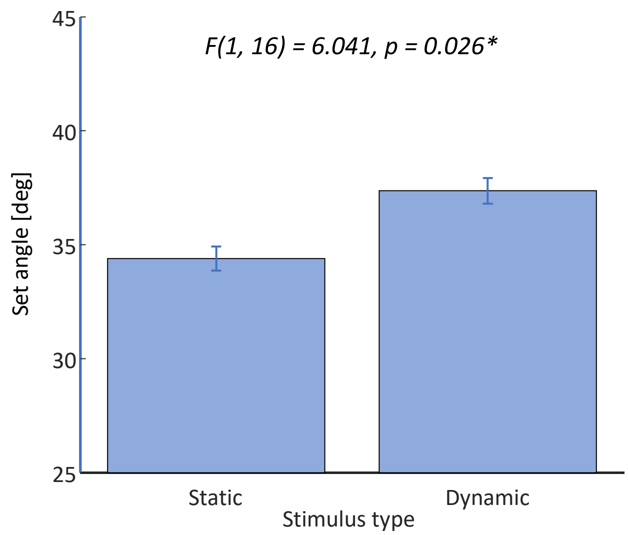

***Figure 4.4**:* The relationship between the stimulus type, being presented as static images or dynamic videos, and the presented angle which was perceived as 45˚. Error bars represent the standard error of the mean (SEM).

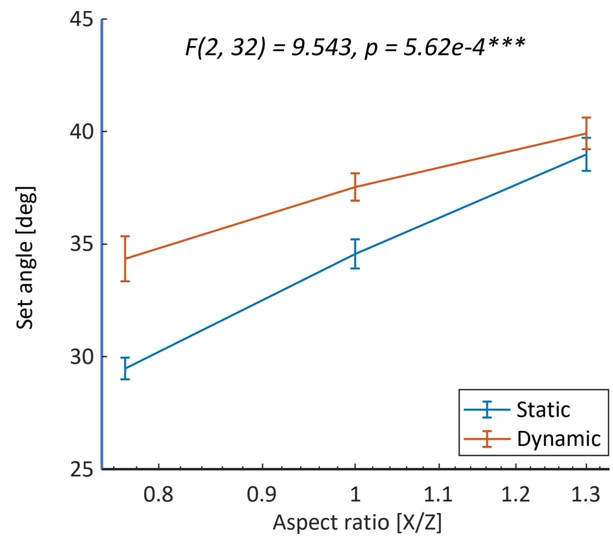

***Figure 4.3**:* Semilogarithmic depiction of the interaction between the stimulus type, being presented as static images or dynamic videos, and the X/Z aspect ratio of the mugs. The slopes of each line demonstrate the degree to which aspect ratio manipulations affect 45˚ orientation perception in each presentation type. Error bars represent the standard error of the mean (SEM).

Additionally, manipulating the type of stimulus viewed – being static or dynamic – and the elevation of the observers' viewpoint also had a significant effect on orientation bias. Increasing the elevation of the viewpoint increased the accuracy of participant estimations, as evident through 45˚ orientation estimates approaching the veridical 45˚ orientation with increases in elevation (Fig. 4.3). Similarly, adding motion to the stimuli increased the accuracy of the perceived orientation estimations (Fig. 4.4). Lastly, the interaction between the stimulus type and the X/Z aspect ratio demonstrates that the effect of shape alterations on orientation perception is greatly exaggerated in static representations and mitigated in dynamic representations (Fig. 4.5).

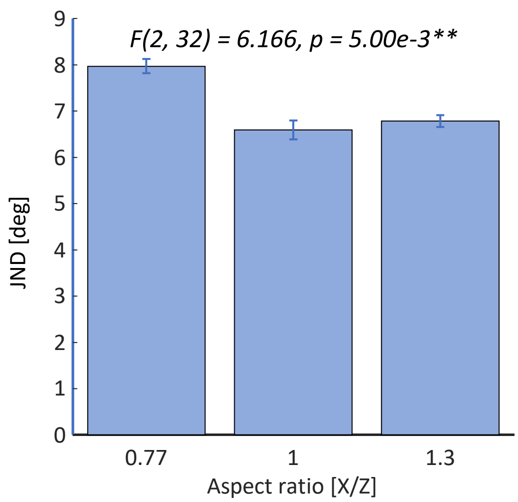

Figure 4.6*:* The relationship between the X/Z aspect ratio of the mugs and the just-noticeable difference (evaluated using the 25% and 75% response thresholds) in degrees required to reliably shift participant judgements. Error bars represent the standard error of the mean (SEM).

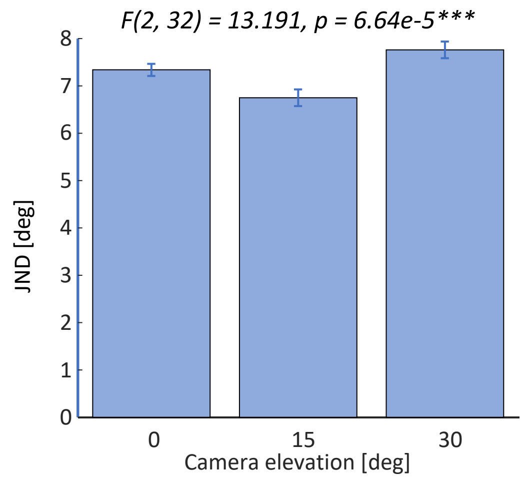

Figure 4.7*:* The relationship between the elevated viewpoint of the mugs and the just-noticeable difference (evaluated using the 25% and 75% response thresholds) in degrees required to reliably shift participant judgements. Error bars represent the standard error of the mean (SEM).

The significant sensitivity results indicated that changing the X/Z aspect ratio of the stimuli away from the expected shape decreased the participants' sensitivity to orientation changes. This effect was particularly strong when decreasing the aspect ratio, thus narrowing the stimuli (Fig. 4.6). Changes in the visual perspective also affected participant sensitivity. Namely, both the 0˚ and 30˚ perspectives, corresponding to a frontal view and a significantly elevated view respectively, caused decreases in discrimination sensitivity. This effect was strongest in the 30˚ elevated condition (Fig. 4.7). Lastly, the interaction between aspect ratio and elevation demonstrates that the most powerful effects within each condition – decreased sensitivity in the X/Z = 0.77 and ELV = 30˚ conditions – seem to compound, causing a much larger decrease in sensitivity (Fig. 4.8).

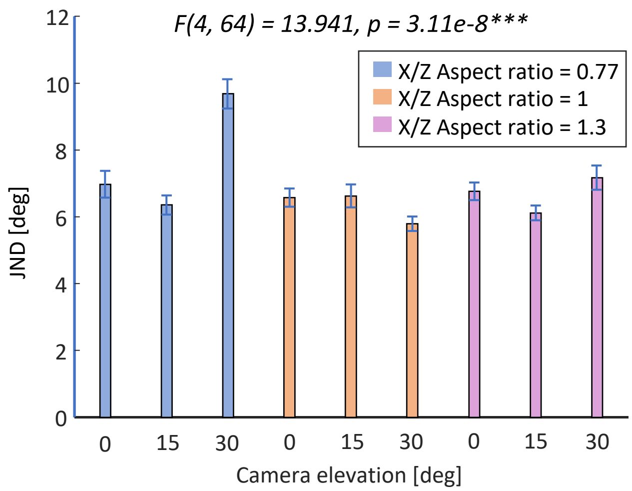

Figure 4.8: The interaction between the elevated viewpoint of the mugs and their aspect ratios and how they affect participant discrimination sensitivity, as measured by the JND. Error bars represent the standard error of the mean (SEM).

Discussion

Altering the cylindrical mugs' dimensions such that they were elliptical caused participants to make predictable errors when estimating orientation. Changing the mugs' aspect ratios such that X > Z or X < Z biased orientation estimates in accordance with the relationship recorded by Troje et al. (2018) and replicated in Experiment I. The orientations of mugs with X/Z aspect ratios like that of a human head (X/Z = 0.77) were systematically overestimated with the same strength as the typical orientation bias observed with pictorially represented faces. The size of this effect is indicative that the expected shape is likely a key contributor to the bias phenomenon being explored, thus strongly supporting H3.1. However, because the bias was still present in the conditions which were meant to be neutral (X/Z = 1), other factors necessarily contribute to this phenomenon as well. In summary, the visual system likely adopts the heuristic of a cylindrical head shape unless explicit information about its shape is provided. However, despite the strength of this effect, this cannot be the sole determinant of the pictorial orientation bias phenomenon.

Manipulating the X/Z aspect ratios of the mugs enabled us to examine whether the other variables explored – the angular elevation of the viewpoint and whether the stimuli were presented dynamically or statically – affected observers' abilities to detect those changes and use that geometric information to adjust their orientation estimates. These explanatory variables produced significant effects which were consistent with the predictions generated by H3.2. For instance, the perception of the orientations of mugs which were presented dynamically were significantly more accurate than those of statically presented mugs (Fig. 4.4). The interaction between the presented stimulus type and the mugs' X/Z aspect ratios strongly suggests that structure-from-motion is the mechanism by which the orientation bias effect induced by objects' abnormal dimensions is mitigated (Fig. 4.5). This figure clearly indicates that not only does adding motion improve the accuracy of orientation estimates, but specifically that aspect ratio alterations affect participants' angular judgements far more in the static condition – as indicated by the steeper slope.

Additionally, increasing the elevation from which the mugs were viewed reduced the orientation bias and thus increased the accuracy of the reported orientation (Fig. 4.3). It is possible that this is due to the elevated vantage point providing additional information towards the true width and depth of the viewed mugs, which could be referred to as structure-from-elevation. We expected there might be an interaction between the mugs' X/Z aspect ratios and viewpoint elevation as well, as such an interaction might indicate that aspect ratio alterations affect participants' angular judgements more when viewed head on and less as the true X/Z dimensions are gradually revealed by the elevation increase. However, this cannot be said definitively as the ANOVA revealed that the interaction between camera elevation and mug shape was statistically insignificant (p = .079; Tab. 4.1). One reason the results may support structure-from-motion more than structure-from-elevation is due to the camera elevation being tested in a smaller angular range than the motion, thus providing less structural information. The total range of the pivot movement in the dynamic condition was 40˚ around the mugs' y-axes, whereas the maximum camera elevation was only 30˚ above the mugs' x-axes. Perhaps if a larger range of camera elevations were tested, this interaction may approach statistical significance as well. Regardless, these results strongly support that providing additional object shape information to the observers, by means of viewpoint elevation or dynamic presentation, informs the geometric prior to be more accurate and thus mitigates shape-induced orientation bias.

Though there were no explicit hypotheses made regarding the sensitivities, analyzing them yields some findings of interest. For example, Tab. 4.2 shows that deforming the X/Z aspect ratio of an object from its familiar shape negatively impacts discrimination sensitivity. Specifically, participants were most sensitive to orientation when viewing cylindrical mugs, which makes sense as it is in line with the existing cylindrical geometric prior. Interestingly, the effect of using the aspect ratio most typical of an average human head (X/Z = 0.77) was much more pronounced than its flatter inverse, so more narrow objects may be harder to discern orientation from than wider ones.

Adjusting the camera elevation also affected discrimination sensitivity. It is possible that the reason this was significant was due to the frontal view condition (0˚) providing fewer depth cues and perhaps being an unfamiliar viewing angle. Furthermore, perhaps more extreme viewing angles, such as the 30˚ elevation condition, are more challenging to discriminate due to the skew between the elevated angle and the rotated angle. Lastly, the significant interaction effect between aspect ratio and elevation on discrimination sensitivity was unexpected but powerful. The strength of this effect likely comes from the way the two most powerful conditions from camera elevation (30˚) and X/Z aspect ratio (X/Z = 0.77) specifically interacted with each other, adding greatly to the strength of the effect.

General Discussion

Each of the experiments conducted in this thesis explored unique hypotheses to explain the orientation bias experienced when viewing pictorial representations of faces. The results of each individual experiment informed the next, and collectively their findings offer a more comprehensive understanding of the perceptual mechanism involved in this experienced bias. Experiment I tested whether the limitations of the pictorial medium were the cause of this bias, such as the twofoldness of pictures and the associated twofoldness-induced depth compression expected in pictorially represented objects. It also tested the effects of depth cues on the strength of the bias, such as stereopsis and motion parallax which are available in visual spaces but not pictorial ones. Lastly, it tested whether the shift in the occluding contour of the presented heads as viewing distance varied was partially responsible for the experienced orientation bias. However, the results of Experiment I showed no significant differences between the orientation biases in visual and pictorial displays. Additionally, the distance effects observed in Experiment I were in the opposite direction to those predicted by the occluding contour shift hypothesis. Ultimately, regardless of the condition the magnitude of the bias reported in this experiment was far smaller than expected.

Upon inspecting these results, we decided to test another potential source of depth compression in pictures, being vantage-point-induced depth compression. This was based on the model proposed by Farber and Rosinski (1978), which suggests that depth compression is a result of the disparity between the observer's vantage point and the picture's center-of-projection (CoP). Thus, Experiment II tested whether the ratio of the vantage point versus the CoP induced depth compression or expansion by the related compression factor *C*. If it were the case that people consistently viewed images from in front of the CoP, this could explain the experience of orientation bias in pictorially represented faces. However, despite this experiment yielding an orientation bias much larger in magnitude than that reported in Experiment I, the changes in *C*, which were reflective of relative changes in the distance between the vantage point and CoP, had a very miniscule effect on the magnitude of the bias.

The differences in the reported orientation bias between Experiments I and II were thought to stem from the different methods of presentation between the two. Specifically, Experiment I used a method of adjustment, allowing stimuli to be rotated to a target orientation, whereas Experiment II used a method of constant stimuli, showing stimuli as static images at specific orientations. This, in combination with the significance of head shape as indicated by Experiment I and Troje et al. (2018), led to the generation of Experiment III's core hypothesis – in the absence of sufficient shape information, the human visual system assumes a cylindrical shape prior for heads, and this incorrect shape prior is the cause for the observed orientation bias. The results of this experiment clearly demonstrated that altering the shapes of mugs away from the expected cylindrical form reliably biased orientation estimates. Furthermore, these biases were larger for static images than they were in short rotating animations. Thus, this confirmed that the additional structure-from-motion provided by the animations informed a more accurate shape to be used for orientation estimation, mitigating the strength of the bias. This explanation also explains why this bias is far less pronounced or experienced in visual space, as structure-from-motion is available through motion parallax.

These experiments collectively identified numerous relevant factors which contribute to the orientation bias phenomenon. Experiment I's results indicated that increasing the simulated viewing distance stimuli were from the observer increased the magnitude of the orientation bias. This effect was minor in the pictorial viewing condition but more apparent in visual conditions which allowed for stereopsis. The changes induced by increasing the viewing distance, or the simulated viewing distance in the pictorial viewing condition, could be explained by the compression of subjective space altering the geometry of the viewed head (Hecht et al., 1999; Koenderink & van Doorn, 2008). Experiment II's results demonstrated that the observer distance from a picture plane could also affect orientation estimates through vantage-point-induced depth compression, though the effect was minor. Perhaps *C* factors when viewing pictured faces are typically much more extreme than those tested, and thus might induce more considerable changes in perceived orientation. Other explanations could be related to the perceived distance in the pictorial condition relating to the size of the retinal image. Perhaps it becomes harder to ascertain reliable geometric information from faces the smaller they are depicted, and this lack of sufficient shape information could lead the human visual system to adopt an incorrect cylindrical shape prior to inform its judgements of angular orientation.

Motion parallax also appears to be a key factor in mitigating the orientation bias phenomenon. Experiment I's results demonstrated the specific main effect of motion parallax reducing the magnitude of the bias phenomenon. In theory, this operates by motion parallax increasing the amount and quality of geometric information available to the observer. In Experiment I motion parallax was self-induced from natural body sway and was likely experienced in a very minor amount. There was also some structure-from-motion resultant from participants adjusting the stimulus faces to the target 45˚ angle, however, this was inconsistent due to the heads' starting orientations being randomized every trial. Though the self-induced motion parallax was not a factor in Experiment III, the results still corroborate the importance of motion parallax by structure-from-motion through rigorous testing of larger angles of rotation, thus significantly diminishing the orientation bias phenomenon.

The factor which had the most pronounced effect on the magnitude of orientation bias was the X/Z aspect ratio of viewed faces – and mugs – as evidenced by Experiments I and III. The results strongly imply that humans' visual systems adopt a cylindrical shape prior in the absence of sufficient geometric information. Due to the relatively large bias specifically induced by altering the shape of the pictorially represented objects, it is likely that this incorrect prior is the main contributor to the orientation bias experienced when viewing depictions of faces. However, the addition of more accurate shape information can greatly mitigate the observed bias phenomenon, which has been demonstrated explicitly through increasing the elevation of the viewpoint and the addition of structure-from-motion. It is probable that this structure-from-motion is responsible for the overall magnitude of orientation bias being so much smaller in Experiment I as opposed to Experiment II. Upon further examination, it seems that the mechanisms by which each of the other significant factors affected the accuracy of orientation estimates was either through providing additional shape information or altering the present shape information. Increasing viewpoint elevation provides extra depth information and allows some top-down perspective of true dimensions (Cornelis et al., 2016). Motion parallax, self-induced through natural sway or simulated through object rotation also provides additional depth information. Vantage-point-induced depth compression operates by optically deforming the shape of the pictorial image. Most simply, increasing the retinal image size makes shape information easier to discern.

Naturally there were a few limitations to the research conducted, the largest of which being the global circumstances regarding the COVID-19 pandemic. Due to restrictions induced by the pandemic, Experiment II was conducted online, and so the results of that experiment are likely less reliable due to the uncontrolled setting. This is particularly important for this experiment as it assumed observers sat a constant and specific distance away from their computer monitors, though this could not be controlled outside of a lab setting beyond giving instructions to the participants. Additionally, due to the pandemic it was impossible to test the factors explored in Experiments II and III in the visual space provided by virtual reality environments, but only the pictorial space of computer screens. Furthermore, though this research highlighted several relevant factors, it still fails to completely explain the orientation bias phenomenon, as is evident from mean angular estimations in each experiment, regardless of condition, falling below the target 45˚ orientation. Perhaps years of accumulated experience viewing pictures – the vast majority of which are compressed in depth – has led to a developed prior when viewing faces or characters in 2-D spaces, thus the global bias might be proportional to the magnitude of the existing prior, which is resultant of the average experienced image depth compression. Analyzing this would be much more challenging, but it could potentially be the root source of the unexplained default orientation bias.

The most robust finding of this research demonstrates that increasing the amount of – or improving the quality of – shape information available to an observer greatly increases the accuracy of their angular estimations. In the context of pictorial orientation bias, adding depth cues through motion parallax, elevation, or animation can greatly reduce the magnitude of this angular overestimation in pictorially represented faces. This is particularly relevant for future psychophysical studies involving eye-contact or attention through screens. These findings can also be applied to consider potential enhancements for teleconferencing, such as incorporating motion parallax to preserve directionality in video calls. By making the vantage point inside of a video call coincide with the user's own vantage point, it is possible to add motion parallax to these displays, and thus, create a 2-D medium capable of mimicking the directionality characteristic of 3-D visual spaces. This could potentially improve communication efficiency, preserve aspects of non-verbal communication such as gesticulation and gaze, facilitate turn-taking behaviors in conversation, and increase the accuracy of the pictorial proxy to visual space. With the proliferation of human interactions with pictorial spaces, it is more important than ever that we understand the structure of these spaces and the factors which differentiate them from visual spaces. Furthering this research will inform technological advancements in video conferencing, increase the accuracy of 2-D educational training and simulations, inform future psychophysical vision research, and provide a framework to create pictorial spaces which act as more viable proxies for the 3-D visual spaces they are often intended to emulate.

References

Busey, T. A., Brady, N. P., & Cutting, J. E. (1990). Compensation is unnecessary for the perception of faces in slanted pictures. *Perception & Psychophysics*, *48*(1), 1–11. https://doi.org/10.3758/bf03205006

Cornelis, E. V. K., van Doorn, A. J., & Wagemans, J. (2016). Reference Frames and 3-D Shape Perception of Pictured Objects: On Verticality and Viewpoint-From-Above. *I-Perception*, *7*(3), 204166951663728. https://doi.org/10.1177/2041669516637286

Cornilleau-Pérès, V., Wexler, M., Droulez, J., Marin, E., Miège, C., & Bourdoncle, B. (2002). Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Research*, *42*(11), 1403–1412. https://doi.org/10.1016/s0042-6989(01)00298-x

Crunelle, M. (1996). A Problem in Perception: Living in Trapezoidal Spaces. *Perception*, *25*(1_suppl), 35. https://doi.org/10.1068/v96p0108

Cutting, J. E. (1988). Affine distortions of pictorial space: Some predictions of Goldstein (1987) that La Gournerie (1859) might have made. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 305–311. https://doi.org/10.1037/0096-1523.14.2.305

Farber, J., & Rosinski, R. R. (1978). Geometric Transformations of Pictured Space. *Perception*, *7*(3), 269–282. https://doi.org/10.1068/p070269

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 256–266. https://doi.org/10.1037/0096-1523.13.2.256

Goldstein, E. B. (1988). Geometry or not geometry? Perceived orientation and spatial layout in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 312–314. https://doi.org/10.1037/0096-1523.14.2.312

Goodale, M. A., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, *14*(2), 203–211. https://doi.org/10.1016/j.conb.2004.03.002

Hecht, H., van Doorn, A., & Koenderink, J. J. (1999). Compression of visual space in natural scenes and in their photographic counterparts. *Perception & Psychophysics*, *61*(7), 1269–1286. https://doi.org/10.3758/bf03206179

Koenderink, J., & van Doorn, A. (2008). The Structure of Visual Spaces. *Journal of Mathematical Imaging and Vision*, *31*(2–3), 171–187. https://doi.org/10.1007/s10851-008-0076-3

Nanay, B. (2011). Perceiving pictures. *Phenomenology and the Cognitive Sciences*, *10*(4), 461–480. https://doi.org/10.1007/s11097-011-9219-x

Nöe, A. (2006). *Action in Perception (Representation and Mind series)* (unknown ed.). A Bradford Book.

O’Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973. https://doi.org/10.1017/s0140525x01000115

Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception & Psychophysics*, *14*(1), 13–18. https://doi.org/10.3758/bf03198608

Powers, M. B., & Emmelkamp, P. M. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *Journal of Anxiety Disorders*, *22*(3), 561–569. https://doi.org/10.1016/j.janxdis.2007.04.006

Reichel, F. D., Todd, J. T., & Yilmaz, E. (1995). Visual discrimination of local surface depth and orientation. *Perception & Psychophysics*, *57*(8), 1233–1240. https://doi.org/10.3758/bf03208379

Rogers, S., Lunsford, M., Strother, L., & Kubovy, M. (2003). The Mona Lisa effect: Perception of gaze direction in real and pictured faces. *Studies in Perception and Action VII*, 19–24.

Satava, R., & Jones, S. (1998). Current and future applications of virtual reality for medicine. *Proceedings of the IEEE*, *86*(3), 484–489. https://doi.org/10.1109/5.662873

Slater, M. (2009). Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3549-3557.

Stiefelhagen, R., & Zhu, J. (2002). Head orientation and gaze direction in meetings. *CHI '02 Extended Abstracts on Human Factors in Computing Systems - CHI '02*. https://doi.org/10.1145/506443.506634

Troje, N., & Vetter, T.(1996). *Representations of human faces* (41). Tübingen, Germany: Max Planck Institute for Biological Cybernetics. http://hdl.handle.net/11858/00-001M-0000-0013-EB0E-2

Troje, N., Rosen, D., Eftekharifar, S. (2018). Facial orientation biases in visual vs. pictorial space. *Vision Science Society meeting, St Pete Beach, FL. [Abstract published in: Journal of Vision 18 (10):514]*

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience*, *8*(10), 1401–1410. https://doi.org/10.1038/nn1553

Wollheim, R. (1998). On Pictorial Representation. *The Journal of Aesthetics and Art Criticism*, *56*(3), 217–226. https://doi.org/10.2307/432361