UNIVERSITY OF CALIFORNIA
Los Angeles

Quantifying the Perceived Length of the Stereokinetic Cylinder

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

by

Yang Zeng Xing

2024

ABSTRACT OF THE DISSERTATION

Quantifying the Perceived Length
of the Stereokinetic Cylinder

by

Yang Zeng Xing
Doctor of Philosophy in Psychology
University of California, Los Angeles, 2024
Professor Zili Liu, Chair

Perception of three-dimensional (3D) structure from the two-dimensional (2D) pattern of image velocities on the retina remains a fundamental issue in vision science. 3D reconstruction is complicated by ambiguous output of local motion estimates derived from direction-selective cells in early visual cortex, as only motion orthogonal to the contour is perceived. Previous research suggests the global integration performed by the human visual system adopts constraints, such as a preference for minimal amount of shape change (Wallach & O'Connell, 1953; Jansson & Johansson, 1973; Ullman, 1979) or slowest and smoothest velocity field (Hildreth, 1984; Yuille & Grzywacz, 1989; Weiss, Simoncelli, & Adelson, 2002), reflecting systematic statistical regularities in the environment to restrict the potential 3D interpretations.

In the current study, we will employ stereokinetic phenomena, which are 2D images that result in the perception of non-veridical 2D and 3D percepts when rotated about an axis perpendicular to the image plane, to probe the constraints underlying the integration scheme and determine how motion information can allow depth. Previous research conducted in

the laboratory has suggested that the visual system applies preferences for minimum motion and minimal deformation (i.e., maximal rigidity) when viewing stereokinetic stimuli (Rokers, Yuille, & Liu, 2006; Xing & Liu, 2018). In order to facilitate the development of computational models that test whether the visual system prefers a 3D object that results in the minimal amount of change and slowest velocity field could be generalized to other stereokinetic stimuli, we have developed various measurement methods to rigorously quantify the perceived depth of the stereokinetic cylinder. Across both experiments measuring observers' perceived depth of the stimulus, the length of the illusory cylinder was constrained by a preference for slow motion and maximal rigidity. To our knowledge, we are the first to quantify the perceived length of the stereokinetic cylinder.

The dissertation of Yang Zeng Xing is approved.

Scott Pratt Johnson

Philip Kellman

Hongjing Lu

Zili Liu, Committee Chair

University of California, Los Angeles

2024

Sorry it took me so long to arrive.

Thank you for staying by my side.

Now, come see what we've done.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

As I reflect on my time at UCLA, I am filled with gratitude for every person who has played a part in helping me get within sight of the checkered flag. But there are a few individuals whose contributions have been particularly instrumental in shaping where I am today.

Firstly, to my family - my parents, my sister, and my aunt - you never wavered in your belief in me. You instilled in me not only a sense of perseverance but a deep appreciation for a journey of self-improvement. Your reminder that progress, no matter how slow, is still progress, has been a guiding principle in my life.

To Dr. Zili Liu, my mentor. I've thoroughly enjoyed the years working with you and getting to know you. You taught me to embrace new perspectives, the importance of every piece of data, and the value of rigorous questioning in experimental approaches. Your patience, intellectual curiosity, and kindness have left a permanent impression on me. You deserve the world, Zili. Thank you so much. We will stay in touch, I promise.

Dr. Philip Kellman, as the PI of my first research lab, you ignited my passion for vision science and set me on a path to graduate study. Thank you for welcoming me into the world of vision science.

Dr. Hongjing Lu, your kindness and inspirational guidance have always been a blessing. Working in your lab was an honor, and our chance encounters have always been a source of joy.

Dr. Scott Johnson, thank you so much for introducing me to the precision of programming stimuli for experiments. It made me want to learn how to program experimental stimuli, which was the catalyst that led me to meeting Dr. Gennady Erlikhman.

To my previous mentors - Dr. Gennady Erlikhman, Dr. Steve Thurman, and Dr. James Kubricht - your influence has been pivotal. The three of you nurtured my interest in vision

science, offered me numerous opportunities to enhance my coding skills, and taught me to appreciate the aesthetics of scientific presentation.

Dr. Luis Jimenez, we've soldiered through so much together. That's why we've forged a bond of enduring friendship and collaboration.

To my extended family — Isaac Prieto, Kelsey Prieto, Dominique Beck, Nico Donovan, Taylor Cartagena, Julian Sar, Preston Gruettner, Brian Wong, Erica Kong, Alex Siegel, Akila Kadambi, Viyehni Fuscher, Maggie Yi, Carolyn Murray — I love you all so, so much.

And last but not least, the students that helped make this dream a reality: Jiantong Liu, Ilene Yang, Alice Wong, Ashley Lee, Avery Jensen, Cody Lejang, Joanna Liu, Kaitlyn Quinn, Lillian Gabrelian, and Michelle Pan, and Gloria Yun. You kids are tough. Together, we achieved what once seemed impossible.

VITA

2012–2014	Research Assistant, UCLA, Human Perception Laboratory, Dr. Philip Kellman
2012–2014	Research Assistant, UCLA, Computational Vision and Learning Laboratory, Dr. Hongjing Lu
2013	B.A. in Psychology, UCLA
2015–2024	Graduate Research Assistant, UCLA, Perceptual Processing and Computation Laboratory, Dr. Zili Liu
2017	M.A. in Psychology, UCLA
2023	Ph.D. Candidate in Psychology, UCLA

PUBLICATIONS

Erlihkman, G., **Xing, Y. Z.**, & Kellman, P. J. (2014). Non-rigid illusory contours and global shape transformations defined by spatiotemporal boundary formation. *Frontiers in Human Neuroscience*, *8*, 1-13.

Xing, Y. & Liu, Z. (2018). A preference for minimal deformation constrains the perceived depth of a stereokinetic stimulus. *Vision Research, 153*, 53–59.

Yu, A., Zhang, R., Silva, A. E., **Xing, Y.**, Thompson, B., & Liu, Z. (2022). Motion opponency at the middle temporal cortex: Preserved motion information and the effect of perceptual learning. *European Journal of Neuroscience*, *56*(12), 6215-6226.

CHAPTER 1

Introduction

Structure from motion (SFM) is the visual system's capacity to deduce the three-dimensional (3D) architecture of objects from their two-dimensional (2D) motions across the retina. This fundamental aspect of human perception enables us to comprehend the depth and form of our surroundings, transforming the flat images that reach our eyes into rich, volumetric scenes as either we or the objects around us shift position. Our understanding of the visual system's interpretation of motion to infer 3D space is largely informed by studies using kinetic depth effect (KDE) displays, consisting either of random dot collections or wireframe objects that represent two-dimensional projections of three-dimensional objects (Wallach & O'Connell, 1953; Green, 1961; Ullman, 1979; Schwartz & Sperling, 1983; Todd, 1984; Ramachandran, Cobb, & Rogers-Ramachandran, 1988; Husain, Treue, & Andersen, 1989; Dosher, Landy, & Sperling, 1989; Treue, Husain, & Andersen, 1991; Norman & Todd, 1993; Hildreth, Ando, Andersen, & Treue, 1995). Through this approach, a random or structured assembly of points moves across a 2D display in a manner that simulates the projection seen from points on the surface of a 3D object in motion. This technique allows observers to perceive a complex 3D structure even in the absence of traditional depth cues, using only the minimal 2D motion cues provided by scattered dots or wireframes.

Computational studies led to two different classes of algorithms for perceiving structure from motion dissimilar in regards to the information that they take as input. Position-based approaches track the positions of points across time to compute 3D positions based on how the points change in position relative to each other between views (Ullman, 1979, 1984;

Grzywacz & Hildreth, 1987). In contrast, velocity-based approaches rely on spatiotemporal derivatives of the optic flow in order to calculate 3D object and motion information (Clocksin, 1980; Longuet-Higgins & Prazdny, 1980; Koenderink & van Doorn, 1975, 1976; Koenderink, 1986). Both types of algorithms encounter issues. Psychophysical evidence suggests that the human structure from motion process uses velocity information rather than position-based information. Psychophysical studies using points with limited lifetimes (<100 milliseconds) have shown that points do not need to stay in view to recover 3D structure and that points must be visible for a certain time rather than exhibit a certain amount of distance traveled when recovering structure from motion; both of these findings are in conflict with what position-based algorithms would suggest (Husain, Treue, & Andersen, 1989; Dosher; Treue, Husain, & Andersen, 1991).

While the investigations into SFM have predominantly centered on the kinetic depth effect (KDE), showcasing how depth and form are perceived through motion-induced transformations, another intriguing phenomenon warrants attention — the stereokinetic effect (Musatti, 1924; Wallach, Weisz & Adams, 1956; Zanforlin, 1988; Proffitt, Rock, Hecht & Shubert, 1992; Todorovic, 1993). First described by Musatti and Duchamp in the early 1900s, the stereokinetic effect refers to the vivid impression of a 3D object that is evoked when certain 2D patterns, often involving nested concentric circles or other geometric forms, are rotated about an axis perpendicular to the image plane. The classic example is a set of nested eccentric circles that, when rotated on a turntable, appear as a cone pointing toward or away from the observer. This occurs despite the absence of explicit depth cues such as contour deformation or perspective foreshortening, typically associated with 3D rotation. Interestingly, the individual circles in these displays do not deform or foreshorten as would be geometrically consistent with their apparent slant. Rather, the depth percept seems to arise primarily from the changing relative positions of the nested contours over time. A variety of stereokinetic stimuli are showcased in Figure 1.1.

Proffitt, Rock, Hecht, and Shubert (1992) proposed that the stimulus basis for the

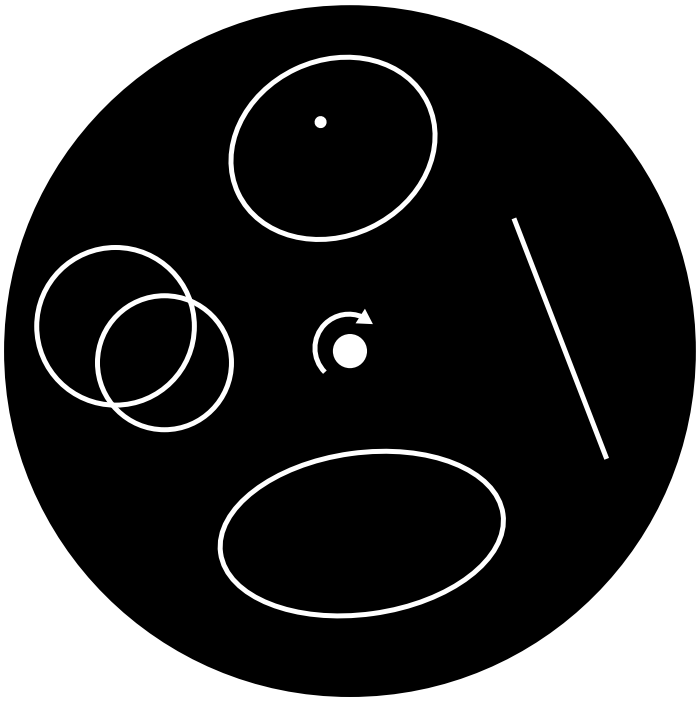

Figure 1.1: **A variety of stereokinetic stimuli.** The stereokinetic effect occurs with a variety of different figures. The image above illustrates how the first stereokinetic stimuli were first displayed to observers. They were cut out of paper and taped onto a rotating turntable that was upright facing the observer. Clockwise from top: Cone, rotating bar, oval, and the focus of this paper: the cylinder.

stereokinetic effect is the presence of between-contour motions consistent with the rotation of a rigid object in depth, in the absence of the appropriate within-contour deformations. Between-contour motions involve the movement and orientation changes of separate contours over time, providing insight into the depth relationships within an object. For instance, when an object rotates, points on different contours at different depths shift relative to one another, with the degree of movement reflecting the depth separation between the contours. Within-contour motions, on the other hand, entail changes in shape and distances between points on the same contour, indicating the surface's orientation relative to the observer. Displays demonstrating the KDE typically feature both types of motion, whereas

those showcasing the stereokinetic effect solely exhibit between-contour motions, signifying object rotation, without concurrent within-contour changes. Proffit et al. 1992 showed that observers experienced compelling depth impressions when viewing stimuli exhibiting simple translations of nested contours relative to each other (e.g., concentric squares), which could not be produced by the traditional turntable method. Their work suggests that the visual system prioritizes relational motion cues over geometric consistency when inferring depth from rotational motions. Previous SFM research indicates that observers' perceptions can be notably impacted by the orientation of an object's rotational axis (Green, 1961; Todd, 1982; Loomis & Eby, 1988, 1999). Perception of depth and rigidity is most pronounced when an object's rotational axis is parallel to the image plane. In fact, when the axis of rotation reaches its maximum slant, becoming precisely parallel with the line of sight, a rotating pattern typically fails to evoke any sense of 3D structure. Stereokinetic stimuli stand out as the singular anomaly within this overarching principle.

In particular, prolonged observation of a 2D rotating ellipse elicits three sequential percepts: (1) the perception of the ellipse rotating rigidly in the image plane, (2) subsequent deformation of the ellipse within the image plane where its contour appears to contract and expand, leading to (3) a final percept of the stereokinetic illusion where the ellipse transforms into the appearance of a rigid 3D circular disk tilted relative to the image plane that rotates both around an axis perpendicular to the image plane and around the surface normal of the disk. Rokers, Yuille, & Liu (2006) proposed that the visual system generates a rigid 3D percept of the rotating ellipse by assuming the motion to be as slow and smooth as possible (Hildreth, 1984; Yuille & Grzwacz, 1989; Weiss, Simoncelli, & Adelson, 2002). In their study, observers were presented with various ellipses and their task was to adjust the perceived rotation around the surface normal of the disk until it disappeared. Rokers et al. 2006, presented theoretical predictions for the perceived motion for the stereokinetic stimulus based on minimizing motion both in 2D and 3D space. Their primary finding was that the perceived motion of the rotating ellipse was nearly identical across all observers sug-

gesting that all observers applied similar assumptions to resolve the perceptual ambiguity when viewing the stereokinetic stimulus. The results indicated that how observers perceived the rotation along the disc's normal directly corresponded to the speed of rotation seen in the image plane, closely matching their theoretical predicts. Consequently, they concluded that when faced with an ambiguous stimulus, the visual system relies on prior assumptions about object motion - specifically assumptions of rigidity, slowness, and smoothness - to generate a stable 3D interpretation.

More recently, in our previous study, Xing and Liu (2018), we demonstrated that the position of the dot within the ellipse influences the perceived height of the stereokinetic cone. When a dot is positioned within the interior of an ellipse and the configuration is rotated around the z-axis, a stereokinetic cone is perceived. The location of the dot significantly influenced the perceived height of the cone, such that the perceived height was tallest when the dot was on the minor axis and shortest when the dot was on the major axis. We hypothesized that the height was constrained by the visual system's preference for a 3D percept that is slowest and maximally rigid. In our experiment, we systematically manipulated the position of the dot, while maintaining its distance to the center of the ellipse, and also varied the aspect ratio of the ellipse.

Observers adjusted a bar positioned on the minor axis of the ellipse. As the ellipse and dot rotated, participants perceived the bar as becoming perpendicular to the plane of the stereokinetic disk. The critical observation was that, with the dot's progression towards the minor axis, the adjustments made by observers consistently indicated their perception of a cone increasing in height. Our results were qualitatively and quantitatively consistent with our hypothesis. The findings illustrated observers' tendency to perceive the apex of the cone at a height that minimized its 3D distance to the surface normal at the center of the circular base, reducing the relative motion between the dot and base. In other words, when presented with an ambiguous stereokinetic stimulus, the visual system prefers the interpretation corresponding to a 3D percept that is slowest (Hildreth, 1984; Yuille & Grzywacz,

1989; Weiss, Simoncelli, & Adelson, 2002) and maximally rigid (Jansson & Johansson, 1973; Ullman, 1979).

Building upon this groundwork, I present two experiments designed to ascertain whether perception of the stereokinetic cylinder is also constrained by a preference for minimal deformation and slow motion. The following questions will serve as guides throughout the thesis: (1) What factors affect the perceived length of the stereokinetic cylinder? (2) Is the perceived length identical across different observers? (3) When presented with the same configuration, do observers perceive the same perceived length of the illusory cylinder each time?

REFERENCES

Andersen, R. A., Treue, S., & Husain, M. (1989). Surface interpolation in three-dimensional structure-from-motion perception. *Neural Computation*, 1(3), 324-333.

Clocksin, W. F. (1980). Perception of surface slant and edge labels from optical flow: a computational approach. *Perception*, 9(3), 253-269.

Dosher, B.A., Landy, M. S., & Sperling, G. (1989). Kinetic depth effect and optic flow - I. 3D shape from fourier motion. *Vision Research*, 29(12), 1789-1813.

Green Jr, B. F. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 62(3), 272.

Grzywacz, N. M., & Hildreth, E. C. (1987). Incremental rigidity scheme for recovering structure from motion: Position-based versus velocity-based formulations. *JOSA A*, 4(3), 503-518.

Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial intelligence*, 23(3), 309-354.

Hildreth, E. C., Ando, H., Andersen, R. A., & Treues, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research*, 35(1), 117-137.

Husain, M., Treue, S., & Andersen, R. A. (1989). Surface interpolation in three-dimensional structure-from-motion perception. *Neural Computation*, 1(3), 324-333.

Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movment of rigid bodies relative to an observer. *Optica acta*, 22(9), 773-791.

Koenderink, J. J., & van Doorn, A. J. (1975). Optic Flow. *Vision Research*, 26(1), 161-179.

Koenderink, J. J., & van Doorn, A. J. (1976). Local structure of movement parallax of the plane. *JOSA*, 66(7), 717-723.

Koenderink, J. J. (1986). Optic flow. *Vision research*, 26(1), 161-179.

Loomis, J.M., & Eby, D.W. (1988). Perceiving structure from motion: Failure of shape constancy. In *Proceedings from the second international conference on computer vision* (pp. 383-391). Washington, D.C.: IEEE.

Loomis, J.M., & Eby, D.W. (1989). Relative motion parallax and the perception of structure from motion. In *Proceedings from the workshop on visual motion* (pp. 204-211). Washington, D.C.: IEEE.

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 208(1173), 385-397.

Musatti, C. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, 3, 105-120.

Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, 53(3), 279-291.

Proffitt, D. R., Rock, I., Hecht, H., & Schubert, J. (1992). Stereokinetic effect and its relation to the kinetic depth effect. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 3.

Ramachandran, V. S., Cobb, S., & Rogers-Ramachandran, D. (1988). Perception of 3-D structure from motion: the role of velocity gradients and segmentation boundaries. *Perception & Psychophysics*, 44(4), 390-393.

Rokers, B., Yuille, A., & Liu, Z. (2006). The perceived motion of a stereokinetic stimulus. *Vision Research*, 46(15), 2375-2387.

Schwartz, B., & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society*, 21, 456-458.

Treue, S., Husain, M., & Andersen, R.A. (1991). Human perception of structure from motion. *Vision Research*, 31(1), 59-75.

Todd, J. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception and Psychophysics*, 36, 97-103.

Todorović, D. (1993). Analysis of two-and three-dimensional rigid and nonrigid motions in the stereokinetic effect. *JOSA A*, 10(5), 804-826.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London B*, 203, 405-426.

Ullman, S. (1984). Rigidity and misperceived motion. *Perception*, 13, 219-220.

Ullman, S., & Yuille, A. (1987). *Rigidity and smoothness of motion* (no. AI-M-989). Massachusetts Institute of Technology Artificial Intelligence Laboratory.

Wallach, H., & O'connell, D. N. (1953). The kinetic depth effect. *Journal of experimental psychology*, 45(4), 205.

Wallach, H., Weisz, A., & Adams, P. A. (1956). Circles and derived figures in rotation. *The American Journal of Psychology*, 69(1), 48-59.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598-604.

Xing, Y., & Liu, Z. (2018). A preference for minimal deformation constrains the perceived depth of a stereokinetic stimulus. *Vision Research*, 153, 53-59.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Computation*, 1(3), 334-347.

Zanforlin, M. (1988). The height of a stereokinetic cone: a quantitative determination of a 3-D effect from 2-D moving patterns without a "rigidity assumption". *Psychological Research*, 50, 162-172.

CHAPTER 2

How Differently Do We See the Same Cylinder?

2.1 Abstract

This study investigates the "stereokinetic cylinder" illusion, where extended observation of two rotating, overlapping circles leads to the perception of a 3D tilted cylinder. We explored factors influencing this perceived cylinder length, consistency of perception across observers, and whether different sensory modalities — monocular linear perspective, binocular stereoscopic vision, and motor reaching — yield coherent measurements of cylinder length. Through experiments involving adjustment of monocular and binocular cues and performing reaching tasks, we focused on the effects of inter-center distance, rotation speed, and circle size on perceived cylinder length. Our findings indicate that only inter-center distance significantly affects perceived length. Correlational analyses reveal a notable consistency in the representation of the stereokinetic cylinder across sensory modalities, indicating a unified internal model that informs both perception and motor actions. This consistency underscores the brain's capacity to integrate diverse sensory inputs into a coherent perceptual output, facilitating accurate interactions with our environment. Crucially, across all measurements, observers consistently perceived a cylinder of finite length, indicating a nuanced balance between the classic rigidity assumption (Jansson & Johansson, 1973; Ullman, 1974, 1984) that predicts an infinitely long cylinder and the preference for slow motion (Hildreth, 1984; Yuille & Grzywacz, 1989; Weiss, Simoncelli, & Adelson, 2002) that suggests a shorter cylinder. The observed individual variability implies different weights assigned to these preferences.

2.2 Introduction

In order to navigate our environment effectively, the visual system rapidly constructs a three-dimensional (3D) interpretation of the world (the distal stimulus) from the two-dimensional (2D) reflection of surrounding surfaces on our retina (the proximal stimulus). In the proximal stimulus, there exist many depth cues that provide information about the 3D layout of the environment. Qualitatively, each cue signals different types of information. Some cues available in the proximal stimulus include binocular disparity, linear perspective, and object motion. Binocular disparity refers to the differences in the horizontal positions of image features because of horizontal separation between the two eyes. Based on linear perspective, the visual system can take advantage of the fact that the proximal stimulus size of an object is inversely proportional to its distance between the object and the observer. When an object rotates, the relative motions of its features can indicate its 3D structure monocularly (Braunstein, 1976, Ullman, 1979).

The visual system's ability to perceive 3D structure and depth relations from 2D motion without additional cues to depth is referred to as structure from motion (SFM). This has been demonstrated experimentally using abstract, simplified sequences of two-dimensional images that imitate how a three-dimensional object in motion would project onto our retina (Wallach & O'Connell, 1953). How 3D structure can be reconstructed from 2D motion has been studied through computational studies and psychophysical studies. Computational studies seek to determine potential algorithms that can be used to recover probable 3D structure and 3D motion based on 2D motion input (Koenderink & van Doorn, 1975, 1976; Koenderink, 1986; Ullman, 1979, 1984; Clocksin, 1980; Longuet-Higgins & Prazdny, 1980; Grzywacz & Hildreth, 1987; Fernandez & Farrell, 2009). Psychophysical studies on SFM have utilized rigid and nonrigid displays to investigate how the human visual system is capable of producing stable 3D percepts despite the inherently ambiguous retinal input (Green, 1961; Ullman, 1979; Schwartz & Sperling, 1983; Todd, 1984; Ramachandran, Cobb, & Rogers-

Ramachandran, 1988; Husain, Treue, & Andersen, 1989; Dosher, Landy, & Sperling, 1989; Treue, Husain, & Andersen, 1991; Norman & Todd, 1993; Hildreth, Ando, Andersen, & Treue, 1995; Aaen-Stockdale, Farivar, & Hess, 2010).

The stereokinetic effect (Musatti, 1924; Wallach, Weisz & Adams, 1956; Zanforlin, 1988; Proffitt, Rock, Hecht & Shubert, 1992; Todorovic, 1993) is a special case of the SFM problem. Typically, a stereokinectic effect stimulus is a two-dimensional (2D) line drawing without a trackable feature such as a corner or an end point of a curve. Instead, such a drawing often consists of circles or ellipses, and the drawing is rotated in the image plane. In theory, such rotation provides no additional information as compared to the static version of the drawing. However, when the drawing is rotated, a three-dimensional (3D) shape is perceived. For example, when a drawing of two circles is rotated (2.1), the two circles are perceived to have different depths, thus forming into a tilted cylinder that is deforming while rotating around the line of sight. This percept is bi-stable in that one circle can be perceived either in front of or behind the other circle. This perception of a 3D shape from a 2D rotating image is called a stereokinetic effect.

Although a stereokinetic effect is created by rotating a 2D drawing, subjectively no single or fixed 2D drawing is perceived to be rotating. Specifically, in the concrete example of two overlapping rotating circles sharing a common axis, neither circle is perceived as rotating around its center. Instead, the center of each circle is perceived to rotate around the rotational axis that is parallel to the line of sight, but each circle itself maintains its orientation and translates without rotating. The simplest explanation for why each circle is not perceived to rotate is related to the correspondence problem (Ullman, 1978; Marr & Ullman, 1981). The motion correspondence problem refers to the challenge of determining which elements in one image or video frame correspond to which elements in another image or frame when there is motion involved. Since each circle is drawn with a solid line, there are no identifiable features on the line. As a result, the visual system is unable to unambiguously match points on the circle across successive moments in time as it rotates. Theoretically, if

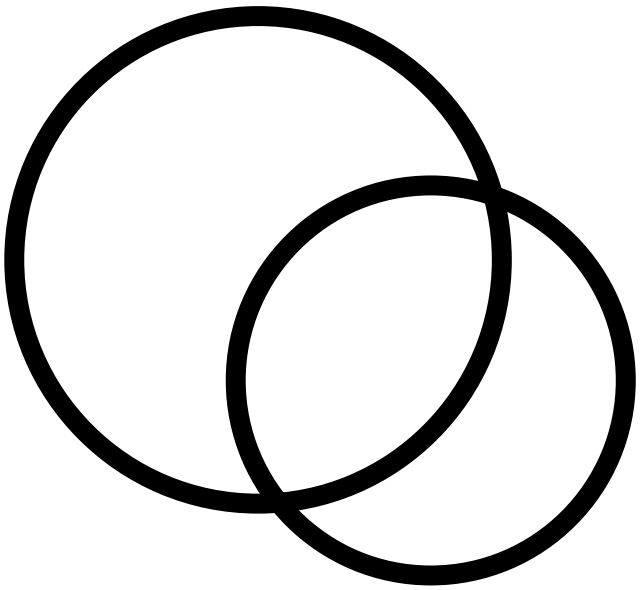

Figure 2.1: **The stereokinetic stimulus.** This figure depicts the two overlapping circles that, when rotated around the line of sight, give rise to the stereokinetic effect. Observers perceive these circles as existing at different depths, thus creating the illusion of a 3D tilted cylinder in motion. The perceptual outcome is bi-stable, allowing for either circle to be perceived as closer or further away from the viewer.

one claims that the circle rotates clockwise with a certain speed, by symmetry, a counter-clockwise rotation with the same speed would be an equally valid percept. With no features to track, all rotational speeds are equally likely. Therefore, the most plausible and stable perceptual solution is that the rotational speed of each individual circle is zero - the circles translate but do not rotate.

Consequently, when two circles separated in depth rotate about their respective centers, they necessarily move relative to each other in the 2D image plane, making the object they form nonrigid. The circles appear to translate and deform relative to each other. The visual system has been suggested to prefer perceiving either a rigid object when possible, or an object that is minimally deforming when rigidity is impossible to achieve (Wallach &

O'Connell, 1953; Jansson & Johansson, 1973; Ullman, 1974, 1984). Separating the two circles in depth reduces such deformation. In fact, when perceived depth difference approaches infinity, as with an infinitely long cylinder, the object appears perfectly rigid because the visual system perceives no relative motion between its ends from any viewpoint.

Although the rigidity assumption could in principle lead to the perception of an infinitely long cylinder, the visually perceived cylinder does not appear infinitely long but instead assumes a well-defined length, as depicted in Figure 2.2. This can be explained by considering additional perceptual constraints beyond rigidity such as the "slow and smooth" motion hypothesis (Hildreth, 1984; Yuille & Grzywacz, 1989; Weiss, Simoncelli, & Adelson, 2002). The "slow and smooth" hypothesis posits that when multiple 3D interpretations are consistent with the 2D retinal projection, the visual system prefers the interpretation that minimizes the speed of motion (the "slow" preference) and that yields the smoothest velocity field (the "smooth" preference). In the case of the stereokinetic cylinder, the "slow" preference implies that the perceived length of the cylinder should be as short as possible according to linear perspective projection. If the cylinder were infinitely long, the far circle would have to be moving infinitely fast to match the retinal velocity of the near circle. A shorter cylinder minimizes the speed of the circles in 3D. Moreover, a finite cylinder yields a smoother velocity gradient between the two circles.

Taken together, the minimal deformation preference described above, which more recently has been termed "smooth" motion preference (Ullman & Yuille, 1987; Weiss, Simoncelli, & Adelson, 2002), combined with the "slow" motion preference, offer a plausible explanation for why observers perceive a stereokinetic cylinder of finite length rather than an infinitely long one. In other words, the perceived length of the cylinder is constrained by a preference for maximal rigidity and slow motion. Previous research in our laboratory focused on extended perception of one type of stereokinetic stimulus, a 2D rotating ellipse with a dot in its interior, revealed the visual system's inclination towards minimal motion (Jansson & Johansson, 1973; Ullman, 1979; Hildreth, 1984; Ullman & Yuille, 1987; Yuille & Grzwacz,

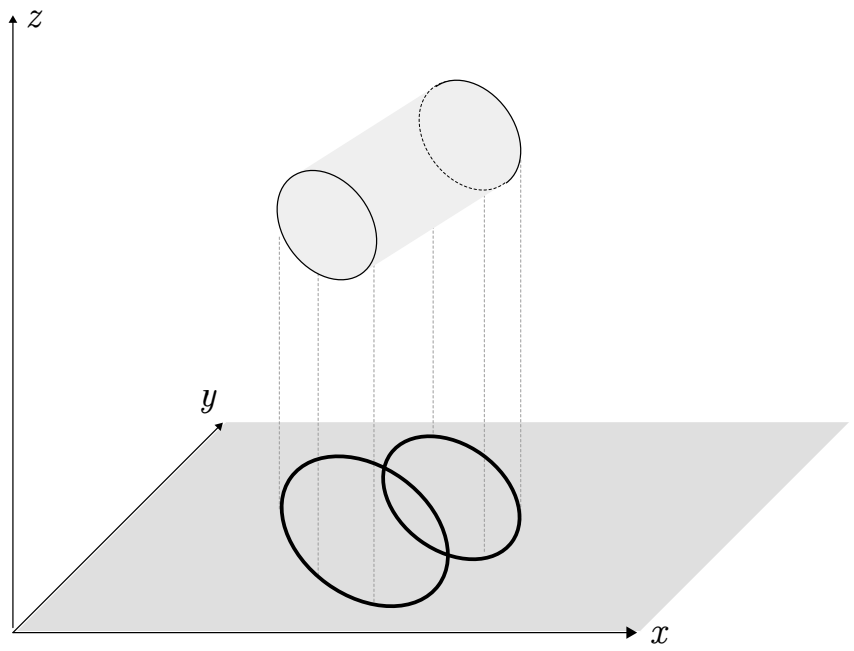

Figure 2.2: **The perceived tilted cylinder.** An illustration of the illusion experienced following extended viewing of two flat, homogeneous, rotating, overlapping rings: a tilted 3D cylinder with a defined height.

1989; Weiss, Simoncelli, & Adelson, 2002), often favoring interpretations characterized by maximal rigidity when presented with stereokinetic stimuli (Rokers, Yuille, & Liu, 2006; Xing & Liu, 2018).

The interplay between the "slow" and "smooth" motion preferences and their contribution to the perception of finite cylinder length remains an open question, without a clear theoretical framework outlining how these preferences might reconcile or vary across individuals. This gap in knowledge forms the basis of our current investigation. By presenting observers with multiple configurations of the two-circle stimulus, we aim to quantify the variations in perceived cylinder length among observers using multiple measurement tools spanning various sensory modalities (monocular linear perspective, binocular disparity, and

motor reaching), setting the stage for future theoretical advancements on how these preferences are reconciled. By doing so, we intend on ascertaining whether different sensory modalities result in congruent depth measurements and to assess the consistency of the cylinder's perceived length, both within subjects and across observers. In turn, by quantifying individual differences in the perceived depth, the current study could provide empirical constraints for extending the "slow and smooth" hypothesis to explain 3D percepts of stereokinetic stimuli. Furthermore, measuring perceived cylinder length across observers allows us to assess whether they converge on the same 3D interpretation, reflecting innate priors, or show individual variability, suggesting idiosyncratic weighting of the "slow" and "smooth" preferences. This in turn can guide the development of computational models specifying how these priors interact to resolve perceptual ambiguities. Although the original "slow" and "smooth" motion theories primarily addressed 2D motion phenomena, our study extends these concepts to the perception of 3D stereokinetic stimuli (Yuille & Grzwacz, 1989; Weiss, Simoncelli, & Adelson, 2002).

Historically, investigations into stereokinetic phenomena have been largely qualitative or have relied on methods with inherent limitations for quantifying perceived depth. For example, Fischer (1956) had participants adjust a gauge only visible to the experimenter to indicate perceived depth. Zanforlin (1988) required participants to simulate grasping the perceived 3D object with their fingers along a graduated ruler. While innovative for their time, these methods present potential confounds that could restrict the accuracy and range of the depth measurements. The lack of direct visual feedback for the participant in Fischer's (1956) gauge adjustment task may have introduced errors or inconsistencies in the depth estimates. Similarly, the physical limitations imposed by the observer's hand size and shape in Zanforlin's (1988) grasping method may have constrained the indicated depth. Moreover, these early studies did not attempt to systematically quantify the perceived size or dimensions of the 3D object, such as the length of a stereokinetic cylinder. This quantitative gap may be partly due to the methodological challenges of accurately measuring perceived

depth in stereokinetic stimuli.

To quantify perceived cylinder length from monocular linear perspective, participants will be presented with various configurations of the two overlapping homogeneous circles. In each trial, one circle is always of a fixed radius and the radius of the other circle, which participants manipulate, is rendered by adding or subtracting a random value from the radius of the fixed circle. Participants are asked to adjust the non-fixed circle until a uniform cylinder is perceived. Given that the circles comprising the ends of a uniform cylinder are identical in radii and the cue of linear perspective, which specifies that an object that is farther away will have a smaller retinal projection than the same object viewed more closely, the circles will never be of the same size if participants *correctly* perceive a 3D cylinder. The circle participants perceive as front of the cylinder should be larger than the circle perceived as the back of the cylinder. The ratio of the radii of the two circles once the observer finishes adjusting the non-fixed circle will be used to calculate the perceived height of the cylinder from monocular linear perspective.

To assess the perceived cylinder depth from binocular stereoscopic vision, we developed a stereo probe where participants manipulated the horizontal binocular disparity of two contrasting sets of concentric diamonds in order to match the perceived distance of front and back of the cylinder from the observer. The horizontal disparities recorded will be used to measure the cylinder length. In regards to the motor measurement, in experiment two, the computer is rotated 90^{o} counter-clockwise away from the observer. A one-way mirror is then positioned 45^{o} relative to the computer monitor, which produces a hologram of the cylinder away from the computer monitor. Observer are asked to reach for the virtual front and back of the cylinder while wearing a dimly lit LED on their index fingers. The difference between their recorded positions of finger locations to the front and back of the cylinder will be used as a measurement of the perceived cylinder depth from motor reaching.

Considering stereokinetic phenomena through the lens of the two-streams hypothesis (Goodale & Milner, 1992) leads to several interesting predictions about how observers would

reach towards the two rotating circles that lead to the illusory 3D cylinder percept. According to their hypothesis, visual information is processed differently depending on whether the observer's goal is purely perceptual or involves motor action. Their theory is motivated by the anatomical separation of visual processing in the primate cerebral cortex into two distinct pathways - the ventral stream projecting to inferotemporal cortex and the dorsal stream terminating in the posterior parietal cortex (Mishkin, Ungerleider, & Macko, 1983). The ventral "vision-for-perception" stream is thought to mediate object recognition and visual awareness, while the dorsal "vision-for-action" stream supports the visual control of skilled actions.

If the two streams operate independently as suggested by the two-streams hypothesis, the motor system guided by the dorsal stream should not be fooled by the stereokinetic illusion and observers' reaching movements should be directed towards the two circles at their actual location, which is at the same depth, rather than reflecting the illusory 3D shape. If observers' reaching movements reflect the perceived depth of the illusory cylinder, this would suggest that the dorsal stream's control of action is influenced by the perceptual representation produced by the ventral stream. This would be inconsistent with a strong dissociation between vision-for-perception and vision-for-action. As such, measuring the reaching movements towards the front and back of the cylinder offers a method to assess the extent of separation or integration between the perceptual and action-guiding visual pathways.

The design of the current study draws inspiration from previous work which examined observer performance across a wide array of stimuli and tasks, emphasizing the importance of repeated measurements (Glennerster, Rogers, & Bradshaw, 1996; Haffenden & Goodale; 2000; Koenderink, van Dorn, Kappers, & Todd, 2001). In particular, Koenderink et al. (2001) investigated the inherent ambiguity in photographs of scenes. Many different real-world scenes could result in the same photograph. Their study was focused on quantifying pictorial relief - how observers perceive depth and form in images - across four stimuli and

four tasks, assessing reliability and consistency in the responses. They aimed to understand how observers' interpretations of the scene align with the concept of a smooth surface and investigate the variance in depth maps across observers and task operationalizations. A key insight from Koenderink's study is the notion of variability in pictorial relief, either among observers for the same stimulus and task or across different tasks for the same stimulus and observer. Surprisingly, in some cases, there was little to no correlation between the indicated 3D structures from the same observer using different response tasks. Eventually, it was found that most of the variability could be reconciled through an affine shearing transformation in depth, suggesting a common approach between their subjects regarding how they resolve visual ambiguities when inferring 3D shape from photographs. The authors highlight the importance of the "beholder's share", which is an observer's unique contribution to resolving scene ambiguities, which often shifts with changes in the task.

In esssence, Koenderink et al. (2001) highlights the importance of accounting for individual perceptual differences when interpreting 3D structures from visual cues like texture, shading, and motion. Especially when presented with ambiguous stimuli such as the stereokinetic cylinder, an observer's unique interpretation can influence the perceived depth. Their research suggests that depth percepts vary not only across individuals but also across different tasks, underscoring that a single task may not fully capture the complexities of perceptual representation. This variance in response patterns can persist even when the same observers evaluate the identical stimulus. Therefore, drawing comparisons across observers requires cautious analysis to avoid masking inherent individual variations. Averaging responses can diminish the visibility of these nuances in perception, hence the importance of examining the individual response patterns to gain a comprehensive understanding of how the visual system processes depth cues.

Relatedly, our study investigates the internal representation of the stereokinetic cylinder's depth by comparing performances across a spectrum of tasks, sensory modalities, and individual observers. By contrasting an individual's perceptual judgments against their motor

actions, and their responses to monocular versus binocular cues, we aim to assess the degree of internal consistency in their perception of the cylinder. Should a consistent response pattern emerge, it would indicate a robust internal model of the cylinder's 3D structure. The coherence between monocular and binocular perceptions would reveal an integration of sensory information, while the alignment between perception and motor actions would reinforce theories of perception-action coupling (Censor, Sagi, & Cohen, 2012; Christensen, Giese, Sultan, Mueller, Goericke, Ilg, & Timmann, 2014; van Andel, Cole, & Pepping, 2018). In addition, the potential of variability in depth perception across observers, despite uniform visual input, could highlight individual differences in interpreting depth. By broadening our analysis to include various observers, tasks, and conditions, we ensure our findings reflect general perceptual strategies rather than idiosyncrasies tied to specific tasks or individuals.

2.3 Experiment 1: Perceptual Measurements of the Perceived Depth of the Stereokinetic Cylinder

While previous studies have qualitatively studied stereokinetic stimuli, there remains a lack of systematic quantitative investigation into the perceived depth and consistency of the effect within and between observers. This experiment was designed to directly quantify the perceived length of the stereokinetic cylinder through a series of perceptual tasks, leveraging various configurations of blue overlapping circles viewed by observers wearing red-blue anaglyph glasses. Central to our investigation is the exploration of how manipulations of inter-center distance, rotational speed, and circle size influences perceived cylinder length. Our approach incorporates direct, quantifiable tasks that enable precise measurement of perceived cylinder length from a monocular linear perspective and binocular disparity. By doing so, we seek to determine if different sensory modalities yield congruent depth estimates and to assess the stability of the cylinder's perceived length, providing empirical constraints for theoretical accounts of depth perception in the stereokinetic effect.

2.3.1 Participants

Seven observers participated in the study (6 females; age range: 18 – 22). All observers were laboratory research assistants familiar with psychophysical experiments. The experiment was approved by the UCLA Institutional Review Board. Informed consent was obtained from all participants, who were treated in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki, 2013).

2.3.2 Apparatus

Stimuli were displayed on a ViewSonic G225f 21" Graphic Series CRT monitor with a resolution of 1280 x 1024 pixels and a 85.024 Hz refresh rate. The viewing distance was 30 cm. Observers used a headrest to stabilize the position of their head. A viewing tube was appended to the computer screen to prevent any additional cues that could affect the perception of depth within the experiment. The computer screen (background luminance of 0.01 cd/m^2) provided the only light source in the room.

2.3.3 Stimuli

The stimuli were developed and displayed using the MATLAB programming language and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). There were three different types of stimuli involved: the blue overlapping circles (forming the stereokinetic cylinder), the sphere wireframe probe (a secondary monocular measurement of the cylinder), and the depth probe defined by binocular disparity (a binocular measurement of the cylinder). Throughout the experiment, the observer wore red-blue anaglyph glasses in order to restrict stimuli to monocular viewing and enable fusion required to manipulate the depth probe.

2.3.3.1 Overlapping Circles

The primary stimulus in this experiment consists of rotating overlapping circles, which, after prolonged viewing, are perceived as a tilted 3D cylinder with a well-defined height. Trial-to-trial we systematically manipulated the inter-center distance between the circles, the size of the front circle (the circle not adjusted by observers), and the rotational speed of the circles to investigate their effects on the perceived length of the stereokinetic cylinder. The inter-center distances between the circles were set at two different measurements: 1 cm (1.91^{o} visual angle) and 2.5 cm (4.77^{o} visual angle). As a reminder, the overlapping rings look like the two circles in Figure 2.1, except are blue in our experiment.

The radius of the front circle, that was not adjusted by observers, was set at either 2.5 cm (4.77^{o} visual angle) or 3 cm (5.72^{o} visual angle). The other circle in each stimulus pair, which observers adjusted, had its radius randomly defined by either adding or subtracting 0.25 cm (0.48^{o} visual angle) from the radius of the front circle at the beginning of each trial. The rotational speed of the overlapping circles was another variable manipulated in this study, with speeds set at either 0.75^{o}/frame or 1.5^{o}/frame around the z-axis, which is perpendicular to the image plane. Both circles were uniformly colored in blue. The luminance of the blue circles was measured at 0.03 cd/m^2 when viewed through the red lens, and at 0.05 cd/m^2 through the blue lens. Due to the use of red-blue anaglyph glasses by observers throughout the experiment, the circles were viewed monocularly, limiting their visual cues to linear perspective exclusively.

2.3.3.2 Sphere Wireframe

In every configuration of the overlapping circles, there are always two intersections. A probe consisting of 12 blue rotated circles, each increasing in rotation by 30^{o}, formed a sphere wireframe, which is positioned randomly at one of the two intersections such that the arc of the front circle provides occlusion cues indicating that the sphere is behind the front

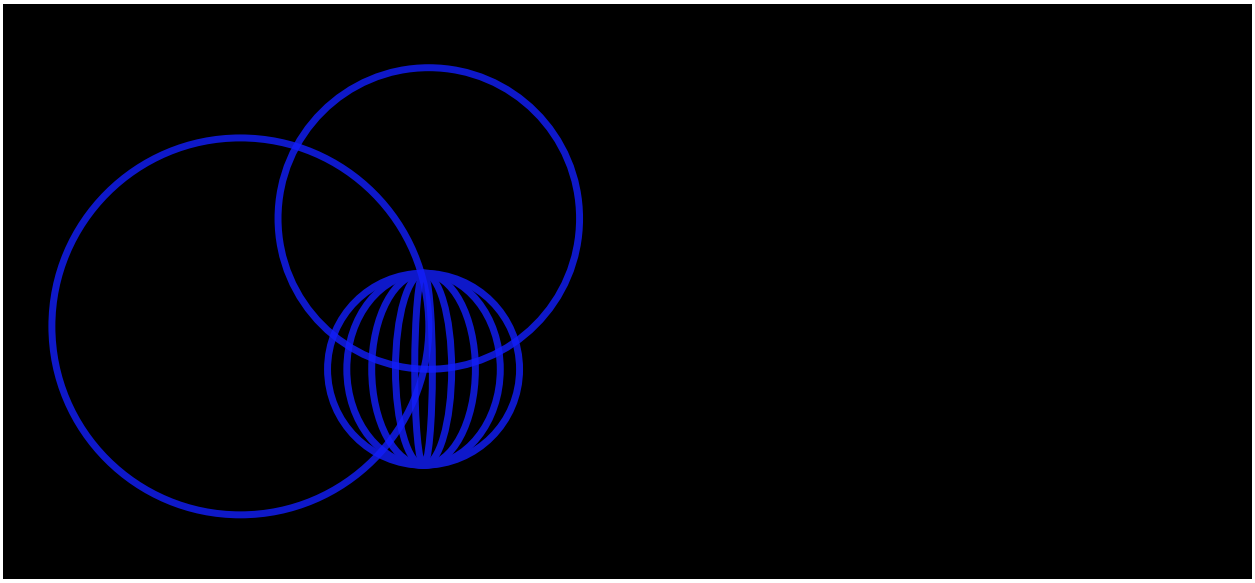

Figure 2.3: **Sphere wireframe probe.** An example of the sphere wireframe probe adjusted by an observer to match their perceived length of the cylinder.

circle of the cylinder. The probe itself rotated around its Y-axis throughout the trial. As the overlapping circles configuration and probe rotate, the probe also provides occlusion cues suggesting that it is in front of the back circle of the cylinder. The sphere wireframe starts at a default diameter of 0.30 cm (0.57^{o} visual angle) and could be increased in size up to 15.24 cm (28.50^{o} visual angle). Viewing through the red lens, the blue sphere probe's luminance registered at 0.03 cd/m^2, while through the blue lens, it increased to 0.05 cd/m^2. Figure 2.3 shows an illustrated example of the sphere wireframe probe.

2.3.3.3 Binocular Disparity Defined Depth Probe

A binocular disparity depth probe composed of red and blue concentric sets of diamonds were programmed in order to examine the perceived length of the cylinder from binocular disparity. For the red set of concentric diamonds, the luminance is measured at 0.05 cd/m^2 when viewed through the red lens, and at 0.03 cd/m^2 through the blue lens. Conversely, the blue set of concentric diamonds exhibits a luminance of 0.03 cd/m^2 when viewed through the

red lens, and 0.05 cd/m^2 when viewed through the blue lens. Each set of concentric diamonds was composed of four identically sized diamonds. The smallest diamond had a width of 1.3 cm (2.48° visual angle), with the radius of each progressively larger diamond increasing by 1.3 cm (2.48° visual angle), to a maximum size of 5.2 cm (9.91°). Observers wore red-blue anaglyph glasses during the experiment enabling each set of diamonds to be visible to only one eye. When viewed through both eyes and fused, they form a single cohesive percept, resulting in the perception of four diamonds instead of eight. By default, the arrangement of the two sets of concentric diamonds was uncrossed. The red set of concentric diamonds began to the left of the blue set of concentric diamonds at a horizontal disparity of -0.34 cm (0.70° visual angle).

It should be noted that during the final portion of each trial when the observer adjusts

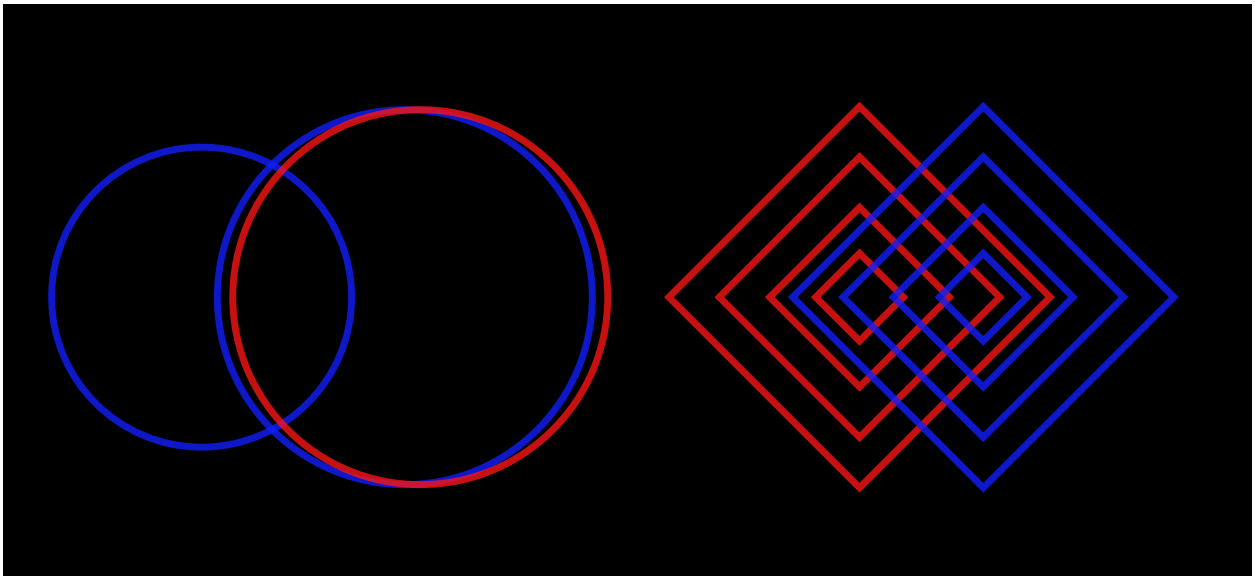

Figure 2.4: **Red circle anchor and binocular disparity depth defined probe.** This imagee illustrates an observer's adjustment of the horizontal disparity between the two sets of diamonds in order to match the perceived distance of the back of the cylinder. Although there are eight diamonds in the image, observers, wearing red-blue anaglyph glasses, fused the two sets into one singular percept composed of four diamonds.

the binocular disparity depth probe, a red circle anchor is affixed to the right of the front blue circle maintaining a crossed horizontal disparity of 0.34 cm (0.70^{o} visual angle) in order to provide some additional visual information regarding where the front of the cylinder is located in 3D space. Through the red lens, the red anchor circle's luminance was measured at 0.05 cd/m^2, whereas through the blue lens, its luminance dropped to 0.02 cd/m^2. Figure 2.4 shows an example of an observer attempting to adjust the horizontal disparity of the two sets of concentric diamonds in order to match the perceived back of the cylinder.

2.3.4 Procedure

Before participating in the main experiments, all observers participated in three calibration activities wearing red-blue anaglyph glasses, each designed with a distinct focus. The calibration sequence began with reaching for real objects at fixed distances, progressed to interacting with a holographic single blue circle seen monocularly, and concluded with reaching for a holographic purple circle visible binocularly. The average of the observer's reaching distance to the front of blue circle, and the purple circle, were used as a measurement of their *perceived* viewing distance when converting the monocular adjustment ratios and depth probe disparity adjustment values to centimeters.

We deliberately used a simplified version of our experimental stimulus during calibration to establish a baseline measurement of each observer's perceptual and motor responses. Understanding how observers perceive and reach for a single circle can more accurately isolate the effects of additional experimental manipulations on perceived depth and motor responses. Reaching for a single circle provides an opportunity to assess and calibrate motor coordination in a straightforward context before introducing the complexity of a bistable percept. This step ensures that any variations in reaching behavior observed in the main experiment can more confidently be attributed to changes in perception rather than unfamiliarity with the task or general difficulties in motor coordination. Reaching for a simpler stimulus allows observers to become familiar with the virtual environment and the mechanics of interacting

with holographic objects.

2.3.4.1 Calibration with Real Objects

The initial calibration with real objects aimed to establish a baseline for observers' spatial perception and reaching accuracy. By having observers reach for objects at known distances, we could assess their ability to accurately estimate distances and execute precise motor actions in a real-world context. This step was essential for calibrating participants' sense of

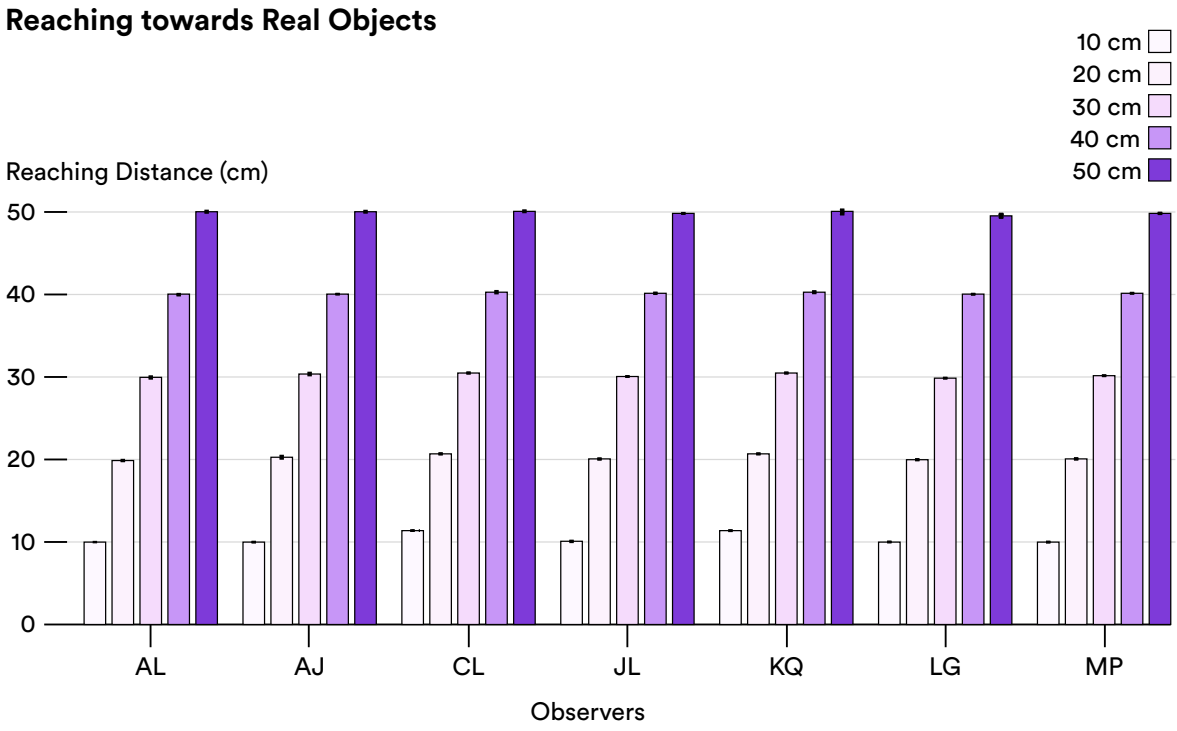

Figure 2.5: **Reaching accuracy across different distances.** This bar graph illustrates the average reaching distances of observers to targets positioned at 10 cm, 20 cm, 30 cm, 40 cm, and 50 cm intervals. Each bar represents the mean reaching distance for one of the five intervals, demonstrating the observers' precision in estimating and reaching towards targets at varying distances. Smaller standard error bars indicate a high level of accuracy and consistency in the participants' reach

space and ensuring their comfort with reaching tasks.

Markers were positioned perpendicular to the observer at intervals of 10 cm, starting from 10 cm and extending up to 50 cm away from the observer. Wearing red-blue anaglyph glasses, the observer sequentially reached for the end of each marker with their index finger, repeating the sequence three times. The purpose of this activity was to familiarize participants with the act of reaching towards targets at known distances, helping to calibrate their sense of space and reach accuracy in a real-world context. It provides a baseline measure of each participant's ability to estimate distances and execute reaching movements accurately. By increasing the distances at intervals of 10 cm, the calibration covers a range of distances that participants might encounter during the experiment, ensuring that they are comfortable and proficient with both near and far reaches without any prior specific training. Observers' average reaching distance to the objects are shown in Figure 2.5.

2.3.4.2 Calibration with a Blue Circle Hologram

Following the real-object calibration, participants engaged with a holographic blue circle, which could only be viewed monocularly. This activity served to acclimate participants to interacting with holographic objects and to assess their depth perception based solely on monocular cues. The simplicity of this stimulus allowed for an evaluation of participants' ability to estimate depth and distance without the aid of binocular cues, setting a foundation for understanding their perceptual adjustments in the subsequent, more complex experimental conditions.

In this calibration step, participants wearing red-blue anaglyph glasses reached towards a holographically projected blue circle, with a radius of 3 cm, matching a key variable from Experiment 1. The task required observers to accurately estimate the location of the holographic circle and reach towards it five times. Observers' average reaching distance to the blue circle is displayed in Figure 2.6.

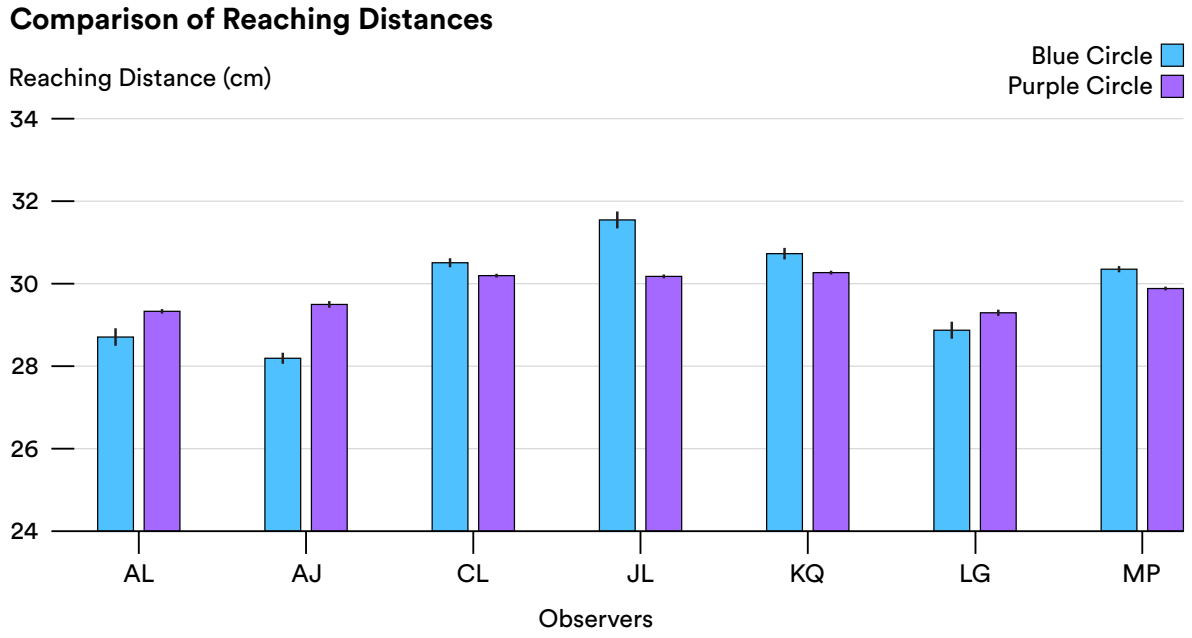

Figure 2.6: **Variations in observers' reaching distances to the holographic circles**. Observers varied when reaching for the blue and purple circles. Performance is more consistent when reaching towards the purple circle, which was visible binocularly. The error bars represent standard error.

2.3.4.3 Calibration with a Purple Circle Hologram

The final calibration activity entailed reaching for a holographic purple circle, where the purple hue was created by combining the blue and red colors utilized in the stimuli of the main experimental conditions. Unlike the blue circle, the purple circle was visible to both eyes. This task directly prepared participants for the experimental conditions involving binocular cues, allowing us to assess their ability to integrate visual information from both eyes for accurate depth perception and motor coordination. By calibrating participants' responses to a binocularly perceived object, we could establish a control for assessing the accuracy of depth perception and reaching movements across the experiments. Moreover, this step served as a verification of whether the hologram was correctly positioned 30 cm from

the observer. Observers reached for the purple circle five times. Observers' mean reaching distance to the purple circle is displayed in Figure 2.6.

Statistical analysis revealed no significant difference in reaching distances between the blue ($M = 30.05$, $SD = 1.15$) and purple circles ($M = 30.10$, $SD = 0.45$); $t(6) = 0.11$, $p = 0.92$, indicating that binocular cues did not notably impact reaching behavior in this context. In contrast, consistency in reaching was higher for the purple circle, as reflected by the lower standard error. A significant difference in the standard errors ($t(6) = -4.93$, $p = 0.002$) between reaching for the purple ($M = 0.05$, $SD = 0.05$) and blue ($M = 0.15$, $SD = 0.05$) circles indicates a marked increase in precision when the circle was visible to both eyes. Moreover, the observed reaching distances for the purple circle were closer to its actual distance of 30 cm, compared to the blue circle, for each observer. This trend may reflect a subtle influence of binocular depth cues on the precision of reaching behavior, despite the lack of statistical significance. Interestingly, subjective perceptions of depth varied among observers, with a majority reporting the blue circle as appearing farther away than the purple circle.

2.3.4.4 Experimental Procedure

At the beginning of each trial, the overlapping circles were presented on the computer screen. One of the circles was briefly flashed three times in order to indicate to the observer that it was the circle that they would manipulate. Immediately afterwards, the configuration was rotated either clockwise or counterclockwise. Once the observer perceived depth between the two circles, they were instructed to maintain the perception of the flashed circle as the back of the cylinder and adjust the radius of this circle by pressing the up or down arrow keys until the configuration appeared to them as a uniform cylinder. Observers were informed that achieving a 'uniform' cylinder did not involve adjusting the back circle to match the size of the front circle. Instead, it was specified prior to experimentation that a 'uniform 'cylinder meant that the front and back appear different in size due to their distinct perceived

distances from the observer, while keeping their objective sizes identical. In essence, because the cylinder is perceived to be 3D, two circles that are objectively identical in size will exhibit different visual angles due to linear perspective. The observer proceeded to the next stage of the trial by pressing the spacebar once the radius adjustment was complete.

After the observer finished adjusting the radius of the back circle to achieve a perceived uniform cylinder, the display remained unchanged, maintaining the adjusted size of the back circle. Subsequently, a sphere wireframe probe appeared randomly at one of the two intersections of the overlapping circles. The initial diameter of the sphere probe at the beginning of each trial was 0.30 cm (0.57° visual angle). The observer's task was to adjust the diameter of the sphere probe such that it matched the perceived length of the stereokinetic cylinder. Upon completing this adjustment, the observer pressed the spacebar to proceed to the final stage of the trial.

In the final portion of the trial, the sphere wireframe probe disappeared, the red and blue sets of concentric diamonds forming the depth probe appeared to the right of the overlapping circles. Simultaneously, a red circle was affixed onto the right of the front circle in order to facilitate binocular viewing of the front of the cylinder. When viewed wearing red-blue anaglyph glasses, the red and blue sets of concentric diamonds were fused such that observers only perceived one set of diamonds rather than two sets of diamonds. Initially, the observers adjusted the horizontal disparity between the two sets of concentric diamonds to match the perceived distance of the front of the cylinder as indicated by the red and blue stereograph pair of circles. Once satisfied with their adjustment, the observer pressed the spacebar key and the depth probe returned to its default position with an uncrossed disparity of 0.34 cm (0.70° visual angle). Afterwards, the observer adjusted the horizontal disparity between the two sets of diamonds in order to match the perceived distance of the back of the cylinder. Once the observer was done, the observer pressed the spacebar key in order to move to the next trial.

2.3.5 Design

Our study utilized a 2 x 2 x 2 within-subjects factorial design to investigate perception of the stereokinetic cylinder. We varied three variables: the inter-center distance between overlapping circles with levels at 1 cm and 2.5 cm, the radius of the non-adjustable circle at either 2.5 cm or 3 cm, and the rotational speed at either 0.75^{o}/frame or 1.5^{o}/frame. To add variability, the radius of the circle adjusted by observers changed by either adding or subtracting 0.25 cm from the fixed circle's radius, resulting in eight unique display conditions. Each participant was presented with each display five times, resulting in a total of 40 trials, with the entire experiment taking an average of 50 minutes to complete. Each observer participated in the experiment three times.

Throughout the experiment, we recorded four dependent variables to assess depth perception. The primary measurement was the final ratio of the back circle's radius to the front circle's radius after adjustment by observers to create a uniform cylinder, later converted into centimeters. The secondary measurement was the final diameter of the sphere wireframe probe, adjusted by observers to reflect their perceived length of the cylinder. The last two measurements focused on perceived cylinder length based on binocular disparity. In the final stage of each trial, observers adjusted the horizontal disparity between sets of concentric diamonds to match the perceived distances of the cylinder's front and back, which were then converted to 3D locations.

2.4 Results

Our first analysis sought to determine whether observers consistently perceived the rotating overlapping circles as 3D cylinders. This question was grounded in the principles of linear perspective, which posits that objects at different distances from an observer project at varying sizes on the retina: distant objects appear smaller than closer ones. In the context of our stimulus — two identical circles representing the ends of a uniform cylinder — linear

perspective suggests that for an accurate perception of a 3D cylinder, the circles should not appear identical in size. Instead, the ratio of the back circle's radius to the front circle's radius should consistently be less than one.

To evaluate this hypothesis, we compiled the adjustment ratio data from all observers across each experimental condition. We performed one-sample t-tests comparing the mean ratio to the hypothesized null mean of 1 for each condition. The analysis indicated that observers consistently perceived depth, as the mean ratio significantly deviated from 1 in all conditions. For an inter-center distance of 1 cm and a rotation speed of 0.75°/frame, circle radii 2.5 cm and 3 cm yielded mean ratios of 0.91 ($SD = 0.03$) and 0.92 ($SD = 0.03$), with $t(6) = -8.06$, $p < 0.001$ and $t(6) = -7.76$, $p < 0.001$, respectively. When the rotation speed was increased to 1.5°/frame for the same inter-center distance, the mean ratios for circle radii 2.5 cm and 3 cm were 0.91 ($SD = 0.03$) and 0.92 ($SD = 0.03$), with $t(6) = -7.61$, $p < 0.001$ and $t(6) = -7.52$, $p < 0.001$, respectively.

At an inter-center distance of 2.5 cm and a rotation speed of 0.75°/frame, circle radii 2.5 cm and 3 cm presented mean ratios of 0.80 ($SD = 0.03$) and 0.81 ($SD = 0.04$), with $t(6) = -14.95$, $p < 0.001$ and textitt(6) = -11.45, $p < 0.001$, respectively. With the rotation speed set to 1.5°/frame at the same inter-center distance, the mean ratios for circle radii 2.5 cm and 3 cm were 0.81 ($SD = 0.04$) and 0.81 ($SD = 0.04$), with $t(6) = -12.84$, $p < 0.001$ and $t(6) = -12.99$, $p < 0.001$, respectively. The significant t-values and virtually zero p-values across all conditions strongly suggest that participants consistently perceived a 3D cylinder, as indicated by the mean ratio significantly differing from 1. This effect was observed regardless of the variations in inter-center distance, rotation speed, and circle size, indicating a robust perception of depth when presented with the different configurations of overlapping circles. This consistent perception of depth, demonstrated by the significant deviation of the mean ratio from 1, is visually represented in Figure 2.7. Here, the clustered bar chart illustrates the mean adjustment values by inter-center distance, rotation speed, and circle size across different experimental manipulations.

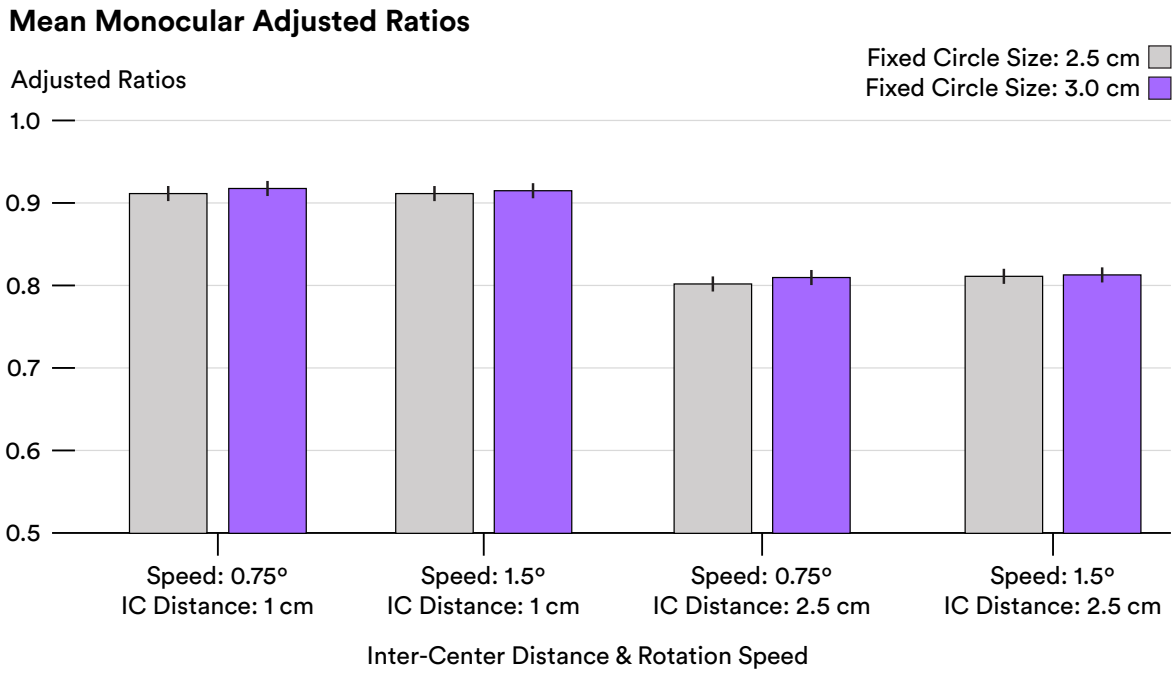

Figure 2.7: **Mean adjustment ratios across all experiment conditions.** The bar chart shows the mean adjustment values made by observers to the back circle's radius to achieve a uniform cylinder perception, grouped by inter-center distance (1 cm and 2.5 cm), with sub-groupings for rotation speed (0.75^{o} and 1.5^{o}) and circle size (2.5 cm and 3 cm). Bars represent the average of all participants' adjustments, with gray bars indicating a circle size of 2.5 cm and purple bars indicating a circle size of 3 cm. The results suggest that the inter-center distance has a significant effect on the mean adjustment values, with smaller adjustments required for a larger inter-center distance.

After investigating whether observers consistently perceived a 3D cylinder, the next analysis was focused on evaluating the observers' ability to reliably adjust the depth probe. Please recall, in each trial, as the depth probe appeared, a red circle anchor was affixed to the front blue circle, to help provide a reference point to observers. Observers were first tasked with aligning the depth probe in 3D space alongside the front of the cylinder. To assess the precision of these adjustments, we aggregated the horizontal disparity between the red and

blue sets of concentric diamonds data across seven observers when matching the front of the cylinder across all experiment conditions.

A $2 \times 2 \times 2$ repeated measures ANOVA was conducted with horizontal disparity as the dependent variable and inter-center distance, rotation speed, and circle size as the independent variables. The ANOVA results indicated that there were no significant main effects for inter-center distance ($F(1,6) = 1.86, p = 0.22$), rotation speed ($F(1,6) = 0.29, p = 0.61$), or circle size ($F(1,6) = 0.15, p = 0.71$). Furthermore, there were no significant interaction effects between inter-center distance and rotation speed ($F(1,6) = 0.02, p = 0.89$), inter-center distance and circle size ($F(1,6) = 0.99, p = 0.36$), rotation speed and circle size ($F(1,6) = 0.034, p = 0.85$), nor for the three-way interaction between inter-center distance, rotation speed, and circle size ($F(1,6) = 2.94, p = 0.14$). The mean horizontal disparities observed across conditions ranged narrowly from 0.35 cm to 0.36 cm, with standard error of 0.002 across all conditions.

The results of the ANOVA suggest that observers' adjustments of the depth probe did not significantly deviate from the objective disparity of 0.34 cm between the blue front circle and its red circle anchor across the different conditions tested. In other words, observers were able to reliably match the depth probe to the front of the cylinder, as evidenced by the non-significant disparities from the objective standard. The depth probe adjustment data are visually summarized in Figure 2.8.

In sum, observers not only consistently perceive a 3D cylinder when viewing the experimental displays, but they are also able to reliably adjust the depth probe. Next, we aimed to identify which experimental variables - inter-center distance, rotation speed, and circle size - significantly influence the perceived depth of the cylinder . We employed three repeated measures ANOVAs each distinguished based on the perceptual measurement that was used as the dependent variable. In the first repeated measures ANOVA, the dependent variable was observers' adjustment ratio while inter-center distance, rotation speed, and circle size served as the independent variables. The analysis revealed a significant main effect of inter-center

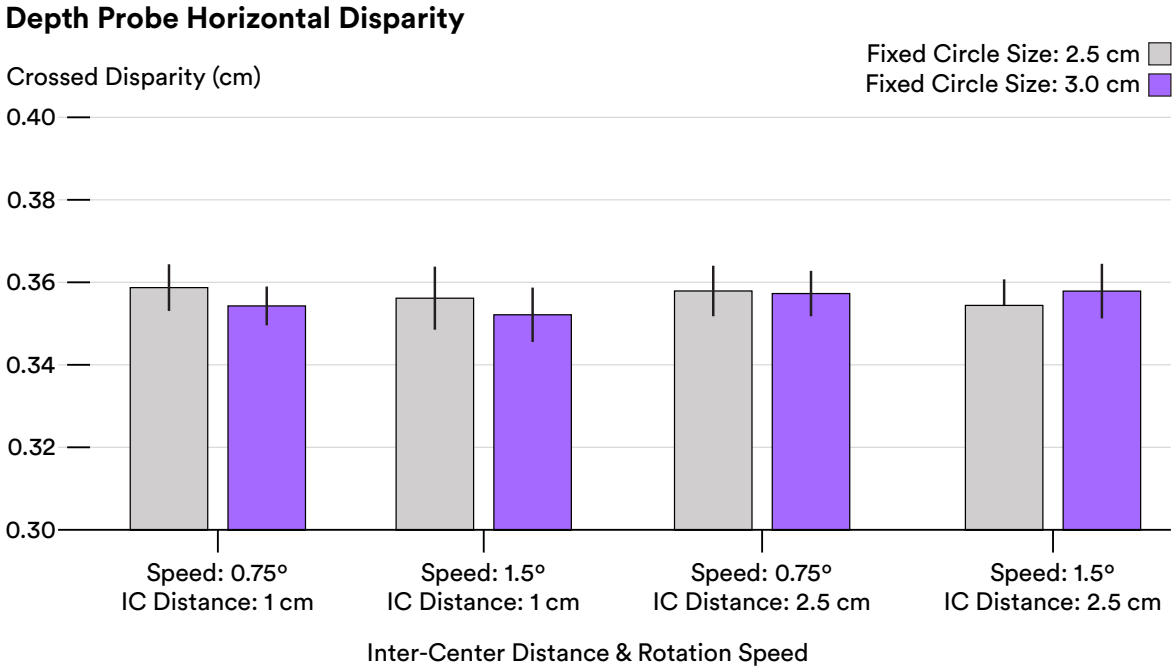

Figure 2.8: **Mean depth probe disparity when matching the front of the cylinder.** This figure illustrates the mean horizontal disparities between the red and blue sets of concentric diamonds as adjusted by observers to match the perceived front of the stereokinetic cylinder. Data are presented for each combination of inter-center distance and rotation speed, across circle sizes of 2.5 cm and 3 cm. Error bars represent standard error. None of the adjustments significantly deviated from the objective standard. The results demonstrate observers' ability to reliably adjust the depth probe.

distance on the adjustment ratios, $F(1, 6) = 76.16, p < 0.001$, indicating that this factor substantially influenced the perceived depth of the cylinder. An increase in inter-center distance from 1 cm to 2.5 cm resulted in a notable decrease in the mean adjustment ratio from 0.92 ($SE = 0.01$) to 0.81 ($SE = 0.01$). Operationally, this indicates that observers proportionally reduced the radius of the back circle in response to greater inter-center distances, suggesting the perception of an elongated cylinder. In essence, as the inter-center distance increased, observers' adjustments corresponded to a perceived increase in the cylinder's length.

In contrast, neither rotation speed, $F(1,6) = 2.27, p = 0.18$, nor circle size, $F(1,6) = 3.10, p = 0.13$, showed a significant main effect, suggesting that these variables alone did not significantly affect the adjustment ratios. Furthermore, the interaction effects between inter-center distance and rotation speed, $F(1,6) = 3.09, p = 0.13$, inter-center distance and circle size, $F(1,6) = 0.06, p = 0.81$, as well as the three-way interaction among all independent variables, $F(1,6) = 0.39, p = 0.53$, were not significant.

These results suggest that while the perceived depth of the cylinder is robustly influenced by the inter-center distance between the circles, it appears to be relatively unaffected by the speed of rotation and the size of the circles when considered independently. The lack of

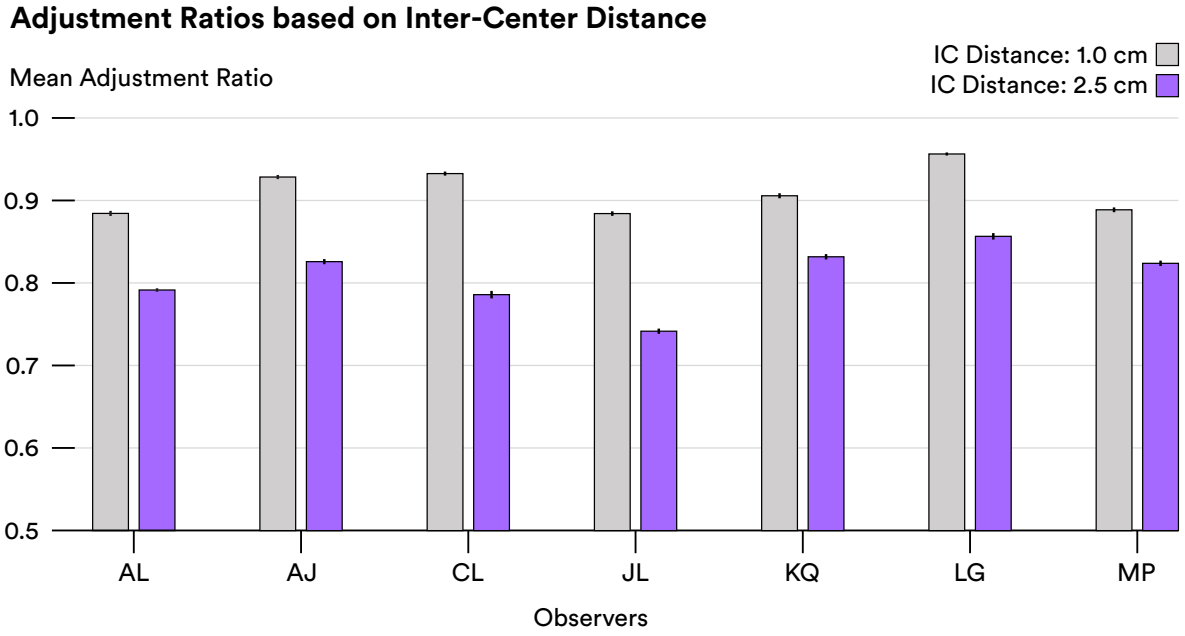

Figure 2.9: **Variability in mean adjustment ratios at an inter-center distance of 1 cm and 2.5 cm.** Assuming the perceived viewing distance of the front of the cylinder is identical across all observers, the perceived cylinder at an inter-center distance of 2.5 cm is significantly smaller than the perceived cylinder at an inter-center distance of 1 cm across all observers. Notably, the y-axis scale ranges from 0.7 to 1.0 and error bars depict standard error measurements.

significant interaction effects further indicates that the combination of these factors does not interact in a way that significantly influences the monocular adjustment ratios. Because the analysis did not indicate a significant influence of the speed of rotation and circle size, observers' monocular adjustment ratios for the two experimental inter-center distances, 1 cm and 2.5 cm, are plotted in Figure 2.9, after collapsing across both rotation speed and circle size.

Next we conducted a repeated measures ANOVA using the absolute difference in centimeters between the crossed and uncrossed horizontal disparities of the depth probe as the dependent variable while inter-center distance, rotation speed, and circle size served as the independent variables. In our experiment, crossed disparities were recorded as positive values, while uncrossed disparities were recorded as negative values. This approach stands in contrast to transforming horizontal disparities into perceived distances and subsequently calculating the perceived cylinder length, based on binocular disparity, as the difference between these two estimates. By directly analyzing the absolute differences in disparities, we avoid the potential complexities and additional variability that could arise from estimating perceived distances. This method provides a more straightforward assessment that is less likely to be influenced by the transformations and assumptions required to calculate perceived distances. Once again, the analysis revealed a significant linear effect of inter-center distance, $F(1,6) = 98.85, p < 0.001$, indicating a strong influence on the dependent measure. As inter-center distance increased, the absolute disparity difference when matching the front and back of the cylinder increased from 0.75 cm ($SE = 0.06$) to 1.24 cm ($SE = 0.1$). This suggests that observers perceived a longer cylinder and positioned the depth probe significantly further away when matching the cylinder's back at the larger inter-center distance.

Conversely, rotation speed did not show a significant linear effect, $F(1,6) = 0.95, p = 0.37$. Similarly, circle size alone did not significantly affect the dependent measure, $F(1,6) = 0.05, p = 0.83$. Furthermore, the interaction between inter-center distance and rotation

speed was not significant, $F(1,6) = 0.04, p = 0.85$, nor was the interaction between inter-center distance and circle size, $F(1,6) = 0.78, p = 0.41$. The interaction between rotation speed and circle size was also non-significant, $F(1,6) = 0.23, p = 0.65$. Finally, the three-way interaction between inter-center distance, rotation speed, and circle size did not reach significance, $F(1,6) = 2.21, p = 0.19$. These results suggest that the perceived depth of the cylinder, as measured by the absolute differences in disparities, is predominantly influenced by the inter-center distance, with other factors playing a less significant role. The impact of inter-center distance on depth probe adjustment is visually represented in Figure 2.10. Here, we observe a clear increase in absolute disparity differences as the inter-center distance extends, corroborating the notion that observers perceive a more profound depth at an inter-

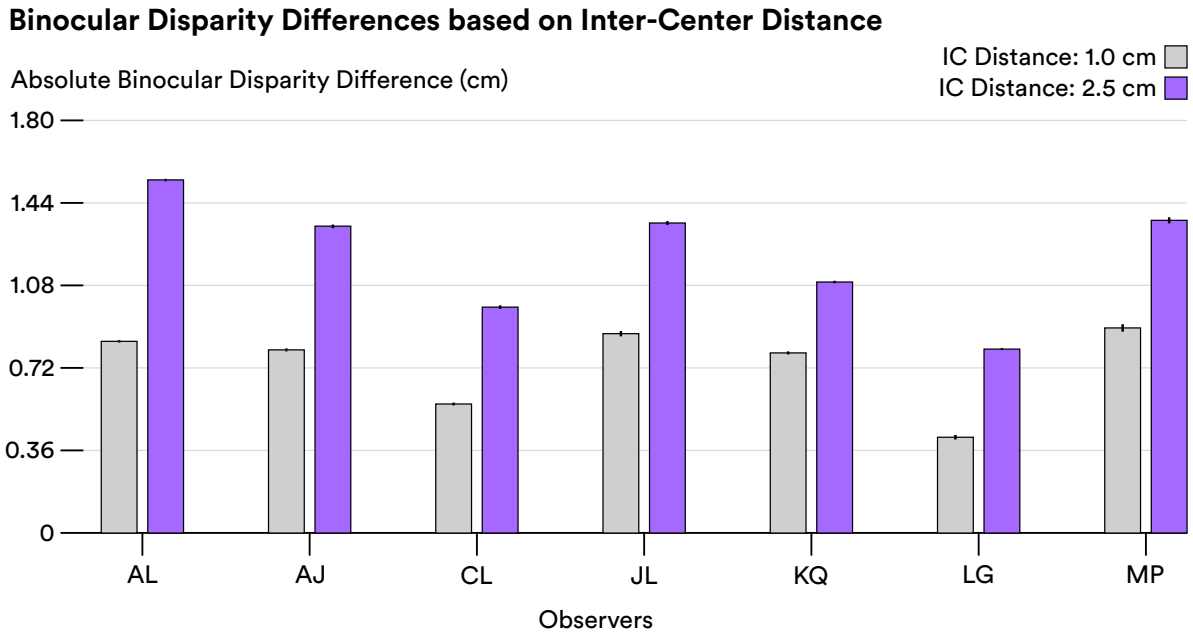

Figure 2.10: **Comparison of absolute depth probe disparity differences at an inter-center distance of 1 cm and 2.5 cm.** The increase in absolute disparity difference as inter-center distance increases suggests that observers positioned the depth probe significantly farther when matching the back of the cylinder at the larger inter-center distance. Error bars depict standard error measurements.

center distance of 2.5 cm compared to 1 cm.

After establishing that inter-center distance is the primary factor influencing the perception of cylinder length as evidenced by the adjustment ratio and binocular disparity data, we next aim to determine if this trend persists within our third perceptual measure: the sphere wireframe probe. We conducted a repeated measures ANOVA using the adjusted diameters of the sphere wireprobe as the dependent variable and inter-center distance, rotation speed, and circle size as the independent variables. The analysis indicated a significant effect of inter-center distance, $F(1, 6) = 46.06, p < 0.001$, which was markedly evident. At an inter-center distance of 1 cm, the mean adjusted diameter of the sphere wireframe probe was 3.04 cm ($SE = 0.25$), whereas at 2.5 cm, this mean diameter significantly increased to 6.80 cm ($SE = 0.46$). This pronounced increase in the sphere's diameter suggests that observers perceived a longer cylinder when viewing configurations with the larger inter-center distance. The mean adjusted diameters of the sphere probe are displayed in Figure 2.11. after collapsing across rotation speed and circle size.

No significant effects were found for rotation speed, $F(1, 6) = 2.24, p = 0.19$, or circle size, $F(1, 6) = 2.87, p = 0.14$. Additionally, there were no significant interaction effects between inter-center distance and rotation speed, $F(1, 6) = 2.31, p = 0.18$, inter-center distance and circle size, $F(1, 6) = 0.64, p = 0.46$, or rotation speed and circle size, $F(1, 6) = 0.04, p = 0.85$. The three-way interaction between inter-center distance, rotation speed, and circle size was also non-significant, $F(1, 6) = 0.58, p = 0.47$. These results corroborate the predominance of inter-center distance in influencing the perceived depth as measured by the sphere wireframe probe, aligning with the earlier findings concerning adjustment ratios and binocular disparities.

Thus far, our analyses have highlighted the pivotal role of inter-center distance in shaping the perceived depth of the stereokinetic cylinder. Across monocular and binocular measures, inter-center distance emerged as the dominant factor, significantly influencing observers' perception of the cylinder's length. Whether assessing the ratio of circle sizes, the absolute

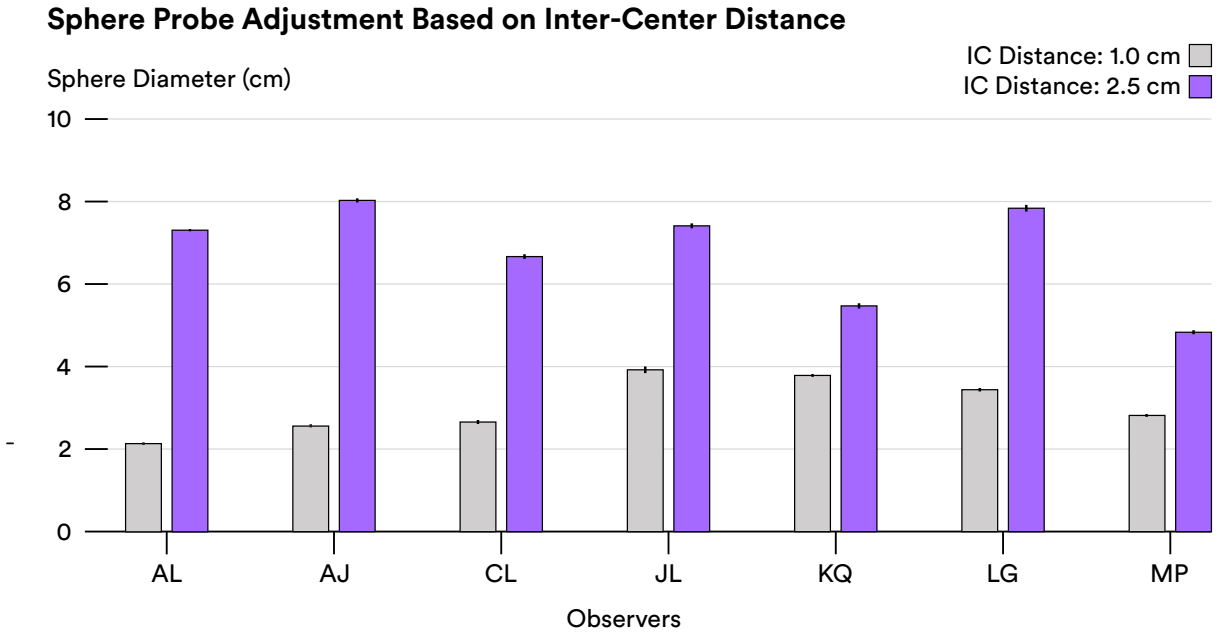

Figure 2.11: **Comparison of sphere diameters at 1 cm and 2.5 cm.** This figure illustrates the contrast in sphere diameters between inter-center distance of 1 cm and 2.5 cm, highlighting the influence of inter-center distance on perceived cylinder length.

disparities, or the adjusted sphere wireframe diameters, the trend remained consistent: as inter-center distance increased, so did the perceived depth.

2.5 Experiment 2: Motor Reaching Measurement of the Perceived Depth of the Stereokinetic Cylinder

In experiment 2, we examine the perceived length of the cylinder through a motor reaching task. The experiment computer is rotated 90^{o} such that the display is perpendicular to the observer's line of sight. A one-way mirror is positioned in front of the monitor at a 45^{o} angle, away from the computer monitor, with the reflective side facing the computer monitor. This arrangement produces a hologram of the cylinder that appears to be floating in mid-air in the observer's line of sight. Observers are asked to reach towards their perceived front and back

ends of the cylinder. The primary goal of this experiment is to measure the perceived cylinder length based on motor reaching. This experiment uses the same independent variables and levels as Experiment 1 to maintain consistency and facilitate comparison between visual and motor-based measurements of the perceived cylinder length.

2.5.1 Participants

All seven observers from Experiment 1 also took part in Experiment 2. This continuity was deliberate, aiming to comprehensively assess the consistency and variation in perceived cylinder length among the same group of individuals using different measurement methods.

2.5.2 Apparatus

With their head position stabilized by a chin rest, observers were presented with the same cylinder configurations as in experiment 1. The presentation was achieved using a CRT monitor positioned to the left of an upright one-way mirror arranged at a 45^{o} angle. The mirror reflected the image displayed on the monitor, creating the illusion that the hologram was suspended in space beyond the mirror. Compared to Experiment 1, the monitor was rotated 90^{o} counterclockwise from its default orientation and moved closer to the observer. The distance between the observer's eyes and the hologram was 30 cm. For monocular perception of the hologram, participants wore red-blue anaglyph glasses.

To aid in reaching for the hologram, a white LED was securely attached to the tip of the participant's index finger. This arrangement allowed the participant to visually track their finger while reaching for the cylinder, while also enabling the experimenter to observe the precise location of their reaching movements. A measuring tape was positioned parallel to the participant's line of sight halfway up the vertical height of the computer screen, immediately to the right of the mirror. This placement ensured that the measuring tape served as an accurate reference for recording the distances reached by the observers. The experimental

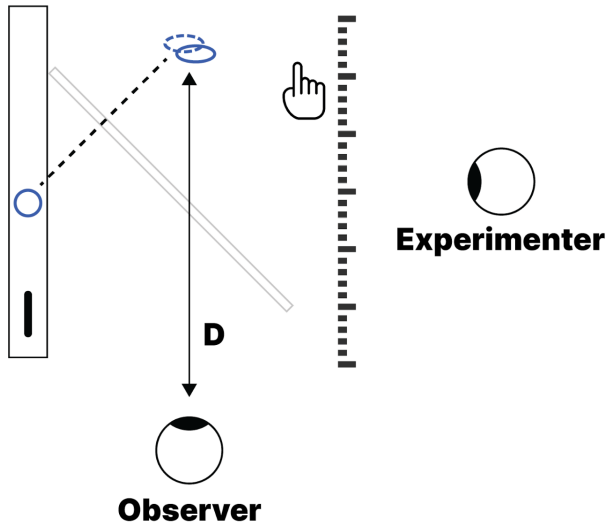

Figure 2.12: **Aerial view of experimental setup.**The computer is rotated 90^{o} counter-clockwise from its default position and moved closer to the observer. The half-silvered mirror is appended to the computer monitor at a 45^{o} angle away from the screen. The illusion is positioned at a distance of D from the observer where D is 30 cm. The experimenter is seated out of view to the right of the observer. The experimenter is facing the measuring tape that is hung parallel to the observer's line of sight in order to record the observer's reaching distances.

setup is illustrated in Figure 2.12.

2.5.3 Procedure

During each trial, first the observer adjusted the radius of the back circle until a perfect cylinder was perceived. Once the observer finished adjusting the back circle, they proceeded to reach for the front of the cylinder with their index finger. The experimenter recorded the distance reached by the observer. Next, the observer extended their reach to touch the back of the cylinder, and the corresponding reached distance was recorded. This process was

repeated for each trial. Throughout the reaching task, observers were asked to reach for the front and back of the cylinder in one continuous motion while maintaining their index finger at the location of the perceived circle they were reaching towards.

2.5.4 Design

Building upon Experiment 1, Experiment 2 adopts a similar 2 x 2 x 2 within-subjects factorial design, systematically manipulating the same three independent variables at identical levels. This approach was intentional and allows for a direct comparison of perceptual and motor-based measurements of the stereokinetic cylinder's perceived depth across different experimental conditions. Each participant encountered each display five times, resulting in 40 trials per participant. On average, the completion time for this experiment was approximately 50 minutes.

Three primary dependent variables were recorded. Similar to experiment 1, we recorded the ratio of the back circle's adjusted ratios to the front circle's radius after participants adjusted the back circle until the cylinder was uniform. This measure continues to serve as a a monocular measurement of perceived cylinder length. After adjusting the back circle, participants were tasked with reaching towards the perceived front of the cylinder. Following the reach towards the cylinder's front, participants extended their reach to what they perceived as the back of the cylinder. The difference between the two reaching measurements (front and back reaches) is analyzed as the perceived length of the cylinder based on motor reaching.

2.6 Results

In both experiment 1 and experiment 2, observers were asked to provide monocular ratio data to gauge their perceived depth of a cylindrical object, albeit presented differently across the two settings. In the first experiment, the cylinder was depicted on a computer screen,

offering a digital representation of depth. Conversely, in the second experiment, the cylinder was presented as a hologram, aiming to offer a more immersive depth perception experience. This contrast raises an intriguing question: does the method of presentation elicit significant differences in observers' perceptual adjustments of depth?

To address this query, a paired samples t-test was employed, comparing the mean adjusted ratios from observers in Experiment 1 (computer screen presentation) with those from Experiment 2 (holographic presentation). The statistical analysis indicated no significant difference between the mean hologram ratios from Experiment 1 ($M = 0.86$, $SD = 0.03$) and the mean computer-adjusted ratios from Experiment 2 ($M = 0.86$, $SD = 0.03$), yielding a $t(6) = -0.22$ and a p-value of 0.83. This outcome suggests that the observers' adjustments in perceived depth were statistically similar across the two distinct presentation modalities. Moreover, the computed effect size ($\eta_p^2 = 0.008$) reveals a very small effect. Consequently, given the statistical equivalence in adjustment ratios between the computer screen and holographic presentations, we combined the adjustment ratios from both presentation types for all subsequent analyses.

After investigating the impact of our independent variables on perceptual measurements of perceived cylinder length, we now examine whether these same factors similarly affect the motor-based measurement. Clearly, inter-center distance has a significant influence on perceived cylinder depth while rotation speed and circle size do not. To determine how inter-center distance affected observers' reaching patterns, a 2 (Inter-Center Distance: 1 cm vs. 2.5 cm) $\times$ 2 (Target Location: front vs. back) repeated measures ANOVA was conducted to assess the influence of inter-center distance and target location on observers' reaching distances. The ANOVA revealed a significant main effect of inter-center distance, $F(1, 6) = 76.83, p < 0.001$ with observers reaching farther for the back circle ($M = 34.5$ cm, $SE = 0.87$) compared to the front ($M = 29.54$ cm, $SE = 0.40$) as the inter-center distance increased. The main effect of target location was also significant, $F(1, 6) = 75.29, p < 0.001$ indicating that reaching distances were influenced by whether observers were reaching towards the

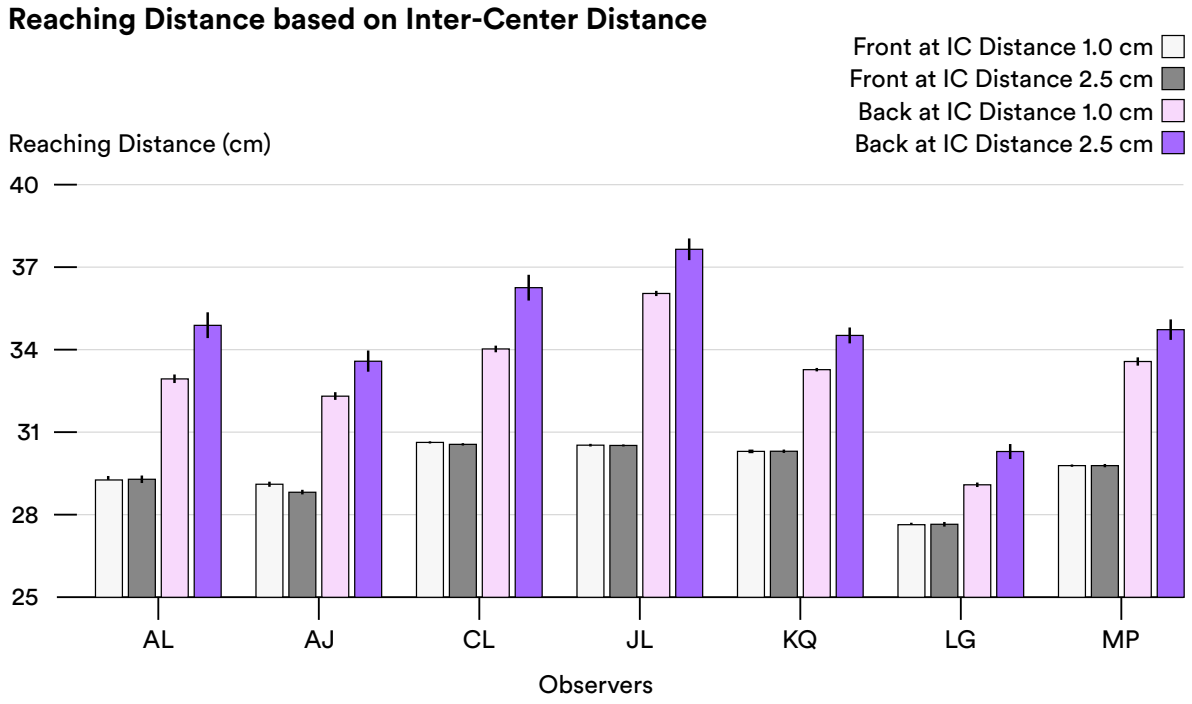

Figure 2.13: **Mean reaching distances across observers based on inter-center distance.** This figure presents the aggregated results of reach measurements for each observer, showcasing front and back reach distances at both inter-center distances. Bars represent the mean distances reached, with error bars indicating standard errors. Significant differences between inter-center distances are highlighted with asterisks.

front or the back of the cylinder. There was a significant interaction between inter-center distance and target location, $F(1, 6) = 95.79, p < 0.001$. Post-hoc paired samples t-tests were conducted to determine the cause of this interaction, revealing that the difference in reaching distances between the front and back of the cylinder was greater at the larger inter-center distance (Pair 2 [Front Reach (ICD2) vs. Back Reach (ICD2)]: $M_{diff} = -4.98$ cm, $SE_{diff} = 0.52$, $t(6) = -9.50$, $p < 0.001$) compared to the smaller inter-center distance (Pair 1 [Front Reach (ICD1) vs. Back Reach (ICD1)]: $M_{diff} = -3.41$ cm, $SE_{diff} = 0.45$, $t(6) = -7.52$, $p < 0.001$). Each observers' mean reaching distance to the front circle and back circle at both inter-center distances are displayed in Figure 2.13.

The post-hoc tests also indicated no significant difference in the reaching distances for the front of the cylinder between the two inter-center distances (Pair 3 [Front Reach (ICD1) vs. Front Reach (ICD2)]: $M_{diff} = 0.05$ cm, $SE_{diff} = 0.04$, $t(6) = 1.10$, $p = 0.31$), suggesting that the perceived location of the cylinder's front remained consistent irrespective of the inter-center distance. However, a significant difference was found for the back reaching distances between the two inter-center distances (Pair 4 [Back Reach (ICD1) vs. Back Reach (ICD2)]: $M_{diff} = -1.52$ cm, $SE_{diff} = 0.16$, $t(6) = -9.59$, $p < 0.001$), confirming that observers perceived the back of the cylinder to be farther away as inter-center distance increased. The findings were further corroborated by significant differences when comparing the reaching distances for the front at one inter-center distance with the back at the other (Pair 5 [Front Reach (ICD1) vs. Back Reach (ICD2)]: $M_{diff} = -4.93$ cm, $SE_{diff} = 0.52$, $t(6) = -9.40$, $p < 0.001$; Pair 6 [Front Reach (ICD2) vs. Back Reach (ICD1)]: $M_{diff} = -3.46$ cm, $SE_{diff} = 0.45$, $t(6) = -7.61$, $p < 0.001$). These results suggest that observers reliably perceive a longer cylinder with an increased inter-center distance, with the lengthening effect primarily observed at the cylinder's back. These findings not only validate the hypothesis that inter-center distance is a determinant of depth perception but also reveal the asymmetrical nature of this effect on the perceived positions of the front and back of the cylinder.

2.6.1 Cross-Modal Correlational Analyses of Perceived Cylinder Length

With inter-center distance established as a primary determinant of perceived cylinder depth, our subsequent analyses pivot towards assessing the consistency of cylinder length perception across various sensory modalities. This phase of the study examines the relationships between perceptions informed by monocular and binocular cues, adjustments in sphere diameter, and motor-based reaching measurements. Such cross-modal correlation analyses are crucial for understanding the degree of coherence in depth perception — do observers' perceptions in one modality align with their perceptions in another? Our goal is to assess whether the different sensory modalities result in a congruent representation of the stereokinetic cylinder.

To this end, we conducted a series of correlational analyses between the different recorded metrics of perceived depth.

2.6.1.1 Perceptual Correlations

We begin by examining the consistency of perceived cylinder length through correlation analyses of the three perceptual measurements: monocular adjusted ratios, absolute binocular disparity differences, and sphere diameter. Binocular disparity differences emerge from the observer's adjustments to the depth probe, aimed at aligning with the cylinder's front and 'back.' Specifically, adjustments matching the cylinder's front result in a crossed disparity, coded positively, whereas adjustments for the cylinder's back lead to an uncrossed disparity, coded negatively. The absolute binocular disparity difference thus captures the total magnitude of depth adjustment, offering a direct measure of perceived depth without the need for transformation into standard length units. This approach allows us to analyze the raw values directly, maintaining the integrity of the observers' depth perception without the potential bias introduced by transforming these measures into cylinder lengths in centimeters.

It's noteworthy that while we refrained from converting monocular adjusted ratios and binocular disparity differences into centimeters to preserve their raw values, sphere diameters were presented in centimeters (cm). This conversion was straightforward, relying on translating the display's pixel density into spatial dimensions, providing a direct measure of perceived size. Specifically, we investigate whether observers who perceive the cylinder as longer in one task tend to do so across other tasks as well, examining the coherence of depth cues and their cumulative effect on the perception of the stereokinetic cylinder's dimensions. All three pair-wise correlations of the perceptual measurements are depicted in Figure 2.14.

Absolute Disparity Differences vs. Monocular Adjusted Ratios: In this analysis, we correlated observers' mean absolute binocular disparity differences with their mean monocular adjusted ratios. A significant negative correlation (r = -0.78, p = 0.04) suggests

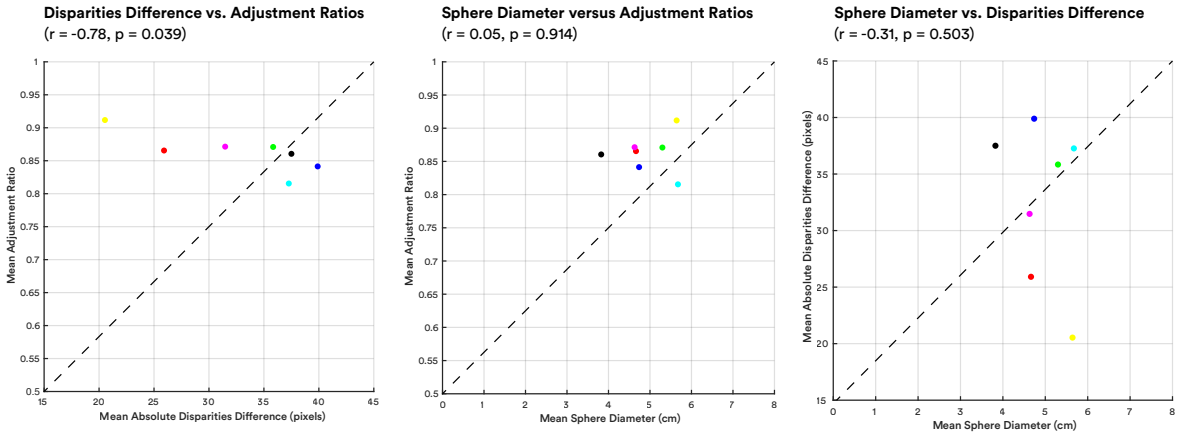

Figure 2.14: **Correlations between perceptual measurements of perceived cylinder length.** Subplot 1 presents the relationship between mean absolute binocular disparity differences and mean adjustment ratios, which are significantly correlated. Subplot 2 compares mean sphere diameters with mean adjustment ratios. Subplot 3 shows the correlation between mean sphere diameters and mean absolute binocular disparity differences.

that as absolute binocular disparity differences increase, indicating greater perceived depth, observers tend to adjust the back circle to be smaller, reflecting a longer perceived cylinder. The regression equation, y = -0.003x + 0.97, with a slope significant at $p = 0.04$, further supports this inverse relationship, indicating a consistent decrease in adjusted ratios with increased disparity differences. As the adjusted ratio decreases, it signifies an increase in the perceived cylinder length.

Sphere Diameter vs. Monocular Adjusted Ratios: The negligible correlation ($r = 0.05$, $p = 0.91$) implies no predictive relationship between the adjustments made to the sphere diameter and the monocular adjusted ratios. This suggests that these two perceptual tasks tap into different aspects of depth perception or that observers are leveraging different pieces of information when performing these particular tasks.

Monocular Sphere Diameter vs. Absolute Disparities Difference The correlation

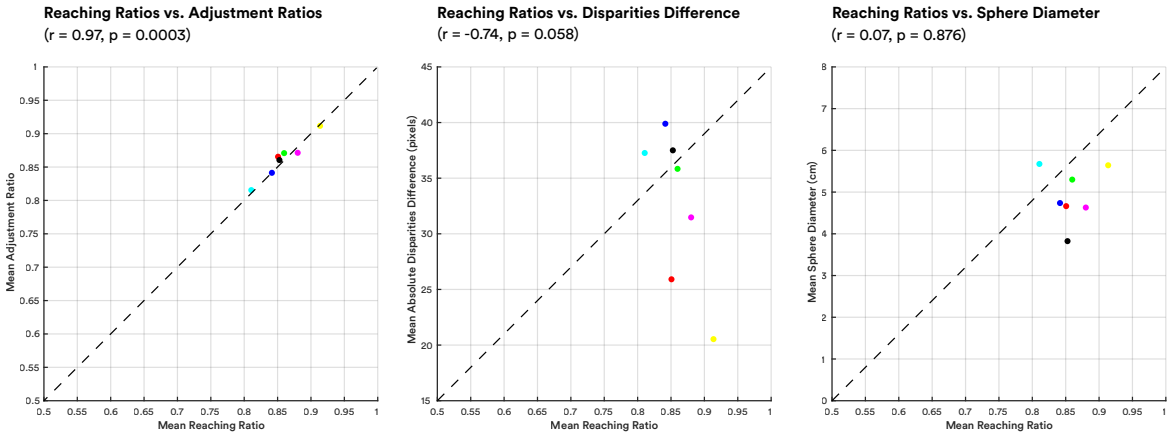

Figure 2.15: **Correlations between reaching ratios and perceptual measurements.** Subplot 1 demonstrates the relationship between mean reaching ratios and mean adjustment ratios, which are significantly correlated. Subplot 2 explores the link between mean reaching ratios and mean absolute binocular disparity differences. Subplot 3 connects mean reaching ratios with mean sphere diameters.

(r = -0.31, p = 0.50) does not suggest a strong relationship between sphere diameter adjustments and binocular disparity differences. This weak negative correlation implies that changes in perceived depth via binocular disparities do not consistently influence the perceived size of the sphere in monocular viewing conditions.

2.6.1.2 Reaching Correlations

In the following analysis, we examine how perceptual measures of cylinder length relate to motor reaching. By correlating mean reaching ratios with perceptual measurements, we aim to discern the alignment between visual perception and physical interaction with the cylinder. We decided to use reaching ratios to account for individual differences in arm length and positioning, ensuring that the data reflects proportional depth perception rather than absolute reach. Correlation analyses between the reaching ratios and the three perceptual measurements of perceived cylinder length are displayed in Figure 2.15.

Reaching Ratios vs. Monocular Adjusted Ratios: A very strong positive correlation ($r = 0.97$, $p < 0.001$) indicates a close alignment between perceptual adjustments in cylinder length and reaching actions. This suggests a high degree of coherence between visual and motor representations of cylinder size, as supported by the regression analysis ($y = 0.89x + 0.10$, slope $p < 0.001$), indicating that reaching adjustments are predictably related to visual adjustments. In other words, the significant correlation suggests a high degree of integration between visual perception and motor action, indicating that the adjustments based on visual perception are closely aligned with how individuals physically interact with space.

Reaching Ratios vs. Absolute Binocular Disparities Differences: The negative correlation (r = -0.74, $p = 0.058$) suggests a trend where increased depth perception (via larger binocular disparity differences) might lead to shorter reaching actions, although this relationship is not statistically significant (slope $p = 0.058$).

Reaching Ratios vs. Sphere Diameter: The very weak correlation ($r = 0.07$, $p = 0.88$) suggests that the adjustments to sphere diameter do not meaningfully predict reaching behavior. The trends observed in the relationships between reaching ratios and absolute binocular disparities differences, as well as the weak correlation with sphere diameter, provide further evidence of the nuanced nature of depth perception. The absence of a strong correlation in these instances could be indicative of the varied reliance on different depth cues depending on the task or the individual differences in cue utilization.

2.6.1.3 Implications

The strong positive correlation between mean reaching ratios and mean monocular adjusted ratios ($r = 0.97$, $p < 0.001$) specifically underscores a coherence in depth perception through monocular vision. This suggests that the way observers adjust for cylinder length based on monocular cues closely aligns with their motor actions during reaching tasks, indicating a precise translation of visual information from monocular perception into motor responses.

The regression equation ($y = 0.89x + 0.10$, slope $p < 0.001$) further demonstrates this relationship by quantifying the predictability of reaching behavior based on visual adjustments. The slope near 1 suggests that for every unit of change in perceptual adjustment, there is a nearly equivalent change in motor action, reinforcing the idea of a tightly coupled perception-action mechanism.

Furthermore, the significant correlation between mean monocular adjusted ratios and mean absolute binocular disparity differences, coupled with the approaching significance between binocular disparities and reaching ratios ($r = -0.74$, $p = 0.058$), highlights a nuanced relationship. While monocular cues directly influence motor action, binocular cues also play a role, albeit in a more complex manner that nearly reaches statistical significance.

The minimal correlation involving sphere diameter adjustments indicates that this measure may engage distinct cognitive processes or perceptual cues from those utilized in monocular adjustments, binocular adjustments, or motor actions. This discrepancy underscores the task-specific nature of depth perception and suggests that different depth cues may be prioritized or processed differently depending on the context or task demands.

2.6.2 Standardization and Conversion of Depth Measurements to Centimeters

To enable a coherent comparison of perceived cylinder lengths across various measurement techniques, standardizing all measurements into a common unit — centimeters — is essential. Monocular adjustment ratios and binocular disparity measurements offer insights into the relative depth perception, revealing the spatial relationship between different parts of the cylinder from the observer's perspective. However, for effective comparison across tasks, these measurements need to be translated into absolute spatial terms. This translation involves converting the perceived depth, initially captured as ratios and disparities, into actual lengths in centimeters. This process requires factoring in the observer's perceived viewing distance, derived from their calibration data. Incorporating the viewing distance into our calculations ensures that the converted measurements accurately mirror the observers'

perceptions. The ultimate goal of this standardization process is to determine if variations in the perceived cylinder lengths among observers indicate individual preferences for the 'slow' motion and rigidity principles, reflecting a differential weighting of these perceptual preferences.

2.6.2.1 Calculating Cylinder Length from Monocular Linear Perspective

To quantify perceived cylinder depth from monocular linear perspective, observers were presented with various configurations of the two overlapping homogeneous circles. In each trial, one circle is always of a fixed radius and the radius of the other circle, which participants manipulate, is rendered by adding or subtracting a predetermined value from the radius of the fixed circle. Observers were asked to adjust the non-fixed circle until a uniform cylinder is perceived, and we recorded in each trial the ratio of back circle radius over the front circle radius. This concept is illustrated in Figure 2.16, where we assume that when perceiving the cylinder, observers perceive the larger circle at the objective viewing distance to the image plane (D) while the smaller circle is perceived at some further distance from the image plane (D + L). Thus, we can use the following formula to calculate L:

$$L = \sqrt{d^2 + \left(\frac{1}{r} - 1\right)^2 \cdot D^2} \tag{2.1}$$

Where L is the length of the cylinder taking into account its tilt, d is the inter-center distance between the two circles, r is the ratio of the two circle radii, and D is the viewing distance from the observer to the image plane. The ratio is defined as the larger circle radius (r_2) divided by the smaller circle radius (r_1) such that r is always less than 1 (i.e., $r = \frac{r_2}{r_1}$). Using this formula, we are able to convert the adjustment ratios provided by the observers into a specific length in centimeters.

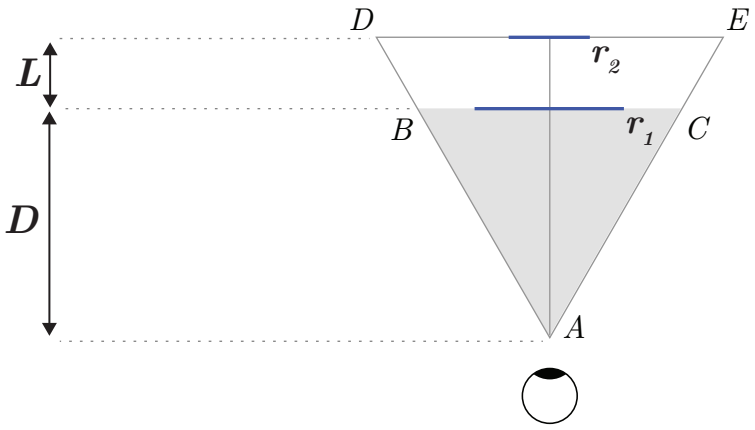

Figure 2.16: **Simplified aerial view of the hypothesized perceived depth of the stereokinetic cylinder based on linear perspective.** The two circles are displayed on an image plane at a viewing distance of 30 cm from the observer, rotating around the z-axis perpendicular to the image plane. After extended viewing, observers constantly report perceiving depth between the two circles such that one appears to be farther away. D is the viewing distance from the observer to the image plane. L is the length of the perceived cylinder. When modeling the perceived depth of the cylinder, we assume that the larger circle (r_1) is perceived to be located directly at the image plane, 30 cm from the observer, while the smaller circle (r_2) is perceived to be at a distance of $D+L$. Notice that $\triangle ABC \sim \triangle ADE$, a relationship we leverage to derive an equation to calculate L, the perceived cylinder length.

2.6.2.2 Calculating Perceived Distance based on Crossed Disparity

Figure 2.17 shows an observer's adjustment of crossed disparity between the red and blue sets of concentric diamonds to match their perceived distance of the front of the cylinder. The observer utilizes red-blue anaglyph glasses with the left lens displaying the red set of concentric diamonds and the right lens displaying the blue set. As a result of this crossed disparity, the depth probe appears to be in front of the computer monitor screen.

Figure 2.18 provides a simplified aerial view illustrating the variables required for cal-

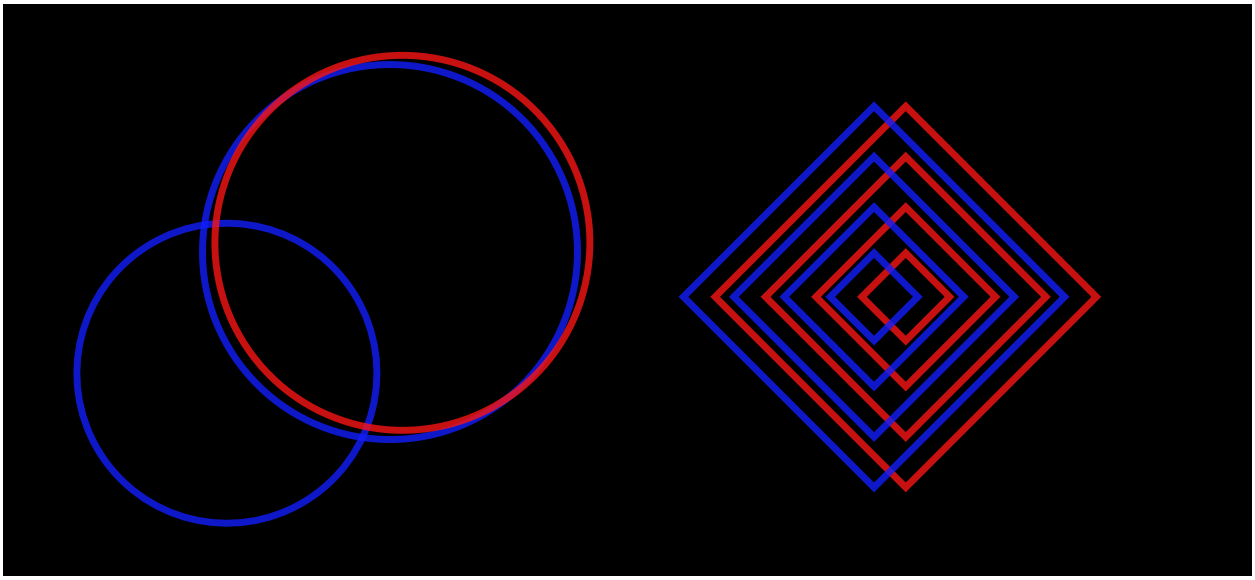

Figure 2.17: **Crossed disparity between concentric diamond sets.** An image showcasing an observer's adjustment of the horizontal disparity between the two sets of diamonds in order to match the perceived distance of the front of the cylinder.

culating the perceived distance of the depth probe from the computer screen. Here, D represents the observer's viewing distance to the image plane, d_i denotes the interpupillary distance, and d signifies the horizontal disparity between the two sets of concentric diamonds. h represents the distance between the depth probe and the computer screen.

In Figure 2.18 the observer is looking directly towards the monitor screen (indicated by the top black line) such that the line segment connecting the observer's eyes is parallel to the monitor. The blue dot and red dot at the top of the figure represent the blue and red sets of concentric diamonds with corresponding lines of sight to each set. The point of intersection between these lines indicates the 3D spatial location of the depth probe, which in this case, is positioned in front of the computer monitor. To determine h, we utilize the similarity of triangles formed by the line connecting the observer's eyes and the horizontal disparity between the two sets of concentric diamonds ($\triangle ABC \sim \triangle CDE$.). As a result of this similarity, we can calculate h using the following formula where $h > 0$:

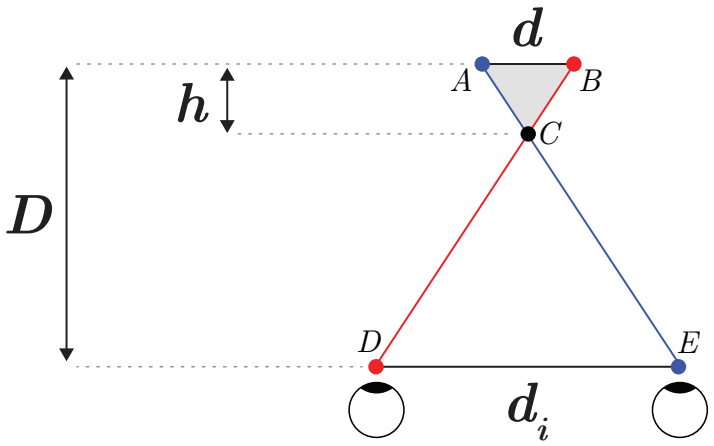

Figure 2.18: **Simplified aerial view of depth probe location when diamond sets are crossed.** D is the viewing distance from the observer to the image plane. d_i is the interpupillary distance. d is the horizontal disparity of the two sets of concentric diamonds. h is the distance between the depth probe and the computer screen. Like our model for calculating cylinder height from linear perspective, we take advantage of $\triangle ABC \sim \triangle CDE$.

$$h = \left(\frac{d}{d + d_i}\right) \cdot D \tag{2.2}$$

h is the perceived distance of the depth probe from the computer screen, d is the inter-center distance between the two circles, r is the ratio of the two circle radii, d_i is the inter-pupillary distance specific to the observer, and D is the viewing distance from the observer to the image plane. This formula allows us to translate the observed horizontal disparity, d, into the perceived distance of the depth probe from the screen, h. After calculating h, we simply subtract it from the viewing distance of 30 cm, D, to determine the distance of the probe from the observer.

2.6.2.3 Calculating Perceived Distance based on Uncrossed Disparity

Please recall in Figure 2.4, we observed an adjustment made by the observer to match the perceived distance of the rear end of the cylinder. This adjustment results in uncrossed disparity between the two sets of concentric diamonds. Calculating perceived distance from uncrossed disparity parallels the process described for crossed disparity. Figure 3.7 illustrates an aerial view of the red and blue sets of concentric diamonds in the uncrossed configuration.

In Figure 2.19, the depth probe is now perceived behind the image plane. Similar to the crossed disparity case, we utilize the geometry of two similar triangles $\triangle ABC \sim \triangle ADE$ to calculate h as follows:

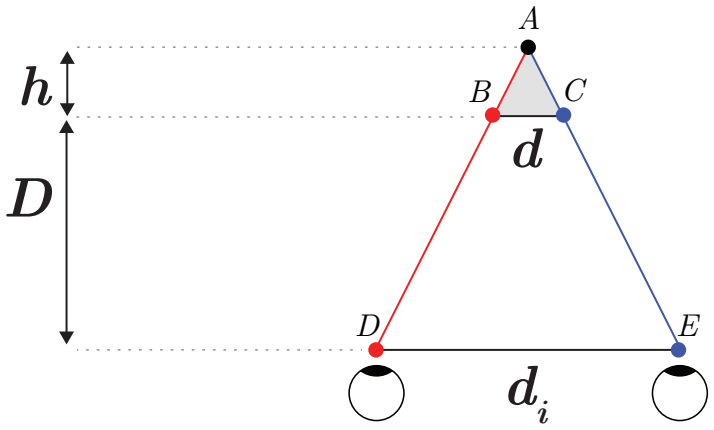

Figure 2.19: **Simplified aerial view of depth probe location when diamond sets are uncrossed.** D is the viewing distance from the observer to the image plane. d_i is the interpupillary distance. d is the horizontal disparity of the two sets of concentric diamonds. h is the distance between the depth probe and the computer screen. As in our previous models, we utilize $\triangle ABC \sim \triangle ADE$ for calculating the perceived 3D location of the depth probe when disparity is uncrossed.

$$h = \left(\frac{d}{d - d_i} \right) \cdot D \tag{2.3}$$

Again, h is the perceived distance of the depth probe from the computer screen, d is the inter-center distance between the two circles, r is the ratio of the two circle radii, d_i is the interpupillary distance specific to the observer, and D is the viewing distance from the observer to the image plane. This formula allows us to translate the observed horizontal disparity, d, into the perceived distance of the depth probe from the screen, h. After calculating h, we simply add it to the viewing distance of 30 cm, D, to determine the distance of the probe from the image plane.

2.6.3 Comparison of Converted Cylinder Measurements

All converted measurements of perceived cylinder depth at an inter-center distance of 1 cm are displayed in Figure 2.20 while Figure 2.21 shows the data at an inter-center distance of 2.5 cm. These measurements, encompassing reaching length, monocular length, binocular length, and sphere diameter of the cylinder, have been standardized to facilitate direct comparisons across different perceptual modalities and motor responses. To investigate the impact of inter-center distance on perceived cylinder length and to assess individual and collective perceptual consistencies, we conducted four separate 2 x 7 mixed effects ANOVAs. Each ANOVA is tailored to one of the converted measurements. This analytical approach is designed to address two critical questions: Do observers perceive significantly different cylinders in terms of length? And, does the inter-center distance uniformly affect each observer's perception, or are there notable variations? By treating inter-center distance as a within-subject factor and the observer as a between-subject factor, we can isolate the effect of spatial configuration on perceived depth while accounting for individual differences in perception. This ensures a analysis of how each converted measurement responds to variations in inter-center distance, providing insights into the consistency of depth perception

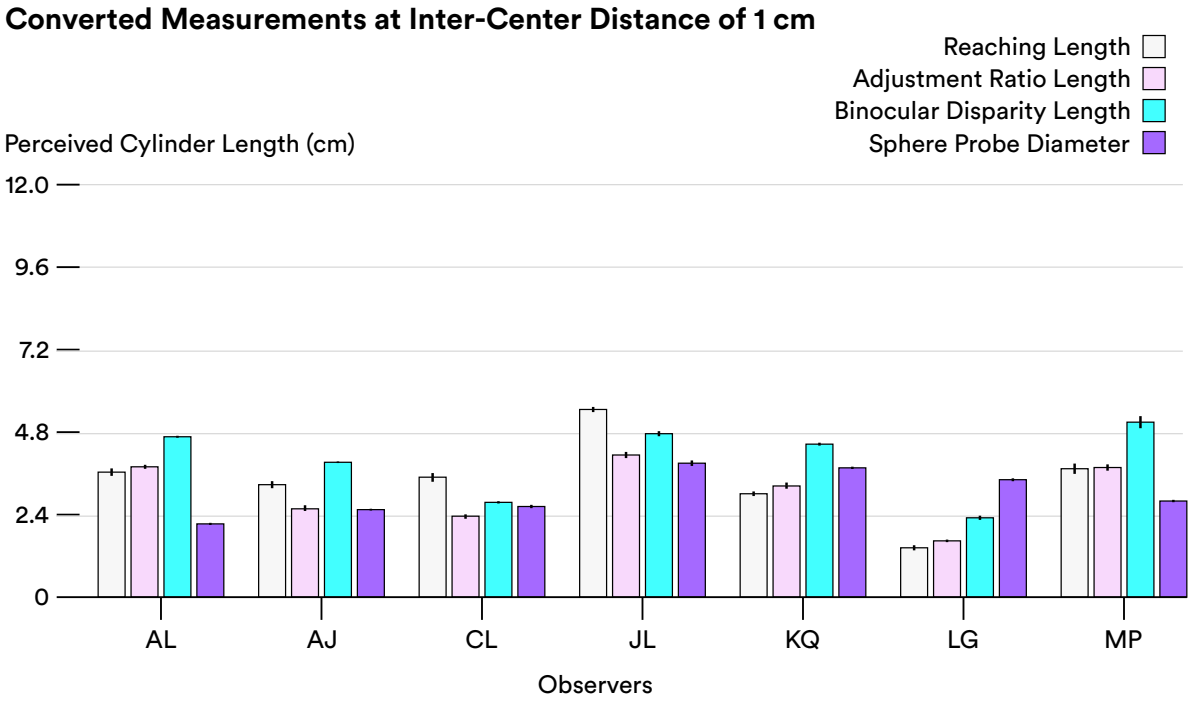

Figure 2.20: **Perceived cylinder length based on four measurement methods at an inter-center distance of 1 cm.** Each bar represents the perceived cylinder length as determined by one of four methods — reaching length, adjustment ratio length, binocular disparity length, and sphere probe diameter — for each observer in the study.

across observers. Rotation speed and circle size were not included as independent variables because previous analyses have suggested that neither of these two variables significantly influence perceived cylinder length.

First, we examined whether observers significantly differed in perceived cylinder length based on monocular linear perspective. A 2 x 7 mixed-effects ANOVA was conducted to compare the effect of inter-center distance on monocular adjusted cylinder length, with observer as a between-subjects factor. There was a significant effect of inter-center distance on monocular adjusted cylinder length, $F(1, 553) = 6,944.04$, $p < .001$, indicating a large effect size and suggesting that the cylinder's perceived length varied significantly with changes in inter-center distance. Mean monocular perceived cylinder length increased from 3.08 cm

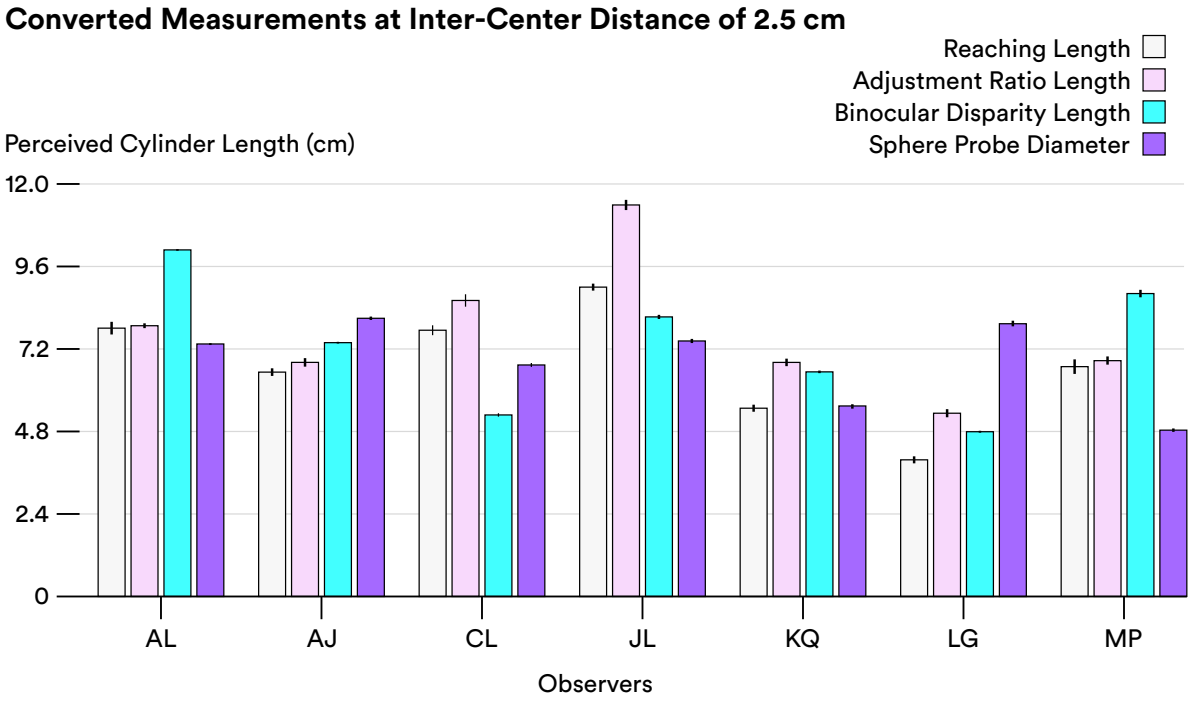

Figure 2.21: **Perceived cylinder length based on four measurement methods at an inter-center distance of 2.5 cm.** Each bar represents the perceived cylinder length as determined by one of four methods — reaching length, adjustment ratio length, binocular disparity length, and sphere probe diameter — for each observer in the study.

($SE = 0.03$) to 7.57 cm ($SE = 0.03$). There was also a significant interaction effect between inter-center distance and observer, $F(6, 553) = 111.81$, $p < .001$, indicating that the effect of inter-center distance on perceived cylinder length differed among observers. In other words, all observers perceived a longer cylinder as inter-center distance increased, but the amount of increase varied across observers. The between-subjects effects revealed a significant effect of observer on monocular adjusted cylinder length, $F(6, 553) = 277.04$, $p < .001$, suggesting substantial individual differences in the perception of cylinder length.

In examining the individual observer means for monocular perceived cylinder length, we observe a range of reported values, reflecting the diversity in depth perception across

observers. AJ perceived a moderate cylinder length with a mean of 4.70 cm, while AL's mean of 5.81 cm indicates a tendency to perceive a longer cylinder. CL also reported a longer cylinder, with a mean of 5.37 cm whereas LG perceived a notably shorter length at 3.43 cm, the shortest among the group. JL's perception was the most distinct, perceiving the cylinder to be significantly longer with a mean of 7.70 cm. KQ's mean estimate of 4.98 cm and MP's mean of 5.32 cm reflect additional variability within this perceptual domain.

Following the monocular analysis, binocular adjusted cylinder length was evaluated using a 2 x 7 mixed-effects ANOVA with inter-center distance as the within-subjects factor and observer as the between-subjects factor. There was a significant main effect of inter-center distance on binocular adjusted cylinder length, $F(1, 413) = 7,911.70$, $p < .001$, indicating a very large effect size. This demonstrates that perceived cylinder length as determined through binocular cues was significantly influenced by the inter-center distance. Mean binocular perceived cylinder length increased from 4.01 cm ($SE = 0.03$) to 7.28 cm ($SE = 0.03$), indicating that on average, observers perceived the cylinder to be longer binocularly in comparison to monocular viewing. The interaction between inter-center distance and observer was also significant, $F(6, 413) = 128.0$, $p < .001$. This significant interaction suggests that the change in perceived cylinder length with inter-center distance varied across observers. Between-subjects effects showed a significant effect of observer, $F(6, 413) = 942.98$, $p < .001$ indicating substantial individual variability in binocular depth perception across the cylinder lengths.

AJ's mean estimation was 5.66 cm indicating a moderate perception of cylinder length. AL reported a significantly longer perceived length at 7.37 cm, suggesting a more pronounced depth perception. CL perceived a notably shorter cylinder length with a mean of 4.03 cm, the shortest within the group, which contrasts sharply with MP's longer mean estimation of 6.96 cm. JL and KQ reported mean lengths of 6.46 cm and 5.48 cm, respectively, reflecting additional variance. LG's mean of 3.56 cm indicates a preference for a shorter perceived length similar to CL's.

Moving on, the sphere diameter is an intriguing measure because it relies on a different set of visual cues compared to the previous linear and binocular perspectives. Observers are tasked with adjusting a two-dimensional representation to match their perception of the cylinder's three-dimensional depth. Previous correlation analyses suggests this measure potentially taps into additional cognitive processes, which may or may not align with the previously examined monocular and binocular depth cues.

The sphere diameter was evaluated for its response to inter-center distance changes and individual observer variability through a 2 x 7 mixed-effects ANOVA similar to the previous measures. There was a significant main effect of inter-center distance on sphere diameter, $F(1, 413) = 20,780.28$, $p < .001$. This significant result shows that as inter-center distance increased, the mean sphere diameter adjusted by observers also increased from 3.05 cm ($SE = 0.02$) to 6.80 cm ($SE = 0.02$). Moreover, the interaction effect between inter-center distance and observer was significant, $F(6, 413) = 451.18$, $p < .001$. This indicates a substantial degree of individual difference in how changes in inter-center distance influenced sphere diameter adjustments. Each observer's unique perception of depth contributed to the variability in how the sphere diameter was perceived and adjusted. Between-subjects effects revealed a significant effect of observer, $F(6, 413) = 310.27$, $p < .001$, suggesting that individual differences are a strong factor in the perception of cylinder length as estimated by sphere diameter.

Observer AJ perceived the cylinder to have an average diameter of 5.30 cm. Observer AL's mean estimate was 4.74 cm and a confidence interval from 4.66 to 4.81 cm, suggesting a preference for a shorter cylinder length than AJ. Observer CL's perception yielded a mean sphere diameter of 4.66 cm, while JL reported a significantly larger perceived diameter at 5.67 cm, the highest among the group. Notably, observer MP perceived the shortest cylinder, with a mean diameter of 3.82 cm, suggesting a distinct variance in depth processing from the other observers.

As we have observed significant variability in how individuals perceive depth through both

monocular and binocular cues, we now turn to the reaching measurements. A significant effect of inter-center distance on reaching lengths was found, $F(1, 133) = 2,541.25, p < .001$. This finding indicates that observers' motor responses in reaching towards the perceived cylinder ends significantly increased with the inter-center distance, echoing the trends observed in the monocular and binocular lengths. Mean reaching length increased from 3.41 cm ($SE = 0.04$) to 6.63 cm ($SE = 0.05$). Additionally, the interaction effect of inter-center distance and observer was significant, $F(6, 133) = 17.10, p < .001$, suggesting individual differences in how observers adjusted their reaching based on inter-center distance. Taken together, these results suggest not only a general trend of increased reaching lengths with greater inter-center distances but also that the magnitude of this increase varied from one individual to another. Between-subjects effects also revealed a significant effect of observer, $F(6, 133) = 228.12, p < .001$, underscoring substantial individual variability in reaching measurements. These findings complement the perceptual data, suggesting that motor-based measures of perceived cylinder length are in line with visual assessments.

The reaching data for each observer demonstrates distinct individual differences in perceived cylinder length. AJ's mean reaching length was 4.80 cm, suggesting a moderate perception of cylinder depth. AL reported a longer reach with a mean of 5.64 cm, while CL was close behind with a mean of 5.49 cm. JL had the longest reach of all, with a mean of 7.19 cm. In contrast, KQ and LG reported much shorter reaching lengths, with means of 4.17 cm and 2.66 cm, respectively, indicating they perceived the cylinder to be shorter than their peers did. MP had a mean reaching length of 5.22 cm, fitting within the middle range of the group. The consistency within each observer's reaching length, as indicated by the small standard error, shows that while individual perceptions of depth vary, each observer's perception is reliable and replicable.

In summary, the four different perceptual modalities yielded varied estimates of the perceived cylinder length. Binocular adjusted lengths tended to result in the longest perceived cylinders, with monocular adjustments and sphere diameters producing moderately long es-

timates, while motor reaching measurements generally reflected shorter perceived lengths. Specifically, the binocular lengths ranged up to 7.28 cm on average, while reaching lengths were more conservative, with the longest mean reaching length being 7.19 cm. Interestingly, the sphere diameter adjustments by observers suggested the most significant individual variability, as indicated by the wide range of means from 3.82 cm to 5.67 cm. This variability in estimates across modalities indicates that each sensory approach leads to distinct perceptual outcomes. Observers' measurements of the cylinder's length are modality-specific, highlighting individual variations in sensory processing.

2.7 Discussion

In the current study, we investigated the perceptual representation of the stereokinetic cylinder by measuring its perceived length across both visual and motor tasks. Inter-center distance emerged as the sole factor influencing perceived cylinder length, echoing findings from previous research that link perceived depth to the size of the display (Proffit, Rock, Hecht, & Schubert, 1992). Notably, as indicated by all measurement types, as inter-center distance increased, so did the perceived length of the cylinder. However, the amount of increased depth perceived by each observer, in every measurement, differed. This raises an intriguing question: does the perceived elongation of the cylinder result from the front appearing closer, or the back appearing farther away? Within our experimental framework, the reaching tasks suggest that while the perceived location of the front circle remains constant, the back circle is perceived as increasingly distant with greater inter-center distances. This interpretation is supported by data showing consistent reaching locations for the front circle across both inter-center distances, but significant variation in reaching towards the back circle as inter-center distance changes.

We found significant inter-observer variability in the perception of cylinder length when presented with the same stimulus and task (Koenderink, 2001). However, despite this vari-

ability, our correlational analyses indicated a substantial degree of consistency in how the cylinder was perceived across our sample. The results illustrate that observers exhibit stable, albeit unique, internal representations of the cylinder's length, with monocular adjustment ratios and binocular disparities difference aligning closely with the reaching measurements. This study's methodology, marked by consistency in observers' responses as indicated by small error bars, reflects the reliability of the measurement tools employed. It should be noted that the consistency is qualitative in nature suggesting a general trend rather than a precise numerical match across measurements.

Interestingly, there was a significant correlation between the monocular cylinder length and the motor reaching measurement of the cylinder, whereas no correlation was observed for the binocular length. This discrepancy may arise because, despite the LED being visible binocularly, the illusion itself is perceived monocularly. As a result, the brain appears to prioritize monocular cues about the cylinder over binocular cues when guiding observers' fingers towards the hologram. Contrary to the two-stream hypothesis (Goodale & Milner, 1992), which suggests that the dorsal stream should not be deceived by the stereokinetic illusion, our findings indicate otherwise. Observers' reaching movements consistently reflected the perceived depth of the illusory cylinder, suggesting an influence of perceptual representation on action, blurring the lines between vision-for-perception and vision-for-action.

In conclusion, our findings spotlight individual preferences for slowness versus rigidity in perception. Contrary to the expectation of perceiving an infinitely long cylinder based on rigidity assumptions (Jansson & Johansson, 1973; Ullman, 1974, 1984), our measurements uniformly indicated a perception of finite cylinder length. This suggests a compromise between preferences for rigidity and slow motion (Hildreth, 1984; Yuille & Grzywacz, 1989; Weiss, Simoncelli, & Adelson, 2002), with significant variations among observers in how these preferences are weighted. The observations accrued from this study establish a foundation for future research to further investigate the interplay between these perceptual preferences and their variability across individuals.

REFERENCES

Aaen-Stockdale, C. R., Farivar, R., & Hess, R. F. (2010). Co-operative interactions between first-and second-order mechanisms in the processing of structure from motion. *Journal of Vision*, 10(13), 1-9.

van Andel, S., Cole, M. H., & Pepping, G. J. (2018). Regulation of locomotor pointing across the lifespan: Investigating age-related influences on perceptual-motor coupling. *PLoS One*, 13(7), e0200244.

Braunstein, M. L. (2014). *Depth perception through motion*. Academic Press.

Censor, N., Sagi, D., & Cohen, L. G. (2012). Common mechanisms of human perceptual and motor learning. *Nature Reviews Neuroscience*, 13(9), 658-664.

Christensen, A., Giese, M. A., Sultan, F., Mueller, O. M., Goericke, S. L., Ilg, W., & Timmann, D. (2014). An intact action-perception coupling depends on the integrity of the cerebellum. *Journal of Neuroscience*, 34(19), 6707-6716.

Clocksin, W. F. (1980). Perception of surface slant and edge labels from optical flow: a computational approach. *Perception*, 9(3), 253-269.

Dosher, B.A., Landy, M. S., & Sperling, G. (1989). Kinetic depth effect and optic flow - I. 3D shape from fourier motion. *Vision Research*, 29(12), 1789-1813.

Fernandez, J. M., & Farell, B. (2009). A new theory of structure-from-motion perception. *Journal of vision*, 9(11), 23-23.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.

Green Jr, B. F. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 62(3), 272.

Grzywacz, N. M., & Hildreth, E. C. (1987). Incremental rigidity scheme for recovering

structure from motion: Position-based versus velocity-based formulations. *JOSA A*, 4(3), 503-518.

Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision Research*, 36(21), 3441-3456.

Haffenden, A. M., & Goodale, M. A. (2000). Independent effects of pictorial displays on perception and action. *Vision Research*, 40(10-12), 1597-1607.

Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial intelligence*, 23(3), 309-354.

Hildreth, E. C., Ando, H., Andersen, R. A., & Treues, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research*, 35(1), 117-137.

Husain, M., Treue, S., & Andersen, R. A. (1989). Surface interpolation in three-dimensional structure-from-motion perception. *Neural Computation*, 1(3), 324-333.

Jansson, G. & Johansson, G. (1973). Visual perception of bending motion. *Perception*, 2(3), 321-326.

Koenderink, J. J., & van Doorn, A. J. (1975). Optic Flow. *Vision Research*, 26(1), 161-179.

Koenderink, J. J., & van Doorn, A. J. (1976). Local structure of movement parallax of the plane. *JOSA*, 66(7), 717-723.

Koenderink, J. J. (1986). Optic flow. *Vision research*, 26(1), 161-179.

Koenderink, J. J., Van Doorn, A. J., Kappers, A. M., & Todd, J. T. (2001). Ambiguity and the 'mental eye'in pictorial relief. *Perception*, 30(4), 431-448.

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 208(1173), 385-397.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6, 414-417.

Musatti, C. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, 3, 105-120.

Norman, J. F., & Todd, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, 53(3), 279-291.

Proffitt, D. R., Rock, I., Hecht, H., & Schubert, J. (1992). Stereokinetic effect and its relation to the kinetic depth effect. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 3.

Ramachandran, V. S., Cobb, S., & Rogers-Ramachandran, D. (1988). Perception of 3-D structure from motion: the role of velocity gradients and segmentation boundaries. *Perception & Psychophysics*, 44(4), 390-393.

Rokers, B., Yuille, A., & Liu, Z. (2006). The perceived motion of a stereokinetic stimulus. *Vision Research*, 46(15), 2375-2387.

Todd, J. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception and Psychophysics*, 36, 97-103.

Todorović, D. (1993). Analysis of two-and three-dimensional rigid and nonrigid motions in the stereokinetic effect. *JOSA A*, 10(5), 804-826.

Treue, S., Husain, M., & Andersen, R.A. (1991). Human perception of structure from motion. *Vision Research*, 31(1), 59-75.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London B*, 203, 405-426.

Ullman, S. (1984). Rigidity and misperceived motion. *Perception*, 13, 219-220.

Wallach, H., & O'connell, D. N. (1953). The kinetic depth effect. *Journal of experimental psychology*, 45(4), 205.

Wallach, H., Weisz, A., & Adams, P. A. (1956). Circles and derived figures in rotation. *The American Journal of Psychology*, 69(1), 48-59.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598-604.

Xing, Y., & Liu, Z. (2018). A preference for minimal deformation constrains the perceived depth of a stereokinetic stimulus. *Vision Research*, 153, 53-59.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Computation*, 1(3), 334-347.

Zanforlin, M. (1988). The height of a stereokinetic cone: a quantitative determination of a 3-D effect from 2-D moving patterns without a "rigidity assumption". *Psychological Research*, 50, 162-172.

CHAPTER 3

Conclusion

SFM stimuli, essentially 2D in nature, create the impression of 3D shape through motion cues alone. Within this category, stereokinetic stimuli specifically rely on rotational movements within the image plane to evoke a 3D perception. Initially, these stimuli may be perceived as flat, 2D images; however, with prolonged viewing, they transform into striking 3D objects, as first noted by Musatti and Duchamp in the early 1900s. This transition from a 2D to a 3D percept suggests that underlying cognitive assumptions significantly influence how these images are interpreted. Once observers perceive the 3D object, reverting back to interpreting the display as a flat image is challenging. Despite extensive study over the past century, the precise cognitive processes that facilitate perception of stereokinetic stimuli has not been identified.

In the current study, we employed multiple measurement tools to quantify the perceived length of the cylinder. Our analysis involved examining the ratio between the two circles based on linear perspective. The perception of a cylinder inherently implies that the retinal image of the distant circle is smaller than that of the nearer one. By having observers adjust the diameter of one of the circles to perceive a cylinder (rather than a truncated cone), we could infer the cylinder's length indirectly from the ratio of the two diameters. The cylinder length across all observers was calculated using their unique viewing distance that was recorded through a calibration procedure where they reached for a single blue circle. Furthermore, we also calculated cylinder length based on binocular disparity. Observers adjusted the binocular disparity of two sets of concentric diamonds in order to match the

monocularly viewed circles' perceived depth. Similarly to the ratio data, we utilized each observer's unique viewing distance when calculating the cylinder length based on binocular disparity.

Another measure involved a spinning wireframe sphere, positioned centrally between the two circles. The design of the wireframe and its rotation ensured its consistent perception as a 3D object, allowing for a direct measurement of cylinder length via the sphere's diameter adjustments. Last but not least, observers also interacted with a hologram of the cylinder, facilitated by a 45^{o} mirror setup. The arrangement enabled physical reaching movements towards the perceived circular planes, with a dim LED light attached to the observer's finger enhancing visibility when gauging precision. The difference in reaching distances between the two circles provided a direct measure of the cylinder's perceived length.

The data analysis indicated a consistent increase in perceived cylinder length with inter-center distance, unaffected by rotation speed and circle size. While individual perceptions of cylinder length varied significantly, there was a robust consistency observed across different observers. Correlation analyses suggested a coherent and stable internal representation of the cylinder's size within individuals.

In conclusion, our findings echo previous lab work on stereokinetic stimuli (Rokers, Yuille, & Liu, 2006; Xing & Liu, 2018). The perception of both the cone and cylinder are influenced by a compromise between the preferences for maximal rigidity (Jansson & Johansson, 1973; Ullman, 1974, 1984) and slow motion (Hildreth, 1984; Yuille & Grzywacz, 1989; Weiss, Simoncelli, & Adelson, 2002). Contrary to the rigidity assumption's prediction of an infinitely long cylinder and slow motion preference's suggestion of a very short cylinder, our measurements repeatedly revealed a perceived cylinder of finite length. This compromise, coupled with the significant differences observed across individuals, indicates a varied weighting of these two perceptual preferences among observers. The evidence gathered underscores the individual variability in balancing the preferences for rigidity and slow motion, which result in the percept of a finite 3D cylinder within every participant.

REFERENCES

Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial intelligence*, 23(3), 309-354.

Jansson, G. & Johansson, G. (1973). Visual perception of bending motion. *Perception*, 2(3), 321-326.

Musatti, C. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, 3, 105-120.

Rokers, B., Yuille, A., & Liu, Z. (2006). The perceived motion of a stereokinetic stimulus. *Vision Research*, 46(15), 2375-2387.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London B*, 203, 405-426.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598-604.

Xing, Y., & Liu, Z. (2018). A preference for minimal deformation constrains the perceived depth of a stereokinetic stimulus. *Vision Research*, 153, 53-59.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Computation*, 1(3), 334-347.