7-22-2011

Stereoscopic Vision's Impact on Spatial Ability Testing

George Takahashi
Purdue University, gtakahas@purdue.edu

Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By George Takahashi

Entitled
Stereoscopic Vision's Impact on Spatial Ability Testing

For the degree of Master of Science

Is approved by the final examining committee:

Patrick Connolly
Chair

James Mohler

Craig Miller

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Patrick Connolly

Approved by: James Mohler 7/13/2011
Head of the Graduate Program Date

Graduate School Form 20
(Revised 9/10)

PURDUE UNIVERSITY
GRADUATE SCHOOL

Research Integrity and Copyright Disclaimer

Title of Thesis/Dissertation:

Stereoscopic Vision's Impact on Spatial Ability Testing

For the degree of Master of Science

I certify that in the preparation of this thesis, I have observed the provisions of *Purdue University Executive Memorandum No. C-22,* September 6, 1991, *Policy on Integrity in Research.**

Further, I certify that this work is free of plagiarism and all materials appearing in this thesis/dissertation have been properly quoted and attributed.

I certify that all copyrighted material incorporated into this thesis/dissertation is in compliance with the United States' copyright law and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless Purdue University from any and all claims that may be asserted or that may arise from any copyright violation.

George Takahashi

Printed Name and Signature of Candidate

7/13/2011

Date (month/day/year)

*Located at http://www.purdue.edu/policies/pages/teach_res_outreach/c_22.html

STEREOSCOPIC VISION'S IMPACT ON SPATIAL ABILITY TESTING

A Thesis

Submitted to the Faculty

of

Purdue University

by

George Takahashi

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 2011

Purdue University

West Lafayette, Indiana

Dedicated to Stacie, for being patient with me as I spent so much time away from home on this endeavor.

ACNOWLEDGEMENTS

The author would like to thank his committee and anyone who put up with him and his need for answers:

Dr. Craig Miller
Dr. Patrick Connolly
Dr. James Mohler
Dr. La Verne Harris
David Braun
Christian Barrett

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

Terms that will be used in this document may require defining or may be abbreviated.

- Depth Cue: indicators that aid in the visual system to infer depth (Schwartz, 2010).
- Ghosting: Also known as crosstalk, is defined as a visual artifact in stereo displays created by the improper filtering between left and right channels (Woods & Newell, 2004)
- IPD: Inter Pupillary Distance, also known as interocular distance, refers to the distance between an individual's pupils (Schwartz, 2010)
- Parallax: Distancing between left and right channel on a stereoscopic display (Autodesk, 2008)
- PSVT: Purdue Spatial Visualization Test is a test that was created by Guay (1976) to be used as a meter for participant's spatial ability without being complicated by analytical processing
- Spatial ability: capacity in the processing of non-linguistic information or the performance on spatial tests (Eliot & Smith, 1983)

ABSTRACT

Takahashi, George. M.S., Purdue University, August 2011. Stereoscopic Vision's Impact on Visual Spatial Ability, Major Professor: Patrick Connolly.

A look into spatial ability testing tools and the variations that past researchers made to focus on key factors that affect test scores, will demonstrate the need for tuning traditional testing methods to accommodate a wider demographic and provide more accurate results. Due to technological limitations of the time, a large variety of past spatial tests were developed by hand-drawings. Within this research, the addition of stereoscopic vision is analyzed to determine the value of said changes on human perception of spatial entities.

CHAPTER 1. INTRODUCTION

The process of perceiving, analyzing and solving a spatial problem is an important and continuous procedure that is performed by engineers, educators, graphics technologists (Jensen, 1986) and by anyone who can perceive and interact with a spatial environment. An individual's spatial ability is defined as one's capability to perform this process and includes several categories in which the process can be divided (Hart & Moore, 1973). In its simplest form, even the arranging of luggage inside of a car trunk can be considered a spatial problem. The comprehension of each component's spatial dimension and operating within this constraint to produce a solution is a prime example of a spatial problem, very similar to the sorting toys used by Örnkloo & Von Hofsten (2009).

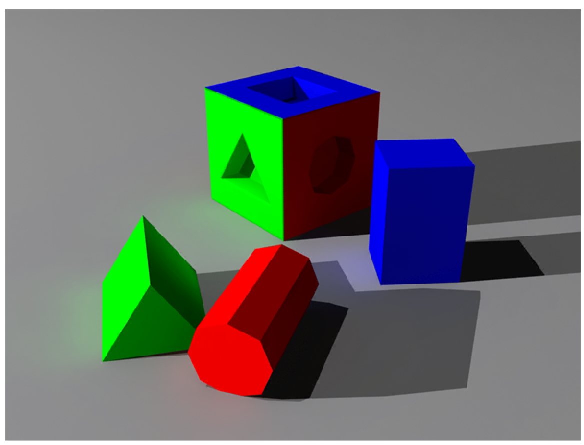

Figure 1.1 Rendering of a spatial sorting toy

To quantify an individual's capacity for solving spatial problems, visual spatial ability tests are commonly given by measuring time or accuracy on a spatial task (Study, 2001). Visual spatial ability tests in particular, traditionally were developed on paper and were usually comprised of pictorial representations of spatial stimuli. An example of a visual spatial ability test is the Purdue Spatial Visualization Test (Guay, 1976). The PSVT illustrates an isometric representation of a three dimensional object and presents a type of rotation to which the participant must correlate to a different three-dimensional model (Study, 2001). Within the realm of visual spatial ability tests, two-dimensional (2D) representations of three-dimensional (3D) objects are often displayed in an isometric view such as The Mental Rotation Test (Vandenberg & Kuse, 1978), PSVT (Guay, 1976), Three Dimensional Cube (3DC) (Gittler & Glueck, 1998), and Mental Cutting test (MCT) (CEEB, 1939). The human vision system is described by Hoffman (1998) to allow the perceived images from the eye to construct depth that might not necessarily be there. In the Necker Cube as seen in Figure 1.2, a drawing published in 1832 by Louis Albert Necker, a sense of depth can be perceived when in reality, they are only an assortment of 12 lines on a 2D medium (Einhauser, Martin & Konig, 2004). In both art and engineering graphics, this wireframe is described to be represented in an oblique perspective, where the perspective is not exactly isometric, yet no additional depth cues are provided (Einhauser, Martin & Konig, 2004). These pictoral representations of 3D objects however, are very different to how humans view the objects they represent. In an attempt to investigate the effect of changing how these pictoral representations are perceived in an academic setting, this research observed the effect of stereoscopy on spatial ability testing.

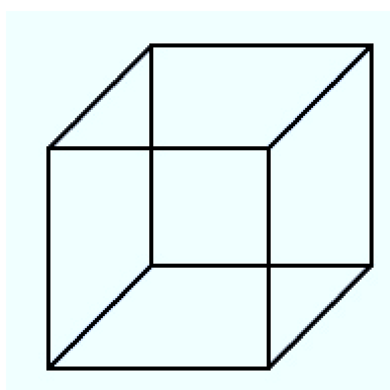

Figure 1.2 Representation of a Necker Cube (Necker, 1832)

1.1. Research Question

What is the effect of stereoscopic vision on spatial ability testing when compared to testing without stereoscopic vision? Stereoscopic vision refers to a depth cue that is achieved by the viewing of an entity with two separate perspectives for each eye, which when processed by the brain and results in a notion of depth. This effect can be simulated by viewing a display that channels a separate perspective for each eye. By producing this effect on a spatial ability test, a change in how the problem is perceived is forced on the viewer without affecting the procedure or the expected answer of the problem. It is the goal of this research to quantify the effect of stereoscopic vision on a spatial ability test.

1.2. Scope

The scope of this research was limited to measuring the spatial ability of college students in the field of Computer Graphics Technologies at Purdue University. The study was conducted on Purdue University campus and did not take into consideration the difference of gender, age, and cultural background. The study was conducted with voluntary participants within an introductory course in the Purdue University Department of Computer Graphics Technology curriculum. The study occurred between January 2011 and May 2011 and was limited to English language tests only. The pre-test and monoscopic post-test was conducted online and used an automated timer and scoring, while the stereoscopic post-test was conducted at the Purdue Envision Center, where specialized active stereoscopic monitors were used.

1.3. Significance

In researching the effects of stereoscopy on spatial ability tests, if stereoscopic vision has a direct influence in spatial performance, I believed it could indicate an issue with traditional spatial tests. It could also support the possibility of utilizing stereoscopy as a remedial tool for students with low spatial ability. Current spatial visualization tests require participants to complete a series of mental problems such as mental cutting tests or mental rotation tests with occlusion as the only depth cue. In order to isolate the impact of visual depth cues on spatial ability tests, this research devised a test that isolates specific cues and created a comparative study on the effects of said cue in the spatial testing process.

1.4. Purpose of the study

This study's purpose was to:

- Determine the difference in testing score between stereo and non-stereo versions of tests based on the Purdue Spatial Visualization Test: Visualization of Rotations.
- Investigate the effects of stereoscopic vision on students with low spatial ability on spatial ability test scores.
- Analyze the effectiveness of stereoscopy in aiding engineering graphics students perform in spatial ability tests and discuss the implications of this effect or lack thereof.
- Observe if a significant difference was discovered between a stereoscopic and non-stereoscopic version of the test, discuss the possible issues in using traditional spatial ability tests with limited depth cues and possible future research to alleviate said issues.

1.5. Hypotheses

The null hypotheses for the study were:

Hypothesis 1: There was no statistically significant difference between the scores of the stereoscopic and monoscopic versions of the Purdue Spatial Visualization Test: Visualization of Rotations taken by the introductory engineering graphics course students.

Hypothesis 2: There was no statistically significant difference between the scores of the stereoscopic and monoscopic versions of the Purdue Spatial Visualization Test: Visualization of Rotations taken by the students of the introductory engineering graphics course who have low spatial ability.

Hypothesis 3: There was no statistically significant difference between the scores of the stereoscopic and monoscopic versions of the Purdue Spatial Visualization Test: Visualization of Rotations taken by the students of the introductory engineering graphics course who have high spatial ability.

1.6. Assumptions

The following assumptions for the study were:

- The participant was placed at the pre-calibrated distance from the screen
- Nausea, headaches, or general discomfort may have occurred when dealing with stereoscopic 3D environments
- The interpupillary distance (IPD) is calibrated to the standard American average
- The participants maintained their heads at a pre-set distance and angle.
- The pre-test and post-tests are an effective tool for measuring the spatial ability of this sample.

1.7. Delimitations

The following delimitations for the study were:

- The research did not accommodate for different IPD
- The study isolated and investigated the effect of stereoscopy and no other depth cue.
- The Purdue Spatial Visualization Test: Visualization of Rotations was used to compare the stereo and mono test.

- Lighting conditions in the room with stereoscopic displays was set to dim to prevent interference with the active shutter glasses.
- IR sources was limited or turned off to avoid interference with the active shutter glasses.
- The 3D objects in the PSVT were rendered individually and then merged into a composition

1.8. Limitations

The following limitations for the study were:

- Hardware and software limitations apply to the specifications described in the methodology.
- The online tests were limited by the participant's internet connection speed and should only be taken on a modern browser.
- The hardware used to produce the stereoscopic effect caused minor ghosting.
- The clarity of the picture on stereo displays as seen by the participant is dependent on the angle of the glasses to the monitor.

1.9. Chapter summary

This chapter placed a frame of reference to help understand the research of this thesis by establishing the problem and the appropriate question to generate an answer. This chapter also defined the scope and significance, along with the assumptions, delimitations and limitations of this research.

CHAPTER 2. LITERATURE REVIEW

This chapter brings forth a summary of relevant literature that applies to spatial ability and stereoscopy. It is intended to create a defined level of knowledge before diving into how the research question will be answered.

2.1. Spatial Ability

Past research on spatial ability including the history, development and the tests used to measure are discussed in the following sections.

2.1.1. Spatial Research

Research in spatial cognition and development of spatial abilities has been shown repeatedly to be a vital component of success in a wide range of engineering, technical, mathematical and scientific positions (Miller, 1992). Interest in spatial abilities were constrained to the fields of psychology, yet recent developments in engineering graphics and visual display has provided educators with new tools to approach spatial research in new ways and by other fields. Engineering graphic educators have meaningful experience that classifies them as experts in the development and quantification of spatial abilities. With the advent of technology and its usage in engineering graphics, educators in this profession have become increasingly interested in harnessing the capabilities for visualization. Early investigations on spatial abilities date back to the early 1900's, when a spatial component was discovered in the process of quantifying human intelligence.

2.1.2. History of Spatial Research

Research in spatial ability can be separated into four phases throughout history. The first phase (1901-1938) was an effort by psychologists to establish and identify the presence of a spatial factor in intelligence (Miller, 1992). Previous testing methods, such as the alpha/beta test conducted by the US Army, were severely biased toward verbally-skilled individuals and considered those without verbal skills uneducated. Following research studies by El Koussy (1935), Kelly (1928) and Thurstone (1950), identified of the spatial factor as a very important aspect of human intelligence (Elliot & Smith, 1983), and prompted a new phase in assessing the importance of spatial abilities and its factors.

The second phase (1938-1961) noted by Elliot and Smith (1983), identified different spatial factors and how they varied from one another. This stage consisted of many large-scale research studies in which a large number of tests arose in spatial relations, visualization, spatial orientation and imagery. These tests can be classified into: (1) the ability to recognize configurations and (2) the ability to manipulate spatial configurations.

The third phase of research (1961-1982), was centered on studies determining the association of spatial abilities with other abilities and the discovery of different sources of variance in testing of spatial abilities as noted by Miller (1992). Some of these differences were correlated to a person's age, sex, environmental upbringing, and hereditary influences. Connection of an individual’s spatial ability to their preferred learning style has also been sought out during the latter stages, specifically the relation between visual and haptic learners and their performance based on the type of test they are administered.

A fourth phase of research can be observed post 1983, in which research takes a shift towards modern technology. Among continuing research from the third period, emerged this new class of testing with the usage of Virtual Reality (VR) and Human Computer Interfaces (HCI) (Study, 2002). Moffat, et al. (2002), and Kauffmann (2000) are examples of this phase in which augmented reality (AR) is used to complement and enhance the visual perception of spatial entities.

2.1.3. Spatial Theories

In defining spatial cognition, many researchers reached conflicting views on how it is explained or differentiated from spatial perception or spatial ability. Psychologists define spatial cognition as "inner space or spatial cognition, the spatial features, properties, categories and relations in terms of which we perceive, store and remember objects, persons, events..." and is the basis for which "...we construct explicit, lexical, geometric, cartographic and artistic representations" (Olson & Bialystok, 1983, p.2). Spatial can be subdivided into the subsystems: spatial perception, spatial thinking, spatial imagination, spatial reasoning, spatial judging and spatial memory (Hart & Moore, 1973).

Much of the conflicting views arise in defining spatial cognition and its separation to spatial perception. Unlike Hart and Moore (1973), some researchers believe that spatial perception shares a direct connection with spatial cognition. As stated by Arnheim (1986), "perceiving and thinking require each other. They complement each other's functions"..." Perception would be useless without thinking; thinking without perception would have nothing to think about" (p. 135). Another distinction required in defining spatial cognition is the differentiation to spatial abilities. The definition of spatial abilities is another topic which brings much debate and controversy. For the purpose of this thesis it is established that spatial abilities refers to an individual's spatial cognition, and when tasked with various spatial ability measurements and tasks, said individual's performance will vary depending on their spatial cognition. Spatial cognition will be understood as the level of comprehension and knowledge on spatial factors, while spatial abilities refer to the capability of an individual to perform a spatial task with this comprehension. It is also noted that spatial ability will be defined as the capacity of an individual to perform a spatial task and is understood as a concept that can be taught and developed rather than a pre-set limit to their capabilities (Sorby, 1999). The distinction is explained by Miller (1992) in defining spatial cognition as the "underlying mental process that allows an individual to develop spatial abilities" (p. 30).

Spatial ability is considered to be the over-arching category that describes the "skill in representing, transforming, generating and recalling symbolic, nonlinguistic information" (Linn & Peterson, 1985, p. 1482). However, when defining spatial perception, it is important to not confuse the varying definitions and usages of the term. According to Thurstone (1950), spatial perception is a factor of spatial ability and is defined as the ability to relate one's orientation to the orientation of the spatial entity, while Carroll (1993) describes Perceptual Speed as the "speed in finding a known visual pattern, or in accurately comparing one or more patterns, in a visual field such that the patterns are not disguised or obscured" (p. 363).

2.1.4. Spatial Abilities

In defining spatial abilities, Lohman and Kyllonen (1983) identified three major components:

- Spatial Relations: "Tests that are parallel forms of one another and the factor emerges only if these or highly similar tests are included in the battery. Although mental rotations are the most common element, the factor probably does not represent the speed of mental rotation; rather it represents the ability to solve such problems quickly by whatever means."(p. 111)
- Spatial Orientation: "The ability to imagine how a stimulus array will appear from another perspective. In the true spatial orientation test, the participant must imagine that they are reoriented in space and then make some judgment about the situation." (p. 111)
- Spatial Visualization: "The tests load on visualization, in addition to their spatial-figural content, share two important features: they are all administered under relative un-timed conditions and most are much more complex than corresponding tests that load on more peripheral factors." (p. 111)

McGee (1979) on the other hand defines spatial visualization as “the ability to mentally manipulate, rotate twist, or invert pictorially presented visual stimuli. The underlying ability seems to involve a process of recognition, retention and recall of a configuration in which there is movement among the internal parts of the configuration, or of an object manipulated in three dimensional spaces, of the folding or unfolding of flat patterns…” (p. 3-4).

Researchers debate on their definition and classification of spatial abilities, however agree on the existence of certain factors and they can be me measured through testing instruments. In the process of developing these test instruments, differences in test scores are noted. According to Liben, Patterson and Newcombe (1981), certain differences appear between test participant's spatial cognition depending on their individual characteristics, cultural heritage, and qualifications. These qualities can be grouped into physical cognitive and socio-emotional. Out of the listed, age is the most influential factor in these categories. Liben states that “depending on the environmental exposure through the various spatial development stages, an individual may be exposed to various environments or environmental experiences that either advance or hider his or her spatial cognitive development and abilities” (Liben, Patterson & Newcombe, 1981, p. 17-19).

2.1.5. Methods of Developing Spatial Cognition

As noted by Sorby (1999), the spatial ability of any particular individual as measured by a spatial test, can fluctuate based on the training and preparation given prior to testing. Construction engineering educators often expose students to simple construction and analysis tasks in which it is necessary to observe an isometric paper representation of a three dimensional object to then later draw the perspective appearance of said object from above, front and side as an exercise of spatial processing. Other tools such as mental cutting plane exercises where a student is asked to identify the appearance of a surface when a two-dimensional representation of a three-dimensional object is intersected by a plane. Another popular method for developing spatial cognition is to require students to mentally analyze the effects of finding the union, intersection, or subtraction of one three dimensional object to another. Many of these tasks however require participants to have basic freehand illustration abilities, which could cause the results to fluctuate below their actual level.

2.1.6. Spatial Tests

An extensive set of pencil and paper tests were used to measure spatial ability in the past. Many of these required the recognition of mental forms or the mental manipulation of visual shapes. Thus, the spatial tests can be separated into a recognition or manipulation groups. Visual memory tests, copying tests, and embedded figures tests are examples of recognition tests. Surface development, paper folding and rotation tests would be examples of manipulation tests. Another characterization involved spatial tests that could be solved within or across a two dimensional plane. Form completion tasks would be within plane tasks and rotation tests would be examples of across plane tests. Another characteristic of these tests focused on the mental transformations; the more transformations, the more complex the test became. Eliot and Smith (1983) grouped paper-and-pencil spatial ability tests into three categories. The first group is divided into recognition and manipulation tests in which they measure visual memory and surface development, paper folding and rotation respectively. The second group is characterized by tests that remain in a two dimensional plane and can be solved without leaving said plane. The third group involves mental transformations.

Study (2001) describes spatial tasks that require the participants to match a stimulus to an identical item are simple in nature, while tasking the participant to recognize a stimulus as viewed from a different angle is more complex. One example of this type of spatial test is the Purdue Spatial Visualization Test (Guay, 1976) seen in Figure 2.1. It is commonly used as a reliable meter for spatial ability without being complicated by analytical processing (Bodner & McMillen, 1986). George M. Bodner, a Purdue chemical educator, utilized this test to analyze introductory chemistry students' spatial ability to predict and later contrast to proficiency in molecular structure arrangement. Other examples of this form of test include the Vandenberg Mental Rotations Test and the Shepard-Metzler Rotations Test. As cited by Study (2001), the PSVT's "internal consistency, KR-20 coefficient, has been reported from .80 to .92" (p. 34).

Additional depth cues that are not present in the PSVT are binocular cues, lighting, and perspective among others (Pinker, 1997). The only monocular cue that is present in the PSVT and many other pen and paper spatial ability tests is occlusion. Without occlusion, the objects represented in the test would appear to be similar to the Necker Cube, a wireframe with transparent faces.

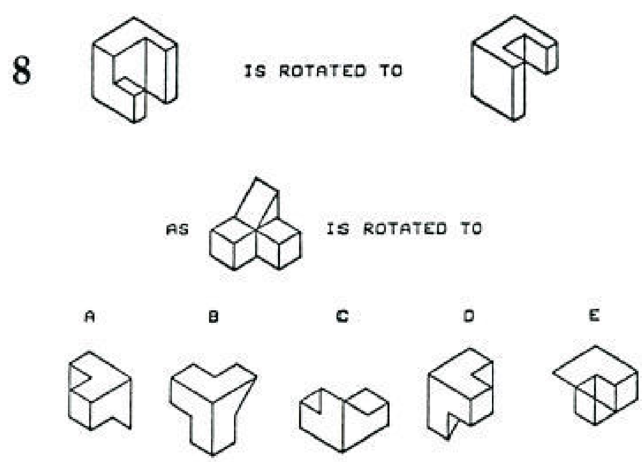

Figure 2.1 Example of the PSVT:R Test

Jianping Yue (2008) claims in the process of designing isometric sketches for the basis of testing spatial ability "... distort true pictorial views, and are prone to drawing errors" (Jianping, 2008 p 36). In order to prevent inaccurate measurement of an individual's spatial capabilities, the information represented within the test must be conveyed accurately and effectively. An example of errors can be found in PSVT-R isometric sketches in a case study of isometric drawing errors by Jianping (2007) in Table 2.1. In the PSVT-R, Jianping outlines seven questions with errors out of the thirty questions in the test that include missing features, misrepresented features, and the inclusion of extra features. It is also seen how two questions represent the same object and rotation in the example rotation.

Table 2.1
Summary of Errors in PSVT-R Test by Jianping Yue (2007)

Item	Question Number	Drawing Number	Error
1	8*	Example rotation	Missing features
2	10*	Example rotation	Missing features
3	13	Example rotation	Missing features
4	13	A	Extra features
5	13	D	Missing features
6	14	A	Missing features
7	14	E	Missing features
8	15	C	Extra features
9	17	Example rotation	Missing features
10	25	B	Misrepresented feature

* Questions #8 and #10 share the same exemplary object and its rotations.

2.2. Perception

Previous studies of spatial ability found a difference in resulting scores by changing how the problem was presented to the viewer. Human perception can be defined as the link between the physiological and cognitive process of retrieving external stimuli for the purpose of processing (Schwartz, 2010). In changing the perceptive process by altering a display method, spatial tasks can be modified. As demonstrated by the research of Jianping (2008) and Tsutsumi et al. (1999), changes in how spatial ability tests were presented, did not significantly alter the end score. However, research by Aitsiselmi and Holliman (2009) claimed a significant difference in test scores by adding the depth cue of stereopsis. By using digital setup, additional control is given to depth cue parameters. In virtual settings unlike in real environments, binocular vision, shading, lighting, perspective and occlusion can be customized for greater control (Pinker, 1997)

2.2.1. Depth Cues

Depth cues, refer to indicators that aid the human visual system to infer depth (Schwartz, 2010). These indicators contribute in providing proper depth perception. The depth cues can be split into two main categories, monocular depth cues and binocular depth cues. Monocular cues refer to the indicators that provide depth information with a requirement of only a single eye. Binocular cues however refer to an indicator that is different for each eye and said difference allows the viewer to infer depth with a requirement of both eyes (Schwartz, 2010). These cues can also be broken up into further groupings of physiological depth cues and psychological depth cues (Teittinen, 2011). Schwartz (2010) describes these depth cues in detail as:

- Monocular depth cues:
    - Physiological
        - Accommodation: refers to an occulomotor cue caused by the muscles that allow an eye to focus on far away objects by morphing the eye lens thinner and thicker.
    - Psychological
        - Perspective: Similar to relative size, perspective refers to the converging of parallel lines as they move farther from the perspective.
        - Relative size: Refers to the size differences between two similar objects, as the objects become bigger, they appear to be closer.
        - Familiar size: When viewing objects of a known size, the object's perceived size when compared to previous encounters with a similar object, one can assume a distance.

        - Occlusion: Also known as Interposition, is a cue that occurs when part of a scene or object is partially obstructing another.
        - Texture: Based on the texture of the object that is being viewed, if the texture seems more densely packed, it would indicate that the viewed texture is farther from the viewer.
        - Lighting/Shading: Based on how light reflects or lights an object and casts shadows, their position and shape can be inferred.
        - Blurring: Similar to occlusion, if a portion of a view is obscured by fog, blur, haze or rain, one can infer depth.
- Binocular depth cues:
    - Physiological
        - Convergence: the occulomotor binocular cue to infer depth by the two eyes rotating inward to focus on an object that is closer.
    - Psychological
        - Stereopsis: By having two separate images projected into the eyes, each with individual separate perspectives, allows the brain to fuse the images into a single view. The depth perception produced by stereopsis only can be processed if the view falls in the fusion distance. If this disparity is too large, the brain will not be able to re-compose the two views, resulting in physiological piplopia, also known as double vision (Schwartz, 2010).

2.2.2. Stereoscopy and 3D displays

In order to present a pictoral representation of a 3D object on a 2D screen, the stereoscopic effect must be simulated. When using a stereoscopic display, the pictoral representation is limited to displaying images on the extents of said screen. Because of this limitation, to achieve stereopsis, displays will separate a left and right image into two separate channels so the end viewer will receive a specific channel to the designated eye. The two main methods to achieve a stereoscopic display can be grouped into passive and active displays (Autodesk, 2008).

- Active stereo:
  - Shutter glasses: Commonly use liquid crystal screens that block light for alternating eyes in synchronization with the images on the computer display. The display system will alternate between left and right channel to be seen by left and right eye, this technique is also known as frame-sequential.
- Passive stereo:
  - Polarized glasses: by polarizing plastic screens, left and right channel can be separated if the display is projected through two separate screens of opposite polarization that match the polarization of the matching lenses in the glasses.
  - Anaglyph glasses refer to color coding each channel to complementary colors, and using colored glasses to filter out the matching colors such as red-cyan glasses.
  - Autostereoscopy: is a display technique that separates a display for each individual eye. This is usually achieved by having two individual displays or by splitting an image per pixel to be directed to each individual eye.

Any stereoscopic display system will be affected by a left-right distancing called parallax. This term is used to describe the separation between the two channels as seen in Figure 2.2. When the displayed object is behind the physical display or render plane, the left and right images are crossed as seen in Figure 5.C and is called Negative Parallax. While if the object is in front of the render plane, the parallax will become positive. When the figures overlay on top of each other, the display is said to have zero parallax. However, if the left and right image separate further than the Inter Pupillary Distance (IPD, also known as interocular separation), the image will become hard to fuse as this does not occur in normal viewing (Autodesk, 2008).

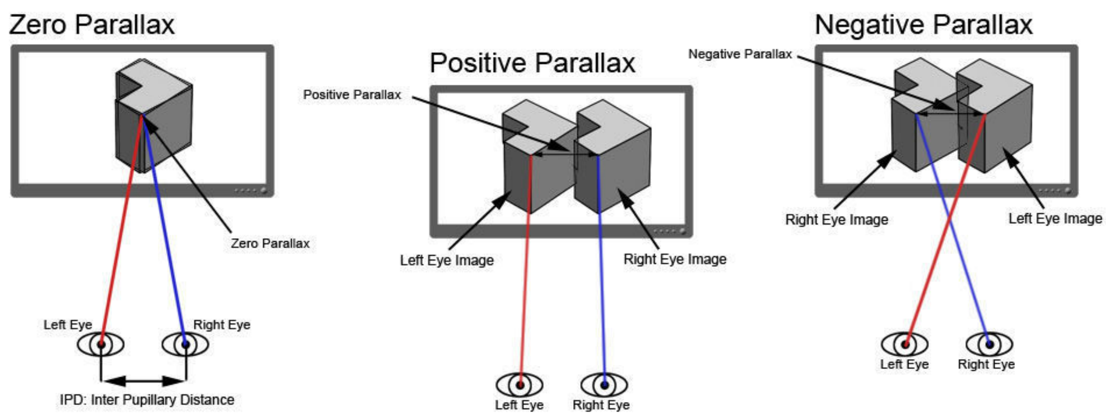

Figure 2.2 Types of Parallax

When shooting a scene in stereo with live cameras or calibrating in a 3D graphical application, the left and right cameras must be calibrated to a separation and rotation that is analogous to that of human eyes. In order to have both cameras focus in on the scene of interest, often the cameras will have to toe-in which will cause additional errors in viewing such as a trapezoidal distortion known as keystoning (Figure 2.3), however this will allow the point of zero parallax to be set to a particular distance that matches the viewer's distance to the screen. Stereoscopic displays display the two channels, left and right, by different methods, however each of these display methods will suffer under some level of crosstalk or ghosting. Ghosting is a critical factor in determining the quality of the stereoscopic display (Woods & Newell, 2004). Crosstalk is defined as the image from a single channel being displayed (however faintly) in the other channel.

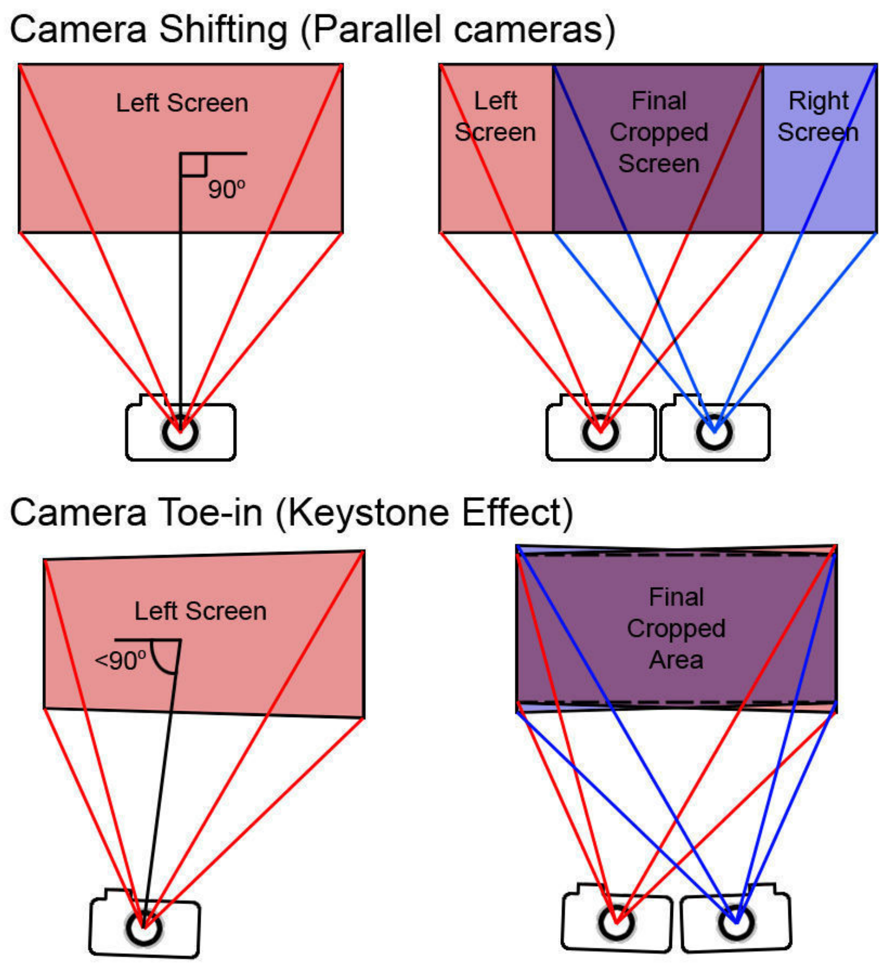

Figure 2.3 Toe-in cameras and Parallel cameras

2.2.3. Stereoscopic Research

In a qualitative study performed by Leroy, Fuchs, Paljie and Moreau (2009), participants who viewed virtual objects on a screen preferred interacting with it in stereo, however the largest factor on perception of shape was caused by using head-tracking. It was also noted that incorrect IPD caused discomfort to the viewers and by association, incorrect distance from the screen would cause increased or decreased parallax. In an attempt to validate the usage of stereoscopic displays, a study conducted by Aitsiselmi and Holliman (2009) measured individual's spatial ability using a replicated computerized version of Vandenberg Mental Rotation Test (Vandenberg & Kuse 1978). This setup was rendered by skewing the individual projection frusta of each camera to match the render plane of the screen as shown in Figure 2.4.

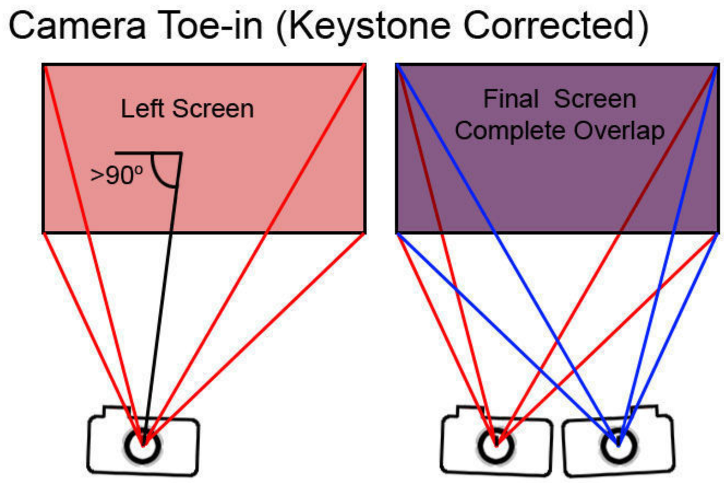

Figure 2.4 Camera Toe-in with corrected keystone

The 32 participants that performed the questionnaire, were split into two equal sized groups and asked to answer 15 mental rotation questions on the auto-stereoscopic display. Results from this study demonstrated an increase in average score of the 3D group up to 14% improvement over the 2D group (Aitsiselmi & Holliman, 2009) with a significance of $p = 0.05$. By taking into account the scores of these participants and the time required to complete it, the research demonstrated how stereoscopic displays aid in "deciphering the shapes in front of them due to the increase in depth information..." (Aitsiselmi & Holliman, 2009 pg 12). However, both time and score values were not inter-correlated and thus, it is hard to validate the usefulness of stereoscopic displays. When observing the score values over time, the research uses the score as the independent variable and the time as the dependant.

This research does look into the time differences between the two groups, however, the time taken to complete the study should be used as a predictor variable to properly adjust the significance of each test. A study by Tsutsumi et al. (1999), observed female participants performing a stereoscopic mental cutting test to analyze their spatial abilities. However, when comparing the average scores of the stereoscopic MCT and regular MCT, "the stereogram did not have any effect on complicated mental image processing tasks, such as transformation of a section to a true shape" and " Low scoring participants in this study could not recognize the test solid and its cutting plane well, and they were unable to construct complete images of the objects, even when they used stereograms" (Tsutsumi et al. 1999 pg 8).

2.3. Chapter Summary

Overall, the effects of stereos copy on spatial testing seem somewhat ambiguous and results from multiple past research indicate conflicting results. Stereoscopic displays, to function properly, must alter how real human perception operates and fake multiple parameters to provide the illusion of depth. In order to create an exact replica of what the human eyes perceive, specific calibration on distances to screen, IPD, screen dimensions, screen resolution, parallax, keystone and toe-in must be taken into account. Additionally, a pre-test that defines a test participant's spatial ability, to then correlate to each versions of the post-test (stereo and non-stereo), would provide insight on how stereoscopy affects spatial testing. By using the pre-test as a predictor variable that establishes a baseline across all test participants, proper comparison between the two groups could be determined.

CHAPTER 3. METHODOLOGY

The purpose of this study was to determine the effect of stereoscopy on spatial ability testing. This chapter outlines the research describing the test framework, population sample, methodology and measurements used for this study.

3.1. Study Design

The research in visual spatial ability was analyzed by quantitative means. This analysis contrasted the results of the devised testing mechanism to a well founded reliable test. For this research, the *Purdue Spatial Visualization Test* (Guay, 1976) was utilized as a source of comparison. In specific, the *Purdue Visualization of Rotations* portion was utilized and is often referred to as ROT. The purpose of this study was to determine the difference in testing score, between the Purdue Spatial Visualization Test: Visualization of Rotations with and without stereoscopic view. In order to create a test in which stereoscopic view could be compared, both stereo and non-stereo versions of the PSVT:R had been re-created and rendered using a modeling package to remove inconsistencies associated with the original sketched PSVT and also establish a similar image quality and characteristic between both versions. A pre-test was also used to establish a baseline of each participant's spatial testing aptitude.

3.2. Sampling

The population to whom the test was administered are Purdue University West Lafayette Campus students from an undergraduate introductory course

Computer Graphics Technologies, CGT 163, in The Department of Computer Graphics Technology (CGT) at Purdue University. The participation in this study was strictly voluntary and anonymous. The only data collected from each individual participant was their test score, and an identification number used to associate the pre-test, and post-test scores. The participants were recruited from within the introductory class as an extracurricular option for class extra credit. Due to the low percentage of female and minority presence in this group, the study was not designed to distinguish results based on gender, age, or ethnicity. Once the participants volunteered, they were divided into two groups. The split amongst the two groups was random and was established before information regarding which group receives which test was revealed. Participants used for this comparison were functionally similar to avoid biased results (Campbell & Stanley, 1966). Students within the CGT 163 class mostly consist of freshmen and sophomores enrolled within mechanical and aeronautical related professions including Mechanical Engineering, Aeronautical Engineering, Aviation Technology, Building and Construction Management, Mechanical Engineering Technology and Computer Graphics Technology. Students within this class are also mostly male within the ages of 18-20.

3.3. Testing Instruments

This study utilizes two spatial ability tests, the PSVT: Visualization of Rotations recreated and digitally rendered with and without a stereoscopic effect. The illustrated objects from the PSVT were converted into three dimensional versions by utilizing the isometric dimensions from within the test to then convert them into full three dimensional objects within a CAD program, and rendered into perspective views. Both versions of the PSVT consist of 30 questions of varying difficulty based on the original version of the test and placed within a time limit (Guay, 1976). The difficulty of the test varies depending on the angles of rotation and the axes upon which the rotations are implemented on. The simplest item requires one rotation of 90˚, the next difficulty requiring 180 about one axis, the third difficulty requiring 90 over two axes, and a fourth difficulty rotating 90 off one axes, then 180 about another (Guay, 1976). The objectives of each test is to determine the angle of rotation and apply an equivalent rotation or set of rotations to a different object, then match the end result to one of five different possibilities that are granted to the participants. The PSVT was selected because of its internal construct validity and it's usage to measure spatial ability (Study 2001). In the following Figure 3.1, the eighth question of the re-rendered non-stereo version of the PSVT is shown.

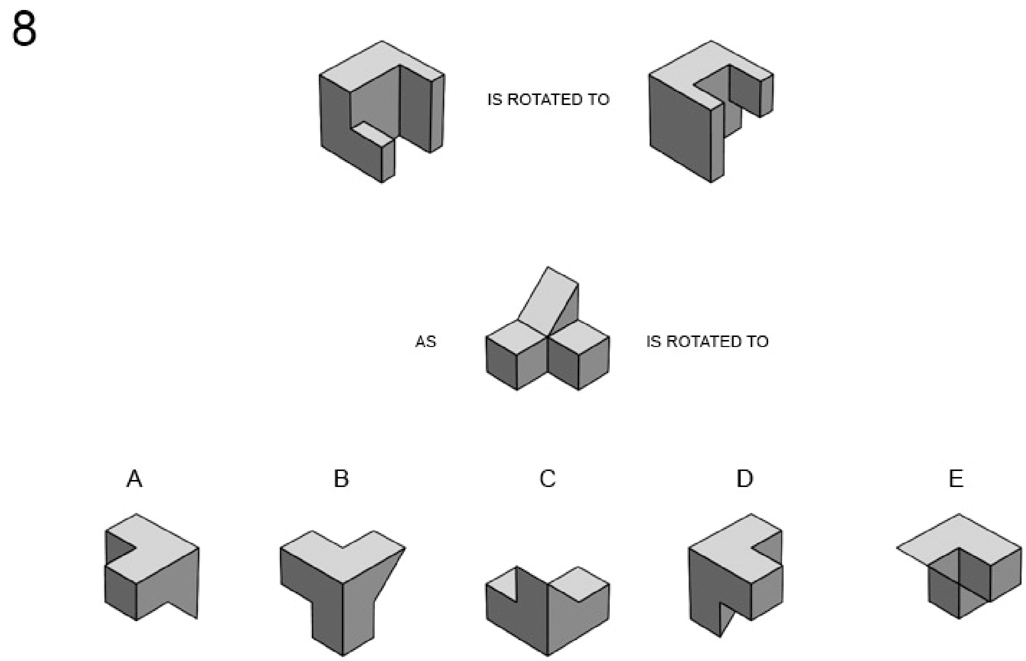

Figure 3.1 Rendered version of the PSVT

The graphics for the PSVT:R post-test were created in Autodesk 3D Studio Max 2011 and each of the components within the problem was rendered in perspective. In the stereo version, two virtual cameras were set at a pre-calibrated distance and then rendered to simulate the average human IPD at 23" from the screen. The distance from the camera to the model was set to 23" assuming the participant would view the screen from an average distance of 23" to the screen. This number was achieved through a set of trials where a comfortable viewing distance was established when using a 22" wide screen monitor. A screen capture of the 3D Studio Max environment can be seen in Figure 3.2. As seen in the following figure, three cameras are directed at the object within view and are toed-in to adjust for proximity and parallax. The cameras were adjusted to correct the keystoning effect.

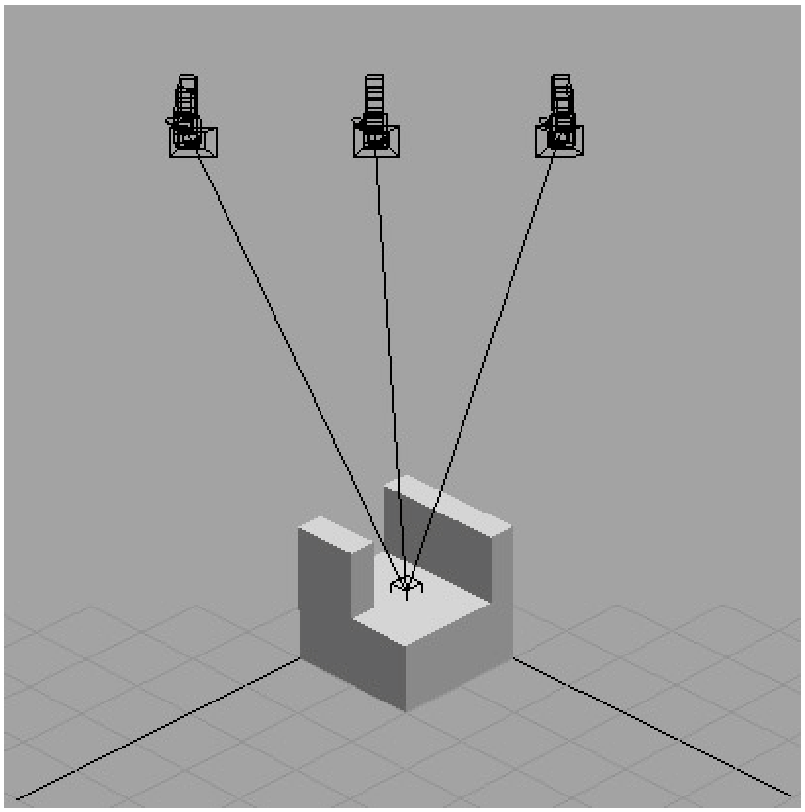

Figure 3.2 Screen Capture of setup in 3D Studio Max 2011

3.4. Testing Methodology

The experiment was conducted by requiring test participants to complete a pre-test to asses their baseline spatial ability, followed by one of two post-tests. The study used a re-drawn version Vandenberg Mental Rotation Test (Vandenberg & Kuse, 1978) as a pre-test, due to it's well documented usage and validity in measuring spatial cognition (Peters et al., 1995). This re-drawn version of the Vandenberg MRT was acquired from the Spatial Intelligence and Learning

Center at http://spatiallearning.org/resource-info/tests_index.html. The Test participants were then divided into two random groups, and were given the spatial ability test that corresponds to their group. One of the groups will take the PSVT:R with stereo and the other will take the PSVT:R without stereo.

The Vandenberg Mental Rotation Test consisted of 20 multiple choice questions, each containing four choices, of which, only two and always two are correct and the other two are erroneous. In order to avoid guessing, the questions required both correct choices to be selected in order to be given a single point for the question. A screenshot of the first question may be seen in figure 3.3. As seen in the figure the question number is illustrated at the top left, along with a timer displaying the amount of time remaining to complete the test. The participants were allowed to select two answers for each question and also were given the option to check their previous work if time still remained. Participants were also given an option to skip to specific questions or opt out of the test completely.

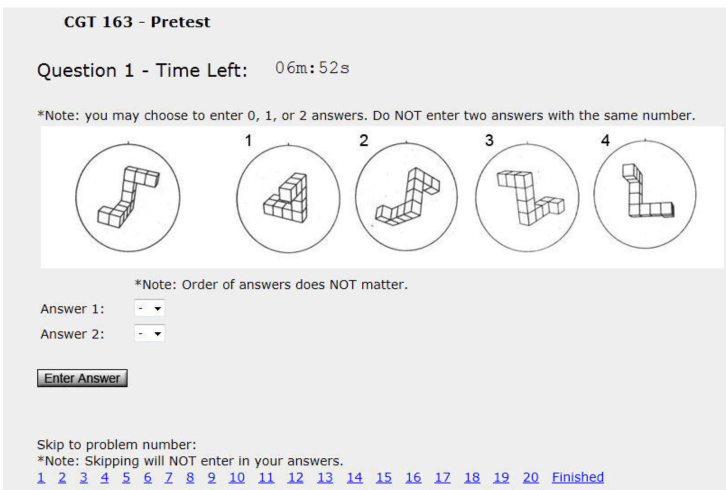

Figure 3.3 Screenshot of the first question in the pre-test

3.5. Data Analysis

The data retrieved from the post-test scores and the pre-test of each participant were analyzed. The values from the stereo and non-stereo versions of the PSVT will be statistically analyzed to find the correlation between these two studies in a quantitative measure by using an ANOVA and also ANCOVA with the pre-test as a predictor variable. ANCOVA or Analysis of Covariance refers to a statistical calculation that takes into account a categorical and continuous predictor variable in order to appropriately assess the effects of one or more factors (StatSoft, 2011). ANCOVA is used to adjust for the differences between the two groups and aid in finding the statistical difference between the means of each group. Participants from the lower spectrum of visualizers will then be re-analyzed separately for the effect of stereo with another ANOVA and ANCOVA. ANCOVA is considered to be a reliable tool for comparison when an additional variable is available to be used as a covariate (McMillan & Schumacher, 2001).

3.6. Chapter Summary

Key information for establishing a frame of research and the testing environment were described. The design, sampling and methodology have been outlined for the study and define the settings in which the participants were tested.

CHAPTER 4. DATA ANALYSIS

This chapter contains the results of the study. The data will first be examined in its entirety, and then will be later split up into groups of participants with low and high spatial ability.

4.1. Sample description

Students from the CGT 163 *Introduction To Graphics For Manufacturing* were presented with the opportunity to participate in this research. Of the 365 students enrolled in the academic spring semester of 2011, 218 participants volunteered and completed all of the requirements necessary to participate in the study. 106 of the participants were grouped into the stereoscopic version of the study, group A, while the remaining 112 were placed in group B, monoscopic version of the study. Participants were placed into their corresponding groups based on the sixth digit of a ten digit identification number given to the students on enrollment at Purdue University. Students with the sixth digit between the values of 0 to 4 were placed in group A, while values between 5 and 9 were placed in group B.

4.2. Pre-test: Vandenberg Mental Rotation Test

The test participants for this research accumulated to a total of 218 and began the testing phase by conducting an online version of the Vandenberg Mental Rotation Test. The results collected from the pre-test are shown in the following histogram, figure 4.1.

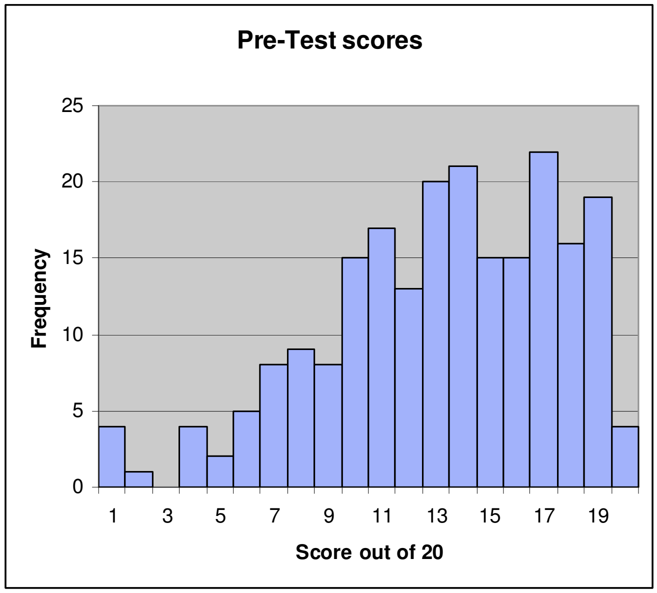

Figure 4.1 Histogram of pre-test combined students

As seen on the histogram of students who participated in the pre-test, the distribution of test scores illustrates a curve that resembles a negative skew. When the results are then split into the groups based on the post test the participant will take, it can be seen that both groups illustrated similar frequencies and produced a histogram with similar skew, seen on figure 4.2.

The participants from the stereoscopic group (group A) had an average score of 13.13 with a standard deviation of 3.98, while the monoscopic group (group B) had an average of 13.24 with a standard deviation of 4.72. Both total scores were taken from a total of 20 points. The combined average of both groups resulted in a score of 13.19 with a standard deviation of 4.37. Both groups, when subjected to the pre-test, were unaware of their group classification and were presented with the exact same pre-test and parameters. As expected, both groups scored similarly and demonstrated a similar distribution.

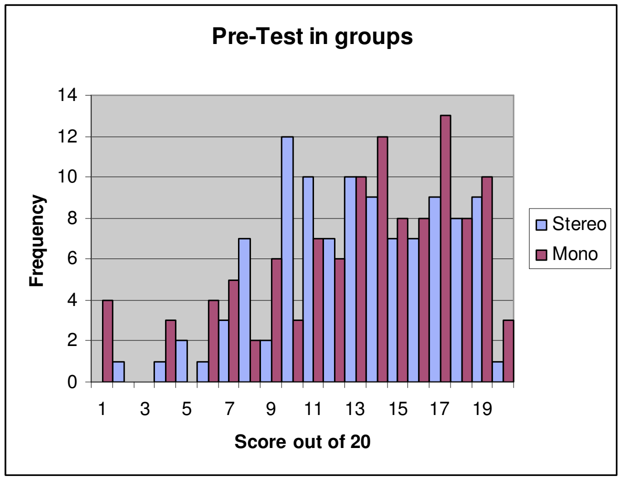

Figure 4.2 Histogram of pre-test as seen in groups

4.3. <u>Post-test: Stereoscopic/Monoscopic PSVT:R</u>

The results from the post-test are illustrated in the following histogram seen in figure 4.3. The average scores for the monoscopic group came to 23.44 with a standard deviation of 5.07, while the stereoscopic group scored an average of 23.74 with a standard deviation of 4.58. A summary of the values may be found on table 4.1. As seen by these two values, both groups scored very similarly, and to establish a baseline comparison, an ANCOVA is calculated with the pre-test (Vandenberg MRT) as a covariate.

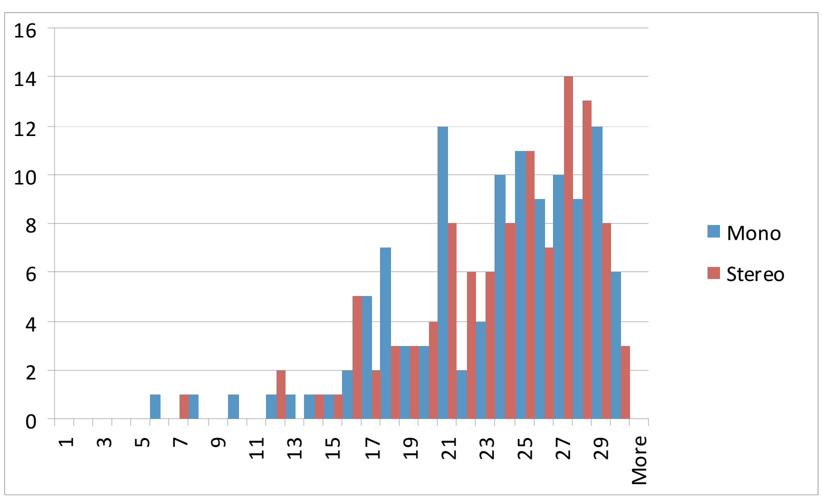

Figure 4.3 Post-Test histogram of the two groups

Table 4.1
Summary or results

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Stereo	106	2516	23.73585	20.93908
Mono	112	2625	23.4375	25.68975

In assuming that the pre-test is a precise and effective tool in calculating spatial ability, it is used as a covariate predictor variable. As observed in table 4.2, based on the score obtained from the pre-test, the correlation between the pre-test and post-test scores are highly significant, $p < 0.001$. However, when using the pre-test as a predictor, the comparison between the two post-tests, stereoscopic and monoscopic, demonstrates no significant difference $p = 0.541$. If a comparison between the two tests were to be done without a baseline variable were needed, a one way ANOVA can be used between the post-test groups with results as seen in table 4.3.

Table 4.2
ANCOVA results

General Linear Model: PostTest versus Group

```
Factor  Type   Levels  Values
Group   fixed       2  1, 2

Analysis of Variance for PostTest, using Adjusted SS for Tests

Source    DF   Seq SS   Adj SS   Adj MS      F      P
PreTest    1  1112.17  1114.18  1114.18  60.86  0.000
Group      1     6.86     6.86     6.86   0.37  0.541
Error    215  3935.99  3935.99    18.31
Total    217  5055.01

S = 4.27866   R-Sq = 22.14%   R-Sq(adj) = 21.41%

Term         Coef  SE Coef      T      P
Constant  16.7464   0.9235  18.13  0.000
PreTest   0.51873  0.06649   7.80  0.000
```

Table 4.3

ANOVA results between post-tests

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.847488	1	4.847488	0.207331	0.649325	3.88487
Within Groups	5050.166	216	23.3804			
Total	5055.014	217				

As seen in Table 4.3, the ANOVA and ANCOVA results both display findings of no significance, however the weight added by adding a covariate changed the significance from $p = 0.649$ to $p = 0.541$. When observing the lower end of the spectrum of spatial ability, and take into account participants who scored less than or equal to 10/20 in the pre-test, a significant difference in means with $p = 0.013$ was observed as seen in Table 4.4. This same analysis was repeated on participants who scored low on the pre-test with a weight added via a covariate variable; results became less significant and are seen in table 4.5. Correlation between pre and post-tests were less significant $p=0.052$, and between group differences of means demonstrates a difference with significance $p=0.058$. The following tests were based on a 95% confidence interval (CI).

Table 4.4

ANOVA results between post-tests of low visualizers

One-way ANOVA: Post versus Group

```
Source  DF      SS     MS     F      P
Group    1   153.1  153.1  6.57  0.013
Error   54  1258.9   23.3
Total   55  1412.0

S = 4.828   R-Sq = 10.84%   R-Sq(adj) = 9.19%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level   N    Mean  StDev   --+---------+---------+---------+-------
1      29  22.828  4.141                    (--------*--------)
2      27  19.519  5.473   (---------*--------)
                           --+---------+---------+---------+-------
                          18.0      20.0      22.0      24.0

Pooled StDev = 4.828
```

Table 4.5
ANCOVA results between post-tests of low visualizers

General Linear Model: Post versus Group

Factor	Type	Levels	Values
Group	fixed	2	1, 2

Analysis of Variance for Post, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Pre	1	144.30	90.16	90.16	3.96	0.052
Group	1	85.73	85.73	85.73	3.76	0.058
Error	51	1161.97	1161.97	22.78		
Total	53	1392.00				

S = 4.77324   R-Sq = 16.52%   R-Sq(adj) = 13.25%

Term	Coef	SE Coef	T	P
Constant	16.918	2.255	7.50	0.000
Pre	0.5688	0.2859	1.99	0.052

4.4. Effect Size

The effect size of the entire participant pool of post-tests using the group's mean and standard deviation resulted in a Cohen's d value of -0.0621 and an effect-size r of -0.0310. Observing the participants who scored low on the initial pre-test, the values returned were Cohen's d of 0.68189 and an effect size r of 0.3227.

4.5. Summary

Results from this study's pre-test and post-tests were outlined in this chapter. Statistical analysis on the resulting data also revealed quantitative findings on the research questions in this study. The next chapter outlines the significance of the findings and draws conclusions for the study. The next chapter will also take a look into the possible implications of these findings in academia and industry along with recommendations for future studies.

CHAPTER 5. CONCLUSIONS

This chapter presents interpretations and conclusions drawn from the data obtained from the research study and provides an analysis on the implications of these conclusions. It also outlines possible reasons for the data values and application of these findings in future research, industry, and academia. Thoughts on the usage of stereoscopy, methods of display and calibration procedures are also discussed.

5.1. Test Completion Time

This research included a completion time variable for both pre-test and post-tests, that kept a record of how quickly a participant completed the test. Initially intended to be used as method to track outliers, was then later discarded as the method of obtaining this value was different for both post-tests and due to the different nature of both versions of the test. The pre-test and the monoscopic version of the post-test were designed to be completed online. This online version allowed the user to constantly keep an eye on the remaining time as a countdown watch was placed at the top of every page. The online tests also only kept record of the last data entry change as the time spent on the test, this was later found inadequate as participants often returned to previous questions to check their work. If the participant took time to review their work, yet did not make any changes, the time spent on the test was not updated from the last entry point. The stereoscopic version of the post-test utilized a specialized stereo viewer and thus required an external source to display the time remaining. Verbal reminders were also given in the stereo version of the test at 5 minutes and 1 minute remaining on time to ensure proper time management.

As mentioned earlier, due to the differences in time measurement, completion time could not adequately be compared between both tests. It did however aid in locating outliers by indicating test members who seemed to complete the test unnaturally fast or participants who ran out of time before completing all the questions. An observation on the stereo version of the post-test, participants frequently seemed to finish the test without consuming all of the time allocated. As noted by some of the participants, the inconvenience of using specialized hardware to complete the stereo version of the post-test, outweighed the convenience of the online monoscopic version of the post-test. The online post-test was made available 24/7 during the allocated week of testing and allowed the participant to take the test at his/her leisure and at their location of choice. The online tests were designed to be accessed by any computer capable of online browsing by utilizing high accessibility web programming standards. The stereo version of the post-test however required the participant to schedule a 30 minute time slot at the Purdue Envision Center and due to the hardware limitations, only two students could be tested per time slot. Testing times also ranged from 9am to 9pm and required the participants to physically be present to participate.

5.2. Thoughts on Participation Honesty and Effort

As mentioned earlier, the recruiting for this research involved presenting an extra credit opportunity for the selected class members of CGT163 during the academic spring semester of Purdue University 2011. This research rewarded the participants with a 3% extra credit towards the final grade of the class. The research was conducted towards the end of the semester when the student's grades were further solidified by the work completed and their performance to date was readily available. Because of the timing for this research and the substantial reward that was given to participants, many students justified their participation as a necessity to improve their grades.

Participation in the study was recorded with four key components, a signed consent form, participation in pre-test, participation in post-test, and a final online login used to associate participation with a name without involving test scores. Each one of these components could be completed with minimal effort by simply signing in or writing the participant's name. Participation in the tests did not require completion and was considered valid as long as something was turned in at the end of the test. Participation was even considered valid if a student turned in a blank answer sheet or withdrew from the test. Throughout the entire testing population, a total of four students voluntarily withdrew from completing the test due to discomfort caused by the stereoscopic display system and were excluded from the data entries, however were awarded with the appropriate compensation.

Participants who did not complete the forms or participate in the tests without formally withdrawing from the test, were not rewarded with the extra credit. Due to the minimal requirements placed to obtain extra credit, a discussion on participation honesty and effort was brought up by some of the participants. These students voiced a concern on other participants abusing the opportunity for extra credit by placing minimal effort in the research and providing fake or dishonest input. However, as seen by the completion times of the participants and the scores obtained, a certain level of honesty in participation can be assumed. A few particular cases did arise during the research, where participants intentionally answered all of the questions with the same answer and completed the tests in mere seconds, however such participants were excluded from the research due to failing to participate in the other required components of participation. If this research were to be duplicated or used as a reference for another study, participation incentives, timing, and test environment would be tweaked in order to ensure common ground across all tests and promote participant honesty in testing.

5.3. Conclusions on Stereoscopy in Spatial Ability Testing

As demonstrated in chapter 4, the research conducted found no significant differences between the stereoscopic and monoscopic versions of the spatial ability test. This would seem to indicate that binocular vision does not aid introductory manufacturing students of CGT163 at Purdue University, academic spring semester of 2011 in performing spatial tasks. Possible reasons for this lack of difference could be caused by the stereoscopic effect not aiding in the image recognition process, or participants with generally high spatial ability are unaffected by display modes. Reasons for differences or lack thereof are uncertain for this specific research, however it is clear that students who have participated in the class of CGT163 did not obtain a significant advantage in taking the post-test with binocular vision.

Mentioned previously in the literature review, stereoscopy or binocular vision is a depth cue used to aid human vision in deciphering encoded visual content into comprehensible shapes and forms. Given that the sample pool of this research consisted of students who are generally adept at performing spatial tasks and had spent the semester performing spatial activities, adding binocular vision might not aid in the perceptive process. However, when observing participants who scored lower on the pre-test, a minor difference emerged between the monoscopic and stereoscopic tests. Previous studies performed by Aitsiselmi & Holliman (2009) and Tsutsumi et al. (1999), indicated a positive significant effect in testing participants with binocular vision. Yet the sample for each of these studies were varied and no data was included on their baseline spatial ability.

A major concern to many previous research tests is the lack of a baseline test to be used as a predictor. By adding a pre-test to properly weigh the participant's spatial ability, the researcher may account for individual differences in spatial ability and account for the increase or decrease in scores based on testing participants with monoscopic or stereoscopic vision. In studies with relatively low sample size and no weight added to the test scores, results become unstable and less predictable. But by accounting for individual testing discrepancies and having a large sample size, results become more representative of the population the group was sampled from (StatSoft, 2011). The small number of participants who scored low on the pre-test and the low significance of the weighted results shown in Table 4.5 illustrate a mild yet important indication of an effect of stereoscopy on spatial ability testing. As seen on Table 4.4, participants who took the stereo version of the post-test generally scored higher than participants from the mono version of the post-test with a high significance value when not weighed by a common model $p = 0.013$. This could possibly indicate that binocular vision has an impact in performance in spatial ability tests for participants with low spatial ability.

5.4. Thoughts on Spatial Ability Testing

In spatial ability testing, the ability to mentally orient, relate and/or visualize a spatial entity is quantified. However, in many of the time tested pen and paper tests administered to students, an additional factor is incorporated as a part of the test. As observed in this research and previous studies conducted by Tsutsumi et al. (1999), Aitsiselmi & Holliman (2009) and Jianping (2008), changes in how the participant perceives the question, changes the end score on the spatial test. If we are to assume that spatial ability of a participant who is to take a post-test using one display method while a separate participant is tasked with a different post-test, both participants will generally not undergo a change in ability to perform mental orientation, relation, visualization or manipulation when compared to each other. Assuming if an individual's spatial ability is unaffected taking a different post-test, if there is a difference in participant's score values, this would indicate the possibility of an external factor that is directly related to the change made on the specific test.

The Vandenberg MRT and the PSVT:R are designed as a speed test and are limited in time in order to artificially avoid a ceiling effect, however by observing the histogram of scores, it is obvious there is an accumulation toward the higher end of the spectrum in the pre-test and both post-tests. Possible solutions to avoid this issue is by shortening the allocated time to complete the test, thereby correcting the skew. For these tests due to the large quantity of participants, even with data that has a slight deviation from normality, statistical procedures can still be applied assuming a normal distribution (Moore, McCabe, & Craig, 2009).

In this research, both post-tests required to complete the same exact mental transformation between each problem, yet a significant difference was encountered when observing participants with low spatial ability. Participants who performed poorly in the pre-test seemed to score higher in the post-test if it was presented to them in stereo. The only difference between the stereo and non-stereo versions of the test was the different perspectives that are given to the stereo post-test to simulate depth by accommodating an individual offset perspective to each eye. This could be caused by stereo systems displaying additional data by seeing the same object from two separate perspectives, or by working as a depth cue and aiding in distinguishing certain concave/convex features of a model (Aitsiselmi & Holliman, 2009).

This minor change in display method, along with other study's changes on display methods, that cause a difference in test score regardless of assumed equivalent spatial abilities across groups, indicate a sub-factor that has an influence on spatial ability testing. This sub factor does not change how the spatial task is completed, however has shown an effect in the proficiency in the test. Because binocular vision is a cue used in perception, the difference between these two post-tests and the sub-factor that is influencing the scores is driven by the perception of the object itself.

5.5. Future Research

Throughout this study certain problems and issues were encountered and should be avoided in future research. This includes the proper selection of test timing and environment. Throughout this study, participants were disorganized and given the freedom to take the online tests at their leisure. This created an imbalance in the testing procedure and provided participants with an avenue to use resources that are otherwise unavailable in an academic test environment.

This study only observed the effect of stereoscopy on the PSVT:R. However, it would be beneficial to also study the effects of multiple depth cues in virtual or real settings as they compare to the traditional pen and paper versions of the test. Jianping (2008) attempted a comparative study between the traditional PSVT and a re-rendering of the PSVT using CAD software, yet the testing parameters and sample size could possibly be improved or expanded to incorporate other depth cues and a more robust methodology.

A possible setup for future research could be designed by replicating this study, targeted at individuals who struggle with spatial problems. To find participants who would qualify for the study sample, the research could include testing high school students ages 14-19 who struggle with spatial problems. To qualify for this research a set of pre-test could be used to quantify a user's spatial ability using the PSVT portions of paper folding and mental cutting test and the Vandenberg MRT in place of the PSVT:ROT. This pre-test can be used to discern participants who struggle with spatial problems while also using this score as a covariate during the analysis. The participants can be divided into four equal sized groups based on proficiency. Additional data on the participants should be recorded including participant age, grade, gender, corrective vision, two functional eyes, and experience with engineering graphics or drafting. The study will not discriminate but should anonymously include information on participants under the effects of psychoactive compounds that may affect perception.

After completing the pre-test, the students will take a post test that is heavily based on the PSVT:ROT with fixed issues and a change in the number of questions or time to complete to avoid a ceiling effect. An added measure to assure equal participation is to prevent early turn in to allow participants to spend any time left over to review previous questions. This future study could take into account the effect of stereoscopy while combined with other depth cues when compared to the original version of the PSVT. The results obtained from this research demonstrated inconclusive evidence that alludes to the possible effect of stereoscopy on perception of representations of three dimensional objects. For the future studies, it is critical to further investigate the role of depth cues in perception on spatial tasks. This particular future study could compare the original PSVT:ROT against a rendered PSVT:ROT that utilizes depth cues interested in analyzing. An alternate setup for this research would compare the original PSVT:ROT against a physical model setup that would require participants to look into a closed box through a set of optics designed to appear identical to the PSVT:ROT.

5.6. Summary

The objective of this study was to observe the effect of stereoscopy on spatial ability tests, and as demonstrated by the results obtained and the analysis on this data, binocular vision has no impact in testing scores of introductory manufacturing students enrolled in the spring semester 2011 class of CGT163 at Purdue University. However, the conflicting significance of the effect of stereoscopy on participants with low spatial ability indicates a need for further study.

LIST OF REFERENCES

LIST OF REFERENCES

Aitsiselmi, Y., & Holliman, N.S. (2009) Using mental rotation to evaluate the benefits of stereoscopic displays. *Proceedings of SPIE*, Vol. 7237.

Arnheim, R. (1986). *New Essays on the Psychology of Art*. Berkeley, California: University of California Press.

Autodesk (2008) Stereoscopic Filmmaking Whitepaper: *The business and Technology of Stereoscopic Filmmaking*. Retrieved 4-20-2011 from images.autodesk.com/adsk/files/stereoscopic_whitepaper_final08.pdf

Bodner, G. M., & McMillen, T. L. B. (1986). Cognitive restructuring as an early stage in problem solving. *Journal of Research in Science Teaching.* 23, 727-737

Campbel,l D.T., & Stanley, J.C., (1966). *Experimental and quazi-experimental designs for research*. Chicago, IL: Rand McNally

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambrige University Press.

CEEB *Special Aptitude Test in Spatial Relations*, developed by the College Entrance Examination Board, USA, (1939).

Einhauser, W., Martin, K., & Konig, P. (2004) Are switches in perception of the Necker cube related to eye position?, *European Journal of Neuroscience* Vol 20, Issue 10, 2811-2818, Nov.

El Koussy, A. A. H. (1935). The visual perception of space. *British Journal of Psychology*, 20, 1-80.

Eliot, J., & Smith, I.M. (1983). An International Dictionary of Spatial Tests. Windsor, UK: *The NFER-Nelson Publishing Company*, Ltd

Gittler, G., & Glueck, J. (1998). Differential transfer of learning: Effects of instruction in descriptive geometry on spatial test performance. *Journal for Geometry and Graphics*, 2(1), 71-84.

Guay, R. (1976), *Purdue Spatial Visualization Tests*; Purdue Research Foundation: West Lafayette, IN

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition: A review. In Stea, B., and Downs, R. (Eds.) *Image and Environment*, Chicago: University of Chicago Press: 226-234.

Hoffman, D. (1998) *Visal Intelligence: How we create what we see*, Norton & Company, Inc

Jensen, J. J. (1986). The impact of computer graphics on instruction in engineering graphics. *Engineering Design Graphics Journal*, 50 (2), pp. 24-33

Jianping, Y. (2007). Spatial Visualization by Isometric Drawing. *Engineering Design Graphics Journal*, v71, n2, 5-19

Jianping, Y. (2008). Spatial Visualization by Realistic 3D views. *Engineering Design Graphics Journal*, v72, 28-138

Kaufmann, H. (2000), Construct3D: A Virtual Reality Application for Mathematics and Geometry Education, *Education and Information Technologies*, Springer Netherlands, Volume 5, Number 4, December, 263-276

Kelly, T. L. (1928). *Crossroads in the mind of man*. Stanford, CA: Stanford University Press.

Leroy, L., Fuchs, P., Paljie, A. & Moreau, G. (2009). Some experiments about shape perception in stereoscopic displays. *Proceedings of SPIE*

Liben, L., Patterson, A. & Newcombe, N. (1981) *Spatial representation and behaviour across the life span: theory and application*. New York. Academic Press.

Linn, M. C., & Peterson, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6)

Lohman, D.F., & Kyllonen, P.C. (1983). Individual differences in solution strategy on spatial tasks. In Dillon, R.F. & Schmek, R.R. (Eds.), *Individual Differences in Cognition*, Academic Press, New York, pp.3-36

McGee, M.G. (1979). *Human spatial Abilities, Sources of Sex Differences*. New York: Praeger

McMillan, J. H., & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.). New York: Longman.

Miller, C. (1992), *The effectiveness of real- and computer-generated models to advance the spatial abilities of visual/haptic engineering students*. Doctoral dissertation, Ohio State University

Moffat, M., Madan, C., Cheng-Tse, C., Gu, J., Hartmut, S., & Raphael, G. (2002), Evaluation of Spatial Abilities through Tabletop AR, *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, Hamilton, New Zealand. Pages: 17 - 24

Moore, D. S., McCabe, G. P., & Craig, B. A. (2009). *Introduction to the Practice of Statistics*. New York, NY: W. H. Freeman and Company.

Necker, L. A., (1832), Observations on some remarkable optical phaenomena seen in Switzerland; and on an optical phaenomenon which occurs on viewing a figure of a crystal or geometrical solid, *London and Edinburgh Philosophical Magazine and Journal of Science*, Vol. 1, No. 5, pp. 329–337

Olson, D. R., & Bialystok, E. (1983). *Spatial Cognition: The Structure and Development of Mental Representation of Spatial Relations*. London: Lawrence Erlbaum Associates.

Örnkloo, H., & Von Hofsten, C (2009) Young children's ability to solve spatial problems involving a choice, *European Journal of Developmental Psychology*, 1740-5610, Volume 6, Issue 6, First published 2009, pp. 685 – 704

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A Redrawn Vandenberg & Kuse Mental Rotations Test: Different Versions and Factors that affect Performance. *Brain and Cognition*, 28, 39-58.

Pinker, S. (1997). *The Mind's Eye*. In How the Mind Works ISBN 0-393-31848-6

Schwartz, S. (2010), *Visual Perception: A clinical orientation*, Fourth edition, McGraw Hill Medical.

Sorby, S. A (1999) Developing 3-D Spatial Visualization Skills, *Engineering Design Graphics Journal*, Vol 63, No 2 pg 21-32

StatSoft, Inc. (2011). Electronic Statistics Textbook. Tulsa, OK: StatSoft. WEB: http://www.statsoft.com/textbook/

Study, N. (2001), *The Effectiveness of Using the Successive Perception Test I to measure Visual-Haptic Tendencies in Engineering Students*. Doctoral Dissertation, Purdue University, Indiana

Study, N. (2002), The Effectiveness of Using the Successive Perception Test I to measure Visual-Haptic Tendencies in Engineering Students. *Engineering Design Graphics Journal*, v66, n2, 17-26, Spring.

Teittinen, M. (2011), *Depth Cues in the Human Visual System*, Retrieved 4-22-2011 from http://www.hitl.washington.edu/scivw/EVE/III.A.1.c.DepthCues.html

Thurstone, L.L. (1950). *Some primary abilities in visual thinking*. Chicago, IL: University of Chicago Psychometric Lab Report No. 59.

Tsutsumi, E., Shiina, K., Suzaki, A., Yamanouchi, K., Saito, T., & Suzuki, K. (1999). A mental cutting test on female students using a stereographic system. *Journal for Geometry and Graphics, 3*(1), 111-119.

Woods, A. T., & Newell, F. N. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *Journal Of Physiology Paris*, *98*(1-3), 147-159. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15477029

Vandenberg, S. G., & Kuse, A.R. (1978). Mental rotations, a group test of three dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599-604.