THE EFFECTS OF STIMULI TYPE, SIZE, AND ARTICULATION ON
ESTIMATING THE LUMINANCE CONTRAST OF REAL AND CRT TARGETS

BY

MATTHEW R. RIDDLE

A Thesis Submitted to the Graduate Faculty of

WAKE FOREST UNIVERSITY GRADUATE SCHOOL OF ARTS AND SCIENCES

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF ARTS

Psychology

May 2011

Winston-Salem, North Carolina

Approved By:

James A. Schirillo, Ph.D., Advisor

Terry D. Blumenthal, Ph.D., Chair

Wayne E. Pratt, Ph.D.

Wayne L. Silver, Ph.D.
ACKNOWLEDGMENTS

There are several people that I need to thank for making my thesis possible. First and foremost, thank you to my advisor, Jim Schirillo, for all the time and effort he put into our discussions and providing feedback on all my drafts. Thank you for taking time to help me work through creating stimuli, analyzing data, and help developing my writing style. Throughout this process, you have helped me grow as a researcher and writer.

Thank you to Terry Blumenthal, Wayne Pratt, and Wayne Silver for being on my thesis committee. I appreciate all of your feedback and insights during the thesis process, which have been very useful in the development of my experiments, as well as in the writing of the final paper. Also, thank you to Jeff Muday for programming the CRT, which was constructed perfectly not only for my thesis, but for future experiments as well.

Last, but certainly not least, thank you to Rumi Tokunaga and Sasha Logvinenko, whose insights and assistance with data analysis early in my thesis process was absolutely invaluable.
TABLE OF CONTENTS

ACKNOWLEDGMENTS .............................................................................................................. ii
LIST OF TABLES AND FIGURES ....................................................................................... vi
LIST OF APPENDICES ........................................................................................................ viii
ABSTRACT .......................................................................................................................... ix
INTRODUCTION .................................................................................................................. 1
  Effects of stimuli type ....................................................................................................... 5
    Estimating lightness ....................................................................................................... 5
    The Illumination-Estimation Hypothesis ..................................................................... 5
  Estimating illumination and lightness with real surfaces and light ....................... 6
    Perceptual scission ....................................................................................................... 6
    Edge substitution experiment ...................................................................................... 8
  Estimating luminance contrast on a CRT monitor .................................................... 10
  Effects of stimuli size .................................................................................................... 10
    Anchoring theory ......................................................................................................... 10
      The anchor – relative area ......................................................................................... 11
  Effects of articulation ..................................................................................................... 11
Pilot data .............................................................................................................................. 12
Present studies .................................................................................................................... 13
Experiment one ................................................................................................................... 13
Experiment two .................................................................................................................. 14
METHOD: EXPERIMENT ONE ......................................................................................... 15
Hypothesis one: observers are inaccurate and variable ......................................

Hypothesis two: observer underestimate increments and overestimate decrements ..............................................................................................................

Hypothesis three: matched paper ratios are more accurate than matched spotlight ratios .............................................................................................................

Hypothesis four: Gilchrist’s area rule predicts that small incremental and decremental paper and spotlight matched ratios should be less veridical than their corresponding large target matched ratios found in pilot data .............

Small incremental matched ratios should be less accurate than large matched ratios ..............................................................................................................

Small decremental matched ratios should be less accurate than large matched ratios ..............................................................................................................
Hypothesis three: according the area rule small targets should be perceived differently than large targets 

Hypothesis four: articulated stimuli should be more veridical than non-articulated stimuli

CROSS-EXPERIMENT RESULTS

Increments

Decrements

GENERAL DISCUSSION

Effects of stimuli type on luminance contrast estimates

Effects of size on luminance contrast estimates

Effects of articulation on luminance contrast estimates

CONCLUSION

Effects of stimuli type

Effects of stimuli size

Effects of articulation

Limitations

Implications

REFERENCES

APPENDIX A

APPENDIX B

APPENDIX C

APPENDIX D

CIRRICULUM VITAE
### LIST OF TABLES AND FIGURES

#### TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Average matched ratios for experiment one</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>Average lightness and contrast matched ratios</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>Average matched ratios for experiment two</td>
<td>37</td>
</tr>
</tbody>
</table>

#### FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Illustration of how illumination times reflectance equals luminance</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>A Mondrian stimulus</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Illustration of the effects of area on lightness</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>A sample of stimuli used in Riddle et al. (2010)</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>Experiment one stimuli</td>
<td>16</td>
</tr>
<tr>
<td>6a</td>
<td>Histogram of incremental paper and spotlight matched ratio distribution</td>
<td>22</td>
</tr>
<tr>
<td>6b</td>
<td>Histogram of decremental paper and spotlight matched ratio distribution</td>
<td>22</td>
</tr>
<tr>
<td>7a</td>
<td>Matched incremental paper and spotlight ratios for all observers</td>
<td>24</td>
</tr>
<tr>
<td>7b</td>
<td>Matched decremental paper and spotlight ratios for all observers</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>Experiment two stimuli</td>
<td>33</td>
</tr>
<tr>
<td>9a</td>
<td>Histogram of incremental (non-articulation) matched ratios distribution</td>
<td>40</td>
</tr>
<tr>
<td>9b</td>
<td>Histogram of incremental (articulation) matched ratios distribution</td>
<td>40</td>
</tr>
<tr>
<td>9c</td>
<td>Histogram of articulated and non-articulated matched ratios distribution</td>
<td>40</td>
</tr>
<tr>
<td>9d</td>
<td>Histogram of all incremental matched ratios distribution</td>
<td>40</td>
</tr>
<tr>
<td>10a</td>
<td>Histogram of decremental (non-articulation) matched ratios distribution</td>
<td>41</td>
</tr>
<tr>
<td>10b</td>
<td>Histogram of decremental (articulation) matched ratios distribution</td>
<td>41</td>
</tr>
</tbody>
</table>
10c Histogram of articulated and non-articulated matched ratios distribution...41
10d Histogram of all decremental matched ratios distribution.........................41
11a Matched incremental small and large matched ratios for all observers......44
11b Matched decremental small and large matched ratios for all observers ....44
12a Histograms of CRT and spotlight incremental matched ratios
distributions (small, non-articulated) ..........................................................52
12b Histograms of CRT and spotlight incremental matched ratios
distributions (large, non-articulated).............................................................52
12c Histograms of CRT and spotlight incremental matched ratios
distributions (large, articulated) ....................................................................52
13a Histograms of CRT and paper incremental matched ratios
distributions (small, non-articulated) ...........................................................54
13b Histograms of CRT and paper incremental matched ratios
distributions (large, non-articulated) ............................................................54
14a Histograms of CRT and spotlight decremental matched ratios
distributions (small, non-articulated) ..........................................................55
14b Histograms of CRT and spotlight decremental matched ratios
distributions (large, non-articulated) ............................................................55
14c Histograms of CRT and spotlight decremental matched ratios
distributions (large, articulated) ..................................................................55
15a Histograms of CRT and paper decremental matched ratios
distributions (small, non-articulated) ............................................................56
15b Histograms of CRT and paper decremental matched ratios
distributions (large, non-articulated) ............................................................56
16 Inaccuracies of spotlight, paper, and CRT stimuli.................................59
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>APPENDIX</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Experiment room pictures</td>
<td>72</td>
</tr>
<tr>
<td>B Experiment one instructions</td>
<td>73</td>
</tr>
<tr>
<td>C Experiment two instructions</td>
<td>74</td>
</tr>
<tr>
<td>D Glossary</td>
<td>75</td>
</tr>
</tbody>
</table>
Matthew R. Riddle

THE EFFECTS OF STIMULI TYPE, SIZE, AND ARTICULATION ON
ESTIMATING THE LUMINANCE CONTRAST OF REAL AND CRT TARGETS

Thesis under the direction of James A. Schirillo, Ph.D., Professor of Psychology, Wake Forest University.

The goal of this thesis was to examine how luminance contrast estimates were affected by varying conditions of stimuli type, size, and articulation. Based on previous research, it was expected that lightness constancy would increase as stimuli moved from a spotlight, to paper, and finally to a CRT monitor. Lightness constancy was also expected to increase with articulation, and to vary as stimuli sized increased or decreased. Two separate experiments were conducted with a total of 80 college students from Wake Forest University. Observers were asked to make lightness and contrast matches of incremental and decremental spotlight, paper, and CRT targets, all of which were either small or large, and articulated or non-articulated. Results showed a strong effect of stimuli type in the predicted direction. Lightness constancy was lowest when matching spotlight targets, and highest when matching CRT targets. Effects of stimuli size were ambiguous. In experiment one, size affected lightness constancy only for decremental targets, not incremental targets; whereas in experiment two, size affected lightness constancy only for incremental targets, not decremental targets. Lastly, no effects of articulation were found.
INTRODUCTION

One question regarding the human visual system that has gone unanswered for over a century is under which conditions, and why, lightness constancy is either achieved or fails. Lightness constancy is the ability to have a surface appear to reflect the same proportion of light independent of the level of illumination falling upon it (see Appendix D: Glossary). This is difficult to discern in that lightness (or perceived reflectance) must be computed from reflected light, or luminance. However, luminance is a product of both the reflectance of a surface and the level of illumination falling upon that surface. Thus, how does the visual system distinguish what proportion of luminance should be dedicated to lightness and what part to illumination?

This problem can be expressed by the following equation

\[ I \times R = L \]

where \( I \) is the illumination falling onto a surface from some light source(s), \( R \) is the reflectance of a given surface, and \( L \) stands for luminance (Figure 1). These physical terms can also be expressed in perceptual terms by the following equation

\[ I_p \times L_p = B \]

where \( I_p \) is the perception of the illumination, \( L_p \) is lightness, or perceived reflectance, and \( B \) stands for brightness, or perceived luminance.

The study of lightness perception has several important broad applications. For instance, how can technology incorporate visual perception
Figure 1. Illustration of how illumination times reflectance equals luminance. The light that reaches the eye (luminance) is a function of illumination (light source) and the percentage of that illumination that is reflected off the paper surface (reflectance). Notice that the light from the sun (designated by the down arrow) is more intense than the light reflected off the paper (designate by the up arrow) because a certain percentage of light is absorbed by the paper.

findings to make improvements in cameras, satellite imagery, devices for the visually impaired, etc? That is, many of these fields are devoted to removing shadows from scenes. Lightness perception provides a means of understanding how lightness, and indirectly, illumination works. Thus, many artists, graphic artists, etc., can also use findings from lightness perception studies to improve upon their work. For example, lightness is perceived reflectance, and therefore tells the observer something about the properties of the objects in view. Consequently, finding systematic errors in lightness perception or cases in which it is veridical is important for object recognition. For example, aircraft pilots must be able to clearly distinguish between many visual targets, all of which carry different meanings and warnings to the pilot. Also, within the field of artificial
intelligence, advanced computers and robots will need to be able to “see” like a human, in order to fully mimic one.

The most immediate application of this thesis concerns current methodology. Approximately twenty years ago it became common to use CRT monitors in lightness perception experiments, given the convenience and high control it lends to a researcher. However, many researchers continued to use more traditional paper and illuminated stimuli. Blakeslee, Reetz, and McCourt, (2008) suggest lightness judgments can vary based on different conditions, making it difficult to compare one experimenter’s results with another’s. It is important to understand more about differences between popular experimental conditions within the field in order to ensure researchers appreciate the implications of choosing certain experimental conditions over others.

This thesis focuses on the perception of illumination, lightness perception, and what is often felt to underlie the latter; namely luminance contrast. Three separate effects that may influence the accuracy of luminance contrast estimates are investigated.

The first effect is the type of stimulus display. Whether a target consists of reflected light, a piece of paper, or light emanating from a CRT screen can impact one’s perception of that target. Previous CRT research suggests that observers can successfully estimate lightness and luminance contrast (Arend & Spehar, 1993). In one such experiment, observers were asked to vary the luminance level of a test patch within a Mondrian stimulus to match a standard patch in another Mondrian (see Figure 2 for an example of a Mondrian stimulus).
Observers were able to make veridical lightness and luminance contrast matches. However, these results have not been replicated using real light and real surfaces (Logvinenko, 2005).

![Figure 2. A Mondrian stimulus, a patchwork of gray rectangles of varying luminances.](image)

Second, Gilchrist et al. (1999) suggests that the size of stimuli can also affect one's lightness perception. Viewing a simple two-part pattern consisting of a light gray and a dark gray surface, observers’ luminance contrast estimates changed as a function of size. The current study examined size effects across a number of different conditions.

Third, the presence of articulation is known to assist in making accurate luminance contrast judgments (Gilchrist, 1994; Gilchrist et al., 1999; Logvinenko, 2005). Articulation is defined as the presence of multiple shades of gray surfaces that can be used as a comparison in judging the lightness of a given surface. Stimuli with articulation are compared to stimuli without articulation across conditions of varying size and stimuli type.
Effects of stimuli type

During the first several decades of lightness perception research, it was a commonly held notion that the visual system determines lightness values based on absolute luminance values and ratios (Gilchrist, 1994; Gilchrist et al., 1999). If this were all that mattered, the type of stimuli used would be irrelevant. All that would matter is how much light reached the retina. However, this notion has long since been rejected, in part because of known effects of stimuli type. This thesis is concerned with three kinds of stimuli: illuminated, paper, and CRT targets.

Estimating lightness

*The Illumination-Estimation Hypothesis*

Before examining the effects of individual stimuli, one first must have an understanding of the complexity and difficulty of the problem stated in the beginning of the introduction (see Figure 1). An intuitive solution is that the visual system can determine lightnesses of surfaces within a scene by first estimating the illuminant. Once the illuminant is known one can compute the lightness of a surface via the Illumination-Estimation Hypothesis (Rutherford & Brainard, 2002). Rutherford and Brainard had observers view two chambers of different illuminations that also contained several objects with surfaces of different reflectances. The first chamber consisted of medium-gray reflectances, and the second chamber consisted of darker-gray reflectances. Observers would first adjust the illumination of one chamber to match the illumination of the second chamber. Observers would then adjust the luminance of a test patch in one chamber to match the surface lightness of a comparable patch in the other
chamber. In effect, observers were making both perceived illumination and lightness matches. If the Illumination-Estimation Hypothesis is correct, the luminances of the two chambers should be equal after both matches are completed.

However, the results of this study contradicted the Illumination-Estimation Hypothesis. First, the illumination judgments were inaccurate because of the difference in reflectances within each chamber – the observers were unable to successfully adjust the illumination in the test chamber to match the illumination in the standard chamber. Secondly, like the illumination matches, the lightness matches also failed veridicality; the observers were unable to accurately match the test patch to the standard patch in lightness.

**Estimating illumination and lightness with real surfaces and light**

The vast majority of experiments (including the experiments in this thesis) include at least one illuminant which requires an accurate approximation. The literature reviewed thus far assumes that veridical judgments of perceived illumination should be highly correlated with veridical lightness judgments. However, this is often not the case.

*Perceptual scission*

Errors of lightness perception might be explained if perceived illumination is in fact *not* correlated with lightness. Lightness might be determined by certain physical characteristics of a scene, such as luminance contrast; while perceived illumination might be determined by a different set of physical characteristics, such as the maximum luminance of a region within the scene. This is known as
perceptual scission, or the separation of the lightness of a surface from the illumination falling onto that surface. Oyama (1968) tested this notion using a display made of three boxes. A standard box with a middle-gray paper was used to evaluate two comparison boxes, one composed of black paper and the other composed of white paper. Within each box could be five different possible levels of illumination as well as five disks with different reflectances, so that within each box there were 25 possible different combinations of reflectance and illumination. Observers made both lightness and illumination judgments.

The results of this experiment indicated that observers did not have perfect lightness constancy when making their judgments. Instead, the matched reflectances of observers’ judgments slightly increased proportional to the standard reflectance and independent of changes in illumination. These results suggest that the physical determinants of lightness are separate and independent of the physical determinants for perceived illumination, which allows for perceptual scission (Noguchi & Kozaki, 1985).

Lightness, according to Oyama (1968) and Kozaki (1973), was most strongly correlated with the relative luminances of a standard disk and background; whereas the determinant for perceived illumination appeared to be the highest luminance present within the field.

There may be additional stimulus correlations that distinguish lightness from perceived illumination, such as different kinds of materials and textures (i.e., flannel, wood, etc.) (Beck, 1961). If lightness and illumination are indeed coupled together, lightness constancy should fail when judgments of illumination are
equal to the luminance of the region. But this was not the case; illumination judgments are inaccurate, suggesting that lightness and perceived illumination are correlated with different aspects of the physical characteristics of a given stimulus.

*Edge substitution experiment*

To consider the difference between illumination and lightness further, Gilchrist, Delman, and Jacobsen (1983) had observers view a display made of two background papers, each with a target in the middle. Observers used a Munsell chart to match on separate trials both the level of illumination and reflectances of the targets. The display could have two conditions. The first was a contrast condition, in which observers viewed the two equally gray targets, one on a black background and the other on a white background. This resulted in a standard simultaneous contrast effect. The second condition was a lightness constancy condition, in which the edge between the two backgrounds was made to appear as if it were an illumination edge as opposed to a reflectance edge.

The results indicated that the different appearance of the background edges in the two conditions had indeed affected the observer’s perception of the targets. In the contrast condition, the target squares were both seen as middle gray and the overall illumination level was perceived as being uniform. However, in the constancy condition, the target squares appeared either white when on the black background or black when on the white background, and the level of illumination was also perceived differently depending on the background. The white background appeared to be under a high level of illumination, while the
illumination appeared to be in shadow for the black background. Thus, although the luminance of the targets and their backgrounds were equal in both conditions, the lightness of those targets differed significantly. This depended upon whether the background edge was perceived either as an illumination or a reflectance edge.

Additionally, errors of luminance contrast often occur in predicted directions. Specifically, estimates of incremental targets tend to be underestimated, while estimates of decremental targets tend to be overestimated (Whittle, 1986).

To summarize, not only does evidence suggest that observers are unable to distinguish between reflectance and illumination edges (Gilchrist et al., 1983), perceived illumination and lightness may in fact be determined by a different set of physical characteristics (Oyama, 1968; Kozaki, 1973; Kozaki & Noguchi, 1976). Additionally, these errors of luminance contrast judgments are not trivial in size. When matching a spotlight target, observers underestimated the physical ratio of the target by approximately 35%. According to Logvinenko (2005), this may be because of partial illumination discounting, rather than a complete discounting of illumination. In other words, the more an illuminant is discounted (i.e., correctly estimated and removed from the luminance equation (see Figure 1)), the more accurate lightness constancy will become. But as these studies indicate, accurate illumination discounting is not something that is readily achieved. However, all of these studies used real light and surfaces, while the same results may not occur using other stimuli, namely CRT monitors.
Estimating luminance contrast on a CRT monitor

In contrast to the above results, there have been instances where observers successfully estimated the lightnesses of a scene. Viewing Mondrian displays on a CRT monitor, Arend and Spehar (1993) found that observers were able to estimate luminance contrasts accurately. In these experiments, Arend and Spehar presented a test and match Mondrian that differed in luminance so that it appeared that the match Mondrian was under a different level of illumination than the test Mondrian. They then varied the luminance of a target area within the test Mondrian, creating the illusion of a reflectance change of a given surface under constant illumination. Under these conditions, observers’ lightness matches and luminance contrast matches between the test and match Mondrian were veridical. Thus, it appears that the accuracy of luminance contrasts is at least partially a function of the type of stimuli.

Effects of stimuli size

Anchoring theory

Gilchrist (1994) showed that observers are able to successfully judge the lightnesses of surfaces across a wide range of illumination conditions. The problem, Gilchrist et al. (1999) argues, is to determine how the visual system transforms a luminance-ratio value into a percept that is recognizable to an observer – are two surfaces dark-gray and mid-gray, or mid-gray and light-gray, etc.? An anchoring rule may help answer this question.
The anchor – relative area

Gilchrist et al.’s (1999) anchoring model states that lightness perception depends on relative luminance and relative area. This thesis will be concerned with relative area. Figure 3 illustrates how an increase in one area relative to another area in the scene can increase the lightness of the former area. This is known as the area rule, which simply states that as the darker region increases in size, it will also increase in lightness.

Figure 3. Illustration of the effects of area on lightness. As darker regions increase in area, they will appear lighter.

Effects of articulation

Logvinenko (2005), as well as Maloney and Schirillo (2002), suggest that other factors, such as articulation, may affect how luminance contrast is estimated. Articulation can be defined as the presence of multiple shades of gray surfaces that can be used as a comparison in judging the lightness of a given surface. In essence, the observers are able to make local luminance contrast estimates across articulated patches which assist when computing the perceived luminance contrasts of a target. In Logvinenko’s articulated stimuli, results indicated an increase in illumination discounting which lead to increased
lightness constancy, although perfect lightness constancy was still not achieved. Gilchrist et al. (1999) also describes two experiments in which articulation resulted in increased lightness constancy using the traditional stimulus of a black background with articulated squares, as well as an articulated Mondrian display. Arend and Spehar (1993) also found an effect of articulation. Thus, there is evidence that articulation can improve luminance contrast judgments on spotlight, paper, and CRT stimuli.

Pilot data

In Riddle, Schirillo, Tokunaga, and Logvinenko (2010), observers matched a crisp luminance edge produced by either a spotlight or a reflectance edge to an array of reflectance edges in a typical room setting (Figure 4). They conducted two separate experiments. In the first, observers viewed spotlights, both decrements and increments, half of which were articulated, half of which had no articulation, and were asked to make luminance contrast estimates. In the second experiment, in addition to the spotlight condition, observers were also asked to make luminance contrast matches of a large gray paper target (with the same size and similar luminance of the spotlight). Overall, observers were unable to accurately estimate luminance contrast. These findings suggested that Arend and Spehar's results may have been a result of viewing a CRT screen, as opposed to real objects and real light.
Present studies

Experiment one

Experiment one was designed to investigate possible effects between paper and spotlight stimuli. In addition to stimuli type effects, the small target results in experiment one were compared to the large target results found in the pilot studies in order to test stimuli size effects. Four hypotheses will be evaluated. The first three hypotheses relate to stimuli type effects, and the fourth hypothesis relates to size effects. First, observers are expected to make inaccurate and variable matches. Previous research suggests that observers are not able to make veridical matches under conditions using real light and surfaces (Beck, 1961; Gilchrist et al., 1983; Kozaki, 1973; Kozaki & Noguchi, 1976; Logvinenko, 2005; Noguchi & Kozaki, 1985; Oyama, 1968; Riddle et al., 2010; Rutherford & Brainard, 2002). Second, lightness constancy errors should underestimate increments and overestimate decrements (Riddle et al., 2010; Whittle, 1986). Third, the average matched paper ratio should be more veridical than the average matched spotlight ratio, given the illumination discounting.
theory put forth by Logvinenko (2005). Fourth, according to the area rule (Gilchrist et al., 1999), the small target matches made in experiment one should be statistically different from large target matches made in the second pilot study.

**Experiment two**

Experiment two was designed to further test stimuli type by using CRT stimuli, instead of papers and spotlights. Experiment two also investigates effects of stimuli size (small vs. large targets), and articulation effects. There are again four hypotheses. First, compared to experiment one, observers are expected to make veridical matches when viewing stimuli on a CRT monitor (Arend & Spehar, 1993). Second, any errors that are present are still expected to be in the direction predicted by Whittle (1986) and Riddle et al. (2010). In particular, incremental errors should underestimate the physical ratio and decremental errors should overestimate the physical ratio. Third, according to the area rule (Gilchrist et al., 1999), the small target matches should be statistically different from large target matches. Fourth, according to previous studies (Gilchrist et al., 1999; Logvinenko, 2005), observers should make more veridical matches when viewing articulated displays than non-articulated displays.

Lastly, looking across experiments one and two, as well as pilot studies, three important questions should have greater clarification. First, in regards to effects of stimuli type, do observers perceive spotlight, paper, and CRT displays differently? Secondly, in regards to effects of stimuli size, do observers perceive small and large targets differently? Third, in regards to articulation effects, do observers perceive articulated displays differently than non-articulated displays?
METHOD: EXPERIMENT ONE

Subjects

All procedures have been approved by the Institutional Review Board of Wake Forest University and were performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Forty observers were recruited from an introductory psychology class at Wake Forest University to earn course credit (18 males, 22 females, and aged 18 – 22). Normal or corrected-to-normal vision was required for participation. Observers with color blindness were not allowed to sign up for experiment participation.

Stimuli

Observers were presented with sheets of either black or white background papers illuminated by a spotlight produced by a projector, or with a gray paper attached to the background paper (see Figure 5). The luminance ratio produced by the spotlight was 4.60:1 for increments and 3.64:1 for decrements. The luminance ratio for the edge produced by the gray paper was 8.53:1 for increments, and 2.78:1 for decrements. The luminance measurements were made with a Minolta (CS-100) photometer. The size of each matching patch printed on the 10 reflectance papers, as well as the four targets (two gray papers, increment spotlight, and decrement spotlight) was 0.72° X 0.72° of visual angle. Thus, the target papers and spotlights were equal in size to the matching reflectance patches. Given that the observer was placed next to the carousel
projector using a chinrest, it was clear that the spotlight was produced by a real light source contained within the room and that the gray papers were not specially illuminated, but were reflectance surfaces (see Appendix A for a picture of the room). Both the spotlight and gray paper conditions were viewed by observers as either increments or decrements.

![Image of experimental stimuli](image-url)

**Figure 5.** Experiment one stimuli.

**Procedure**

Observers were asked to estimate both the contrast and lightness of the spotlight and gray paper in separate sessions, determined at random. It was decided to measure both contrast and lightness specifically to ensure observers were able to judge the spotlight targets “reflectance” qualities. Each session lasted approximately 15-20 minutes.

Five different top arrays of paper were used for each of the incremental and decremental conditions, presented in random order. Each condition received two judgments determined at random, one for the top-right 20 squares and one for the top-left 20 squares.

In one session the 40 naïve observers matched the luminance contrast (see Appendix B for instructions) of the bottom test patch to two of 40 different
reflectance edges. In a second session the same 40 observers matched the lightness (see Appendix B for instructions) of the region lit by the spotlight or gray paper to two of the same 40 reflectance edges. Thus, all observers made two judgments per individual stimulus configuration (increment versus decrement and spotlight versus real paper), for both a luminance contrast and lightness instruction condition. This resulted in a total of 80 judgments made by each observer.
A one-way ANOVA was employed to test the relationship between lightness and contrast matching for both increments and decrements. In addition to one-way ANOVAs, Kruskal-Wallis tests were also used as the distributions were not expected to be normal due to previous studies (Logvinenko, 2005). Kruskal-Wallis tests were more appropriate to test differences in lightness and contrast matches due to the smaller N, as well as the fact that the Kruskal-Wallis test ranks the data, unlike one-way ANOVAs. In the incremental gray paper conditions the average matched ratio was found to be 7.63:1 for lightness matches and 7.71:1 for contrast matches. In the incremental spotlight conditions the average matched ratio was found to be 5.44:1 for lightness matches and 5.49:1 for contrast matches. An independent-samples Kruskal-Wallis test revealed no statistical differences between incremental lightness and contrast conditions ($p = .601$). In the decremental gray paper conditions the average matched ratio was found to be 3.18:1 for lightness matches and 3.15:1 for contrast matches. In the decremental spotlight conditions the average matched ratio was found to be 4.53:1 for lightness matches and 4.72:1 for contrast matches. The independent-samples Kruskal-Wallis test revealed no statistical differences between decremental lightness and contrast conditions ($p = .625$). Because there were no significant statistical differences between lightness and contrast instructions the two conditions were combined for all subsequent
analyses. Therefore, the average combined ratios for incremental paper and spotlight matches were 7.67:1 and 5.46:1, respectively. The average combined ratios for decremental paper and spotlight matches were 3.16:1 and 4.62:1, respectively.

Table 1 shows the average matched ratio for each condition along with the amount of underestimation or overestimation that was observed compared to the physical ratio of the target stimuli. Regarding increments, paper matches (matched ratio of 7.67:1) were underestimated by 9.7%, whereas spotlight matches (matched ratio of 5.46:1) were overestimated by 18.8%. A planned comparison revealed a statistically significant difference between paper and spotlight ratios, \( t(3196) = 32.34, p < .001 \). However, because the physical ratios were not equivalent across these conditions (the physical ratios were 8.53:1 and 4.6:1 for paper and spotlight targets, respectively), this test is not particularly meaningful, as the matched ratios should be different from one another. Therefore, another planned comparison was performed using the percentage of over or underestimation from the physical ratio, rather than the average matched ratios. This test revealed a statistically significant difference between the paper and spotlight stimuli, \( t(3196) = -19.45, p < .001 \).

Regarding decrements, observers overestimated the paper stimuli by 13.8%, as well as spotlight stimuli by 27.1%. As with increments, because the decremental paper and spotlight physical ratios differed (2.78:1 and 3.64:1 for paper and spotlight, respectively), a planned comparison was performed using
the difference from the physical ratio (using percentages), which again revealed a significant difference, \( t(3196) = -8.97, p < .001 \).

Overall, observers were more accurate with paper stimuli than with spotlight stimuli (-9.7% and 13.8% for paper vs. 18.8% and 27.1% for spotlight). Also, increments were found to be more accurate than decrements across both paper and spotlight stimuli. On average, increments were found to be inaccurate by approximately 14.3%, whereas decrements were found to be inaccurate by approximately 20.5% \( (t(3196) = -15.29, p < .001 \). The most accurate condition was the incremental paper target (underestimation of 9.7%). The most inaccurate condition was the decremental spotlight (overestimation by 27.1%).

Table I

**Average matched ratios for experiment one**

<table>
<thead>
<tr>
<th></th>
<th>Increments</th>
<th>Decrements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of paper target = 8.53:1</td>
<td>Ratio of spotlight target = 4.6:1</td>
<td>Ratio of paper target = 2.78:1</td>
</tr>
<tr>
<td>Observed Ratio</td>
<td>Difference from Actual Ratio</td>
<td>Observed Ratio</td>
</tr>
<tr>
<td>Paper</td>
<td>7.67:1</td>
<td>-9.7%</td>
</tr>
<tr>
<td>Spotlight</td>
<td>5.46:1</td>
<td>18.8%</td>
</tr>
</tbody>
</table>

The incremental histogram shows the distributions of both matched ratios for the paper and spotlight targets (Figure 6a). The x-axis represents the average matched ratio, and the y-axis represents the total number of matches for a given
ratio. Paper matches are represented by blue bars and spotlight matches are represented by red bars. The blue (paper) and red (spotlight) arrows signify the physical ratios. One can clearly see that the paper distribution falls mostly below (i.e., underestimates) the blue arrow, or the actual physical ratio of the paper stimuli. The distribution for the spotlight stimuli is also slightly more variable compared to the paper distribution. Spotlight ratios ranged from 1.9:1 to 14.6:1 with a standard deviation of 1.75, whereas the paper ratio ranged from 2.3:1 to 18.2:1, with a standard deviation of 1.51.

Likewise, Figure 6b illustrates the distributions for decremental paper and spotlight matched ratios. Here, the two distributions almost overlap each other, but notice that the physical ratios are closer to each other than they are in the increment condition, thus one would expect the distributions to overlap if observers had perceived paper targets similarly to spotlight targets. Although both distributions fall above (i.e., overestimate) their respective physical ratios, the spotlight distribution is wider and overestimates its physical ratio by a greater amount than does the paper distribution. Observers were also more variable in their spotlight matches. Matched spotlight ratios ranged from 2.5:1 to 14.8:1 with a standard deviation of 1.35, whereas matched paper ratios ranged from 2.0:1 to 6.4:1 with a standard deviation of just 0.54. This aspect can also be seen descriptively by noticing the high and narrow peaks of paper matches. The vast majority of paper matches fell between 2.0:1 and 4.0:1, while spotlight matches peaked between 3.0:1 and 6.0:1.
Figure 6a. Histogram of incremental paper (blue bars) and spotlight (red bars) observed ratio distributions.

Figure 6b. Histogram of decremental paper (blue bars) and spotlight (red bars) observed ratio distributions.
Next, one-way ANOVAs were computed for each observers’ matches across both increments and decrements and were then plotted based on each individual’s average matched paper ratio and matched spotlight ratio, as shown in Figures 7a and 7b. Each green circle represents one observer’s matched ratios, where their matched paper ratio is plotted along the x-axis, and their matched spotlight ratio is plotted along the y-axis. A filled in circle indicates statistically different matched paper and spotlight ratios. A hollow circle indicates that the observer’s matched paper ratio did not statistically differ from their matched spotlight ratio. The blue line designates the actual physical ratio for paper targets and the red line designates the physical ratio for spotlight targets. Note that because the physical paper and spotlight ratios differed for both increments and decrements, whether or not a given observer’s ratios statistically differed from one another has little meaning. Even if an observer made all perfect veridical matches, their paper and spotlight ratios would still differ. Thus, the important aspects of these two graphs are where each observer falls in relation to the physical ratios. A perfect ratio (i.e., veridicality) would appear at the intersection of the red and blue line.

In Figure 7a, notice that almost every observer falls to the left of the actual paper ratio (blue line) and thus are underestimating paper targets; whereas in Figure 7b, with several exceptions, observers are falling to the right and thereby overestimating the paper target.
Regarding spotlight targets, both Figure 7a and 7b indicate a similar pattern, namely that almost every observer falls above the red line, indicating overestimation.

**Figure 7a.** Matched incremental paper and spotlight ratios for all observers.

**Figure 7b.** Matched decremental paper and spotlight ratios for all observers.
DISCUSSION

There were four hypotheses regarding experiment one. First, observers were expected to show a high degree of inaccuracy and variability in their estimates of both increments and decrements. Second, based on pilot data, observers would underestimate increments but overestimate decrements. Third, matched paper ratios should be more accurate than matched spotlight ratios. Fourth, according to Gilchrist et al.'s (1999) area rule, small incremental and decremental paper and spotlight matched ratios should be less veridical than their corresponding large target matched ratios found in pilot data.

**Hypothesis one: Observers are inaccurate and variable**

Observers were indeed quite variable and inaccurate in their matches. Figure 6a and 6b illustrate wide distributions with several large peaks that range from two to four units of matched ratios. This inconsistent and inaccurate matching by observers is in keeping with previous findings (Logvinenko, 2005; Riddle et al., 2010). Conversely, when using a CRT screen Arend and Spehar (1993) suggested observers are able to make veridical matches. Thus, there may be a perceptual difference between images on a CRT screen and images that are being projected on tangible objects, such as a piece of paper. When viewing a CRT screen, the light being projected is directly received by the retina; whereas when viewing a paper or spotlight the light source (i.e., a projector) does not have a direct pathway to the retina, but rather is being reflected off of another
surface. Imagine a light bulb that emits a certain amount of light, say 50 cd/m². Now imagine a white surface that has a reflectance of exactly 100%, meaning it will reflect the exact amount of light that falls upon it, and is placed directly adjacent to the light bulb. An observer's retina will not be able to tell the difference between the light bulb and the white surface, given that both are emitting 50 cd/m². However, according to Land (1983) the cortex plays a critical role in lightness perception. Therefore, it may be possible that the cortex computes the light that comes directly from a light source differently than it does for light that is reflected off of another surface. The differences between matched ratios of CRT and paper/spotlight targets will be investigated and discussed further in experiment two.

**Hypothesis two: Observers underestimate increments and overestimate decrements**

The predicted difference between increments and decrements was partially validated. Specifically, observers underestimated the incremental paper target while overestimating the incremental spotlight target and both decremental targets (see Table 1). It is difficult to elucidate why there was a discrepancy among incremental targets, although a plausible explanation for decremental targets does exist. According to Whittle (1986), observers typically overestimate decrements because the ratios of decrements tend to be small, in an absolute sense, compared to increments. The decremental ratios of this current experiment are in the range of 2.78:1 to 3.64:1. These ratios are quite small and any error will most likely fall above this range, rather than below. That is, there
are few underestimating errors that can be made, so that observer’s underestimating errors are capped, whereas errors of overestimation can potentially reach infinity.

However, one could posit that this phenomenon may hold true for increments as well. The incremental ratios in this experiment range from 4.6:1 to 8.53:1. So while it is true that there are certainly more errors of underestimation to be made compared to decrements, there are still more potential overestimating errors to be made for increments, just as there were for decrements. Observers actually did overestimate the incremental spotlight target by 18.7%. The underestimating witnessed in the incremental paper target (10.1% underestimation) could be the exception to the rule. However, given that pilot data and Logvinenko (2005) suggested that observers should underestimate increments, this conclusion is unlikely. As previously stated, there is a cap on errors of underestimation. The decremental targets were overestimated primarily because their physical ratios were small, as was the incremental spotlight target, where the physical ratio of this stimulus was relatively small at 4.60:1.

**Hypothesis three: Matched paper ratios are more accurate than matched spotlight ratios**

According to Gilchrist et al. (1983), observers will treat a reflectance edge differently than an illumination edge, even if both targets have identical luminance ratios to their backgrounds. Although pilot data (Riddle et al., 2010) suggested that observers treat paper and spotlight targets similarly, it was hypothesized in the current experiment that paper targets should be perceived
more veridically than spotlight targets. The results support this hypothesis. Observers were more accurate with paper stimuli than with spotlight stimuli (-9.7% and 13.8% for paper vs. 18.8% and 27.1% for spotlight).

In Logvinenko (2005) observers fail to completely discount an illuminant, and thereby misjudge the lightness of a surface. When viewing spotlight stimuli, an observer is forced to discount a greater amount of the illumination to make an accurate estimate. With paper stimuli, there is less illumination, and thus inaccurate discounting of the illuminant will not have as great of an effect on one’s estimates.

However, in the second pilot study (where observers viewed large paper and spotlight targets) observers’ paper matched ratios were only marginally more accurate than spotlight matched ratios (-24.2% and -2.1% for paper vs. -24.1% and 7.5% for spotlight). So why were observers more accurate with paper stimuli compared to spotlight stimuli in one experiment, yet not in a second? Gilchrist et al.’s (1999) area rule suggests stimuli size can affect lightness perception. As the target size decreased from the second pilot study to the current study, where targets equaled matching reflectance squares in size, it is possible that the change in target size resulted in a change in lightness constancy errors.

**Hypothesis four: Gilchrist’s area rule predicts that small incremental and decremental paper and spotlight matched ratios should be less veridical than their corresponding large target matched ratios found in pilot data.**

The area rule was also used to postulate that small targets would be less veridical than large targets; however, the data were ambiguous on this point. The
hypothesis was not supported for increments, but was supported for decrements and so the hypothesis will be discussed in two sections, one for increments and the second for decrements.

**Small incremental matched ratios should be less accurate than large matched ratios**

This hypothesis was not supported by the current results. Large incremental targets were underestimated by an average of 24.2% and small targets were near veridicality at an overestimation of 4.3%. Note that the darker region on incremental stimuli is actually the background, not the target. The incremental backgrounds increased in size from pilot study two to the current study; hence the background should have appeared lighter when the target was small. This should have the effect of decreasing the average matched ratio (note that the actual size of background paper did not change across experiments, but rather the target that covered the background paper decreased in size, thereby increasing the amount of background paper that was in sight for experiment one). To illustrate this further, take a large incremental target (hence a small background) with a physical ratio of 10:1 that is underestimated by observers by 25%, so that the matched ratio is 7.5:1. When the background is increased in size, thereby increasing its perceived reflectance, it should have the effect of increasing the right side of the ratio, which would result in underestimating the physical ratio even more than the 7.5:1 matched ratio. However, this is not what occurred, instead observers’ matched ratios increased and were much more accurate with their estimates.
There is no clear explanation as to why this occurred. In Gilchrist’s area rule experiments (Gilchrist et al., 1999) there were no discernible backgrounds and targets, but rather a two-part pattern painted on a single surface. In the current set of experiments, there is a piece of paper or spotlight that covers the background paper, so that the target clearly sits in front. Because observers were instructed to focus primarily on the target and not the background, the area rule may not have applied to this situation, as it was the background that should have undergone the perceptual change, not the target.

**Small decremental matched ratios should be less accurate than large matched ratios**

This hypothesis was supported by the current results. Large decremental targets were near veridicality at an overestimation of 2.7% and small targets were overestimated by 20.3%. Note that the darker region on decremental stimuli is now the target, unlike increments. The decremental targets decreased in size from pilot study two to the current study, and thus the reciprocal of the area rule should be used: as the darker region decreases in area, it decreases in lightness (i.e., becomes darker). Hence the target should have appeared darker when the target was small. This should have the effect of increasing the average matched ratio. To illustrate this further, take a large decremental target (with a relatively small background) with a physical ratio of 10:2 that achieves perfect constancy by observers, so that the average matched ratio is also 10:2. When the target is decreased in size, thereby decreasing its perceived reflectance, it should have the effect of decreasing the right side of the ratio (note that here, unlike
increments, the right side of the ratio represents the target and the left side represents the background), which would result in overestimating the physical ratio. This is what happened.
METHOD: EXPERIMENT TWO

Subjects

All procedures were approved by the Institutional Review Board of Wake Forest University and have been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Forty different observers were recruited from an introductory psychology class at Wake Forest University to earn course credit (19 males, 21 females, and aged 18 – 21). Normal or corrected-to-normal vision was required for participation. Observers with color blindness were not allowed to sign up for experiment participation.

Stimuli

Observers were presented with images of either black or white backgrounds on a CRT monitor with four large targets and four small targets (see Figure 8). Stimuli were generated by a Dell Dimension 9200 and presented on an accurately calibrated Radius Press-View 17SR 17-in. color monitor. The 832 x 624 pixel screen produced achromatic stimuli at CIE chromaticity \( x = 0.27, y = 0.28 \). The scan rate was 75 Hz non inter-laced. The chromaticity of each phosphor was measured spectroradiometrically. The red, green, and blue guns were linearized by use of an 8-bit lookup table. Luminance was approximately constant (~3%) within the central region of the screen that displayed the test and comparison patterns.
Articulation was present on two large targets (one increment and one decrement) as well as two small targets (one increment and one decrement). On the other two large targets and two small targets, no articulation was present. The luminance ratio produced by increments was 8.98:1 and 3.24:1 for decrements. The luminance measurements were recorded with a Minolta (CS-100) photometer. The size of each matching square, each articulation square, and the four small targets was 1.99° x 1.99° of visual angle. The size of the four large targets was 13.43° x 10.20° of visual angle. Thus, the small targets are equal in size to the matching reflectance squares, but not for the large targets. The sizes of all the targets and matching reflectance squares are approximately proportional to those used in experiment one.

![Figure 8. Experiment two stimuli.](image)

**Procedure**

Observers were asked to estimate both the contrast and lightness of the small and large targets in separate sessions, determined at random. A chin rest
was used in order to ensure each observer was of equal distance away from the
monitor (see Appendix A for a picture of the room). Each session took
approximately 20 minutes.

Five different top arrays consisting of 40 small matching squares were
used for each of the incremental and decremental conditions, presented in
random order. Each condition received two judgments determined at random,
one for the top-right 20 squares and one for the top-left 20 squares. In addition,
before each new stimulus page, observers viewed an adaptation screen for
approximately three seconds. The luminance of each adaptation screen was
computed using the average luminance of the previous page’s 40 matching
squares, which always resulted in a middle gray appearance.

In one session the 40 naïve observers matched the luminance contrast
(see Appendix C for instructions) of the bottom test patch to two of 40 different
matching squares. In a second session the same 40 observers matched the
lightness (see Appendix C for instructions) of the bottom test patch to two of the
same 40 matching squares. Thus, all observers made two judgments per
individual stimulus configuration (increment versus decrement and large versus
small target patch), for both a luminance contrast and lightness instruction
condition. This resulted in a total of 160 judgments made by each observer in
each session.
A one-way ANOVA was employed to test the relationship between lightness and contrast matching for both increments and decrements. In addition to one-way ANOVAs, Kruskal-Wallis tests were also used as the distributions were not expected to be normal due to previous studies (Logvinenko, 2005). Table 2 lists the average ratios for all increment and decrement conditions. For example, when making lightness matches on small incremental targets, observers’ average matched ratio was 8.68:1. When making contrast matches for the same stimuli, observers’ average matched ratio was 8.57:1. All tests, using both one-way ANOVAs and Kruskal-Wallis tests revealed no statistically significant differences between lightness and contrast matches of the same stimuli. Combining all stimuli together (i.e., all small and large targets) a Kruskal-Wallis test indicated no significant differences in lightness and contrast matches for both increments ($p = .494$) and decrements ($p = .842$), thereby suggesting that observers made lightness matches as if they were contrast matches, or vice versa. Therefore, all subsequent analyses will not distinguish between lightness and contrast instructions, unless otherwise noted.

Once lightness and contrast matches were combined, new average ratios were calculated. Table 3 lists these averages. Regarding increments, large articulated targets were underestimated by 2.8% and small articulated targets
Table II

*Average lightness and contrast matched ratios*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stimulus</th>
<th>Incremental Ratios*</th>
<th>Decremental Ratios*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lightness</td>
<td>Small</td>
<td>8.68:1</td>
<td>3.49:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Small</td>
<td>8.57:1</td>
<td>3.50:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Non-articulation</td>
<td>8.84:1</td>
<td>3.51:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Non-articulation</td>
<td>8.83:1</td>
<td>3.48:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Articulation</td>
<td>8.53:1</td>
<td>3.47:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Articulation</td>
<td>8.31:1</td>
<td>3.52:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Large</td>
<td>8.91:1</td>
<td>3.43:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Large</td>
<td>8.81:1</td>
<td>3.45:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Non-articulation</td>
<td>9.04:1</td>
<td>3.45:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Non-articulation</td>
<td>8.93:1</td>
<td>3.46:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Articulation</td>
<td>8.77:1</td>
<td>3.40:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Articulation</td>
<td>8.69:1</td>
<td>3.43:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Non-articulation</td>
<td>8.94:1</td>
<td>3.48:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Non-articulation</td>
<td>8.88:1</td>
<td>3.47:1</td>
</tr>
<tr>
<td>Lightness</td>
<td>Articulation</td>
<td>8.65:1</td>
<td>3.44:1</td>
</tr>
<tr>
<td>Contrast</td>
<td>Articulation</td>
<td>8.50:1</td>
<td>3.48:1</td>
</tr>
</tbody>
</table>

*All tests were not found to be statistically significant*

were underestimated by 6.2%. Observers achieved near veridicality when matching large non-articulated targets, barely overestimating by 0.1%, while underestimating small non-articulated targets by 1.6%. A planned comparison test indicated several significant differences among incremental stimuli. First, it should be noted that the physical ratios across all incremental stimuli were identical (8.98:1) and therefore comparing matched ratios is in fact meaningful,
unlike experiment one where the physical ratios differed and thus matched ratios were not used in the planned comparison tests. The same holds true for all decremental stimuli, where the physical ratio was 3.24:1.

Table III

*Average matched ratios for experiment two*

<table>
<thead>
<tr>
<th>Increments</th>
<th>Articulated Background</th>
<th>Non-articulated Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of Large and Small Target = 8.98:1</td>
<td>Observed Ratio</td>
<td>Difference from Actual Ratio</td>
</tr>
<tr>
<td>Large</td>
<td>8.73:1</td>
<td>-2.8%</td>
</tr>
<tr>
<td>Small</td>
<td>8.42:1</td>
<td>-6.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decrements</th>
<th>Articulated Background</th>
<th>Non-articulated Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio of Large and Small Target = 3.24:1</td>
<td>Observed Ratio</td>
<td>Difference from Actual Ratio</td>
</tr>
<tr>
<td>Large</td>
<td>3.42:1</td>
<td>5.6%</td>
</tr>
<tr>
<td>Small</td>
<td>3.50:1</td>
<td>8.0%</td>
</tr>
</tbody>
</table>

Regarding increments, matched ratios for large articulated targets were significantly more veridical than matched ratios for small articulated targets, $t(6392) = -5.20, p < .001$. Likewise, matched ratios for large non-articulated targets were also closer to veridicality than small non-articulated targets, $t(6392) = -2.51, p = .012$, suggesting that observers performed better with large targets than with small targets. When combining articulated and non-articulated backgrounds, the same pattern again appears, as all matched ratios for large targets (8.86:1) were closer to the physical ratio than all matched ratios for small targets (8.63:1), $t(6392) = -5.45, p < .001$. Lastly, combining small and large
targets showed a significant difference concerning articulation vs. non-articulation. All non-articulated matched ratios (8.92:1) achieved a more veridical performance than all articulated matched ratios (8.58:1), t(6392) = 7.84, p < .001. To summarize, the results suggest that observers were more accurate when matching large targets and non-articulated targets.

Regarding decrements, both large and small articulated targets were overestimated by 5.6% and 8.0%, respectively. Large and small non-articulated targets were both overestimated by 6.8% and 8.0%, respectively. However, unlike increments, planned comparison tests did not reveal any statistically significant results. Small and large articulated matched ratios were statistically non-significant, t(6392) = 1.40, p = .161, as were small and large non-articulated matched ratios, t(6392) = .65, p = .517. When combining articulated and non-articulated backgrounds, small and large matched ratios did not differ, t(6392) = 1.45, p = .147, nor did articulated and non-articulated matched ratios when combing small and large targets, t(6392) = .479, p = .632. To summarize, unlike increments, observers overestimated decremental stimuli as a group, although there were no statistical differences found between any of the decremental stimuli.

Overall, observers were more accurate with large targets than with small targets (8.86:1 and 8.63:1) for increments, but no difference in performance was found for small and large decrements. Similarly, observers were more accurate with non-articulated targets than they were with articulated targets (8.92:1 and 8.58:1) for increments, but no difference in performance was found for non-
articulate and articulated decrements. Also, observers tended to underestimate increments, but overestimate decrements. Observers were also more accurate with increments. On average, increments were found to be inaccurate by approximately 2.7%, whereas decrements were found to be inaccurate by approximately 7.10% \(t(6392) = -22.87, p < .001\). The most accurate condition was the incremental large non-articulated target (overestimation of just 0.1%). The most inaccurate condition was the decremental small articulated and non-articulated targets (both were overestimated by 8.0%).

The incremental histograms (Figures 9a, b, c, and d) show the distributions of observed ratios for non-articulated small and large targets, articulated small and large targets, as well as the combined non-articulated and articulated targets, and combined small and large targets. The x-axis represents the matched ratio, and the y-axis plots the total number of matches for a given ratio. Small target matched ratios are represented by blue bars and large target matched ratios are represented by red bars. The black arrow signifies the actual physical ratio, which was 8.98:1 for all incremental stimuli, and 3.24:1 for all decremental stimuli. As previously stated, all incremental planned comparisons revealed a significant difference between small and large matched ratios, and non-articulated and articulated matched ratios; however, descriptively analyzing the histograms reveals little to no differences, suggesting that while there are statistical differences, the effect sizes are quite small. The distributions in each of the four histograms nearly fall on top of each other and slightly to the left (i.e., underestimation) of the physical ratio of 8.98:1. The variability of each distribution
is also very similar. For example, in Figure 9d, the small target matched ratios’ distribution ranges from 1.85:1 to 16.5:1 with a standard deviation of 1.67; whereas the large target matched ratios’ distribution ranges from 1.76:1 to 16.5:1 with a standard deviation of 1.57. In Figure 9a, both the small and large target distributions have standard deviations of 1.67. In Figure 9b, the small and large target distributions have standard deviations of 1.66 and 1.47, respectively. Lastly, the articulated and non-articulated distributions in Figure 9c have standard deviations of 1.57 and 1.67, respectively.
Likewise, Figures 10a, b, c, and d illustrate the distributions for
decremental matched ratios. Unlike increments, the planned comparisons did not
find any statistically significant differences between decremental targets.
Descriptively, the decremental histograms also support this notion. Notice how
the spread of the distributions are nearly all identical. The variability of each
distribution is also narrow and quite similar. In Figure 10a, the small and large
target matched ratios’ distributions have standard deviations of 0.47 and 0.49,
respectively. In Figure 10b, the standard deviations of the small and large target
distributions were 0.56 and 0.50, respectively. In Figure 10c, the standard
deviation of the articulated target distribution is 0.53, whereas the standard

*Figure 10a –d. Histograms of decremental matched ratios distributions.*
deviation of the non-articulated target distribution is 0.47. Lastly, the combined small and large target distributions in Figure 10d both have standard deviation of 0.50.

Next, one-way ANOVAs were computed for each observer’s matches across both increments and decrements and were then plotted based on each individual’s average matched small target ratio and matched large target ratio, as shown in Figures 11a and 11b. As in experiment one, each green circle represents one observer’s observed ratios, where their matched small target ratio is plotted along the x-axis, and their matched large target ratio is plotted along the y-axis. A filled in circle indicates statistically different matched small and large target ratios. A hollow circle indicates that the observer’s matched small target ratio did not statistically differ from their matched large target ratio. The blue line designates the actual physical ratio for small targets and the red line designates the physical ratio for large targets. Note that the physical ratios, unlike experiment one, were equal for both increments (8.98:1) and decrements (3.24:1), and, thus, whether or not an observer’s matched ratios statistically differed is now in fact meaningful. A non-significant pair of ratios would indicate that the observer perceived small targets similarly to large targets, while a significant pair of ratios would indicate that the observer perceived the targets differently. A perfect ratio (i.e., veridicality) would appear at the intersection of the red and blue line. The black “veridical trendline” is also of importance now. If an observer falls on this line, this implies that they misjudged the small and large targets by an equivalent amount. If an observer falls above this line, this implies
greater small target accuracy, while an observer who falls below this line implies
greater large target accuracy.

In Figure 11a, notice that the majority of observers fall to the left of the
small target ratio (blue line) and thus are underestimating paper targets; whereas
in Figure 11b, with no exceptions, observers are falling to the right and thereby
overestimating small targets. Regarding large targets, both Figures 11a and 11b
indicate a similar pattern to the small targets, that is to say, the majority of
observers underestimated (i.e., fall below the red line) increments, but
overestimated (i.e., fall above the red line) decrements. Also, note that the
removal of the two potential outliers seen in Figure 11a do not change the
statistically significant difference found between small and large incremental
targets.
**Figure 11a.** Matched incremental small and large matched ratios for all observers.

**Figure 11b.** Matched decremental small and large matched ratios for all observers.
DISCUSSION

There were four hypotheses regarding experiment two. First, observers were expected to achieve near veridicality in their matched CRT ratios. Second, based on pilot data and the results from experiment one, observers were expected to underestimate increments but overestimate decrements. Third, according to the area rule, matched ratios for small targets should be perceived differently than large ratios. Fourth, according to Gilchrist et al. (1999) and Logvinenko (2005), matched ratios of articulated stimuli should be more veridical than non-articulated stimuli.

**Hypothesis one: Observers veridically match CRT ratios**

Observers were indeed quite accurate in their matches. Across all conditions, observers’ accuracy ranged from perfect veridicality (0.1% for non-articulated incremental large targets) to 8.0% (small articulated and non-articulated decremental targets). This is consistent with another CRT study in which observers achieved veridicality (Arend & Spehar, 1993). Compare this range to that of experiment one, where observers ranged from -9.7% to 27.1%. This suggests that a perceptual difference exists between images made of light emanating from a CRT screen and light that is reflected off of paper.
Hypothesis two: Observers underestimate increments and overestimate decrements

The predicted difference between increments and decrements was validated. With the lone exception of non-articulated large incremental targets, observers underestimated incremental targets and overestimated decremental targets (see Table 3). As for the non-articulated large incremental stimuli, observers technically overestimated the targets by 0.1%; however, given that the effect size is negligible, this result should be interpreted as achieving veridicality, not overestimation.

It was suggested in experiment one that decrements were overestimated not because there was an inherent tendency to overestimate decrements, but rather that decrements have a propensity to have small physical ratios, which was again the case in experiment two (e.g., 3.24:1). As previously stated, there are simply more possible overestimating errors to be made when a physical ratio is small, but there are few underestimating errors that can be made. Thus, observer’s underestimating errors are capped whereas errors of overestimation can reach infinity.

In keeping with previous research, as well as experiment one, incremental targets had a higher physical ratio (e.g., 8.98:1) and thus were not overestimated, but rather were slightly underestimated (Logvinenko, 2005; Riddle et al., 2010).
Hypothesis three: According to the area rule small targets should be perceived differently than large targets

Gilchrist et al.’s (1999) area rule was only partially supported by experiment two results. Regarding increments, observers averaged a matched ratio of 8.63:1 for large targets (underestimation of 1.4%) and 8.86:1 for small targets (underestimation of 3.9%). Referencing Table 3, one can see that the same pattern holds true within the articulated and non-articulated conditions as well. In both instances, the large targets were closer to veridicality than small targets. On the surface, this result seems to corroborate Gilchrist’s area rule in that small and large targets were indeed treated differently by observers. However, one needs to next consider the fact that the area rule also specifies that as a darker region increases in size, it will not only be perceived differently, but will appear lighter than it was as a smaller size. The incremental results follow this prediction. To see why this is true, recall that the darker region in the incremental condition is the background, which increases in size when moving from the large to the small target stimuli. Thus, the small target background should be perceived as lighter than the large target background. The background luminance is represented by the right side of the matched ratios. The average large target matched ratio is 8.86:1, when an observer then moves on to a screen that shows a small target, the right side of this ratio should increase, thereby making the overall ratio smaller. Which is true - the average small target matched ratio is 8.63:1.
However, the area rule under decremental conditions was ultimately not supported. Observers averaged a matched ratio of 3.50:1 for small targets and 3.44:1 for large targets (recall the physical ratio for decrements is 3.24:1). The same pattern was also true within the articulated and non-articulated conditions. According to the area rule, large target matched ratios should be less veridical than small target matched ratios. As the darker region (which is the target in decrements) increases in size it should become perceptually lighter, and because all decremental targets were overestimated, the larger decremental targets were overestimated to an even greater degree. However, differences between small and large target matched ratios were found to be statistically non-significant. Thus, although the matched ratios were found to be in the predicted direction, the differences were negligible, and as a result the area rule was not supported under decremental conditions.

Hypothesis four: Articulated stimuli should be more veridical than non-articulated stimuli

Surprisingly, articulated images resulted in less veridicality than non-articulated images under incremental conditions, despite previous research that suggests that articulated scenes serve as unconscious comparison points for an observer (Gilchrist et al., 1999; Logvinenko, 2005; Maloney & Schirillo, 2002). Observers’ average articulated matched ratio was 8.58:1, an underestimation of 4.5%. Conversely, non-articulated matched ratios averaged 8.92:1, an underestimation of just 0.75%, a difference that was statistically significant. This also holds true for both small and large target stimuli individually. So it seems in
this particular case, that articulation served to distract observer’s judgments – a surprising result indeed – rather than assist observer's judgments. It is possible that, because luminance contrast estimates on CRT screens are already accurate without articulation, the added stimuli only diverts attention away from the task and thus makes observers’ estimates less accurate, whereas with paper and spotlight stimuli, as used in Gilchrist et al. (1999) and Logvinenko (2005), articulation can serve to improve the inaccurate estimates of observers.

On the other hand, articulation was not a distraction for observers when viewing decremental stimuli. The average matched ratio for articulated stimuli was 3.46:1, an overestimation of 6.8% and for non-articulated stimuli the average matched ratio was 3.48:1, an overestimation of 7.4%. However, this difference was not statistically significant and thus it appears there was no effect of articulation present.
CROSS-EXPERIMENT RESULTS

A one-way ANOVA was employed to test the relationship between lightness and contrast matching for both increments and decrements for pilot studies one and two. In addition to one-way ANOVAs, Kruskal-Wallis tests were also used as the distributions were not expected to be normal, as was the case with the current studies of this thesis. Regarding pilot study one, when making lightness matches on articulated incremental targets, observers’ average matched ratio was 7.38:1. When making contrast matches for the same stimuli, observers’ average matched ratio was 7.95:1. Regarding decrements, observers averaged a ratio of 5.72:1 for lightness matches, and 6.01:1 for contrast matches. A Kruskal-Wallis test indicated no significant differences in lightness and contrast matches for both increments (p = .113) and decrements (p = .103), thereby suggesting that observers made lightness matches as if they were contrast matches, or vice versa.

In the second pilot study observes averaged a ratio of 6.48:1 for non-articulated incremental paper targets when making lightness matches, and 6.63:1 when making contrast matches. When viewing incremental spotlight targets, observers averaged a ratio of 7.24:1 for both lightness and contrast matches. When viewing decremental paper targets, the average matched ratio was 2.74:1 for lightness matches, and 2.84:1 for contrast matches. When viewing decremental spotlight targets, the average matched ratio was 3.44:1 for both
lightness and contrast matches. A Kruskal-Wallis test indicated no significant differences in lightness and contrast matched for both increments ($p = .177$) and decrements ($p = .335$), thereby suggesting that observers made lightness matches as if they were contrast matches, or vice versa.

Note that all four experiments, the two pilot studies and the two primary studies, did not find any statistical differences between lightness and contrast matches.

Next, the effects of target type were investigated. One-way ANOVAs were used to test for statistical differences between CRT, paper, and spotlight targets. Because the physical ratios across experiments were not always equal, the average matched ratios were not used in the one-way ANOVAs. Rather, the difference between the matched ratios and physical ratios (as measured by percentages) was used. This is a more accurate way of determining whether or not observers matched one kind of stimuli differently than another kind of stimuli.

**Increments**

First, differences between CRT targets and spotlight targets were tested. Figure 12a, b, and c show the histograms for incremental CRT and spotlight target distributions. The CRT distributions are shown in blue (the physical CRT ratio is indicated by the blue arrow) and the spotlight distributions are shown in red (the physical spotlight ratio is indicated by the red arrow). The average matched ratio for incremental small CRT (non-articulated) targets was 8.84:1 (physical ratio was 8.98:1), an underestimation of 1.6%; whereas the average matched ratio for small spotlight targets was 5.46:1 (physical ratio was 4.6:1), an
overestimation of 18.7%, $F(1, 1598) = 184.55, p < .001$. The average matched ratio for large CRT targets was 8.99:1 (physical ratio of 8.98:1), an overestimation of just 0.1%; whereas the average matched ratio for large spotlight targets was 7.24:1 (physical ratio of 9.54:1), an underestimation of 24.1%, $F(1, 1598) = 562.73, p < .001$. Lastly, the average matched ratio for large articulated CRT targets was 8.73:1 (physical ratio was 8.98:1), an underestimation of 2.8%, whereas the average matched ratio for large articulated spotlight targets was 8.94:1 (physical ratio of 15.0:1), an underestimation of 40.4%, $F(1, 1538) = 1496.76, p < .001$. 

Figure 12a – c. Histograms of incremental matched ratios distributions comparing CRT and spotlight targets. CRT distributions and corresponding physical ratios (indicated by an arrow) are represented in blue. Spotlight distributions and corresponding physical ratios (indicated by an arrow) are represented in red.
Next, incremental CRT targets were compared to incremental paper targets, both small and large. As with the CRT and spotlight histograms, CRT data is associated with the color blue, and paper is associated with the color red (Figure 13a and 13b). The average matched ratio for incremental small CRT (non-articulated) targets was 8.84:1 (physical ratio was 8.98:1), an underestimation of 1.6%; whereas the average matched ratio for small paper targets was 7.67:1 (physical ratio was 8.53:1), an underestimation of 10.1%, $F(1, 1598) = 87.85, p < .001$. The average matched ratio for large CRT targets was 8.99:1 (physical ratio of 8.98:1), an overestimation of just 0.1%; whereas the average matched ratio for large paper targets was 6.56:1 (physical ratio of 8.66:1), an underestimation of 24.2%, $F(1, 1598) = 610.55, p < .001$. Note that the CRT stimulus that contained small articulated targets does not have a comparable set of stimuli from a previous experiment, and thus no cross-experiment comparison can be made. This is true for both increments and decrements.

To summarize, observers tended to match CRT increments quite accurately compared to spotlight and paper target matches. On average, observers' CRT matches were inaccurate by approximately 1.5%. On the other hand, when matching spotlight targets observers were inaccurate by approximately 27.7%, and when matching paper targets observers were inaccurate by approximately 17.2%. Also, observers tended to slightly underestimate incremental CRT targets as well as both incremental spotlight and
paper targets. The one noticeable exception is seen in Figure 12a, where observers actually overestimated small non-articulated spotlight targets.

**Figure 13a – b.** Histograms of incremental matched ratios distributions comparing CRT and paper targets. CRT target distributions and corresponding physical ratios (indicated by an arrow) are represented in blue. Paper target distributions and corresponding physical ratios (indicated by an arrow) are represented in red.

**Decrements**

Figure 14a, b, and c show the histograms for decremental CRT and spotlight target distributions. As with increments, the CRT distributions are shown in blue (the physical CRT ratio is indicated by the blue arrow) and the spotlight distributions are shown in red (the physical spotlight ratio is indicated by the red arrow). The average matched ratio for decremental small CRT (non-articulated) targets was 3.5:1 (physical ratio was 3.24:1), an overestimation of 8.0%; whereas the average matched ratio for small spotlight targets was 4.62:1 (physical ratio was 3.64:1), an overestimation of 26.9%, $F(1, 1598) = 186.23, p < .001$. The average matched ratio for large CRT targets was 3.46:1 (physical ratio of 3.24:1), an overestimation of 6.8%; whereas the average matched ratio for large non-articulated spotlight targets was 3.44:1 (physical ratio of 3.2:1), an overestimation
of 7.5%, $F(1, 1598) = .389, p = .533$, a statistically non-significant difference.

Lastly, the average matched ratio for large articulated CRT targets was 3.42:1 (physical ratio was 3.24:1), an overestimation of 5.6%, whereas the average matched ratio for large articulated spotlight targets was 5.86:1 (physical ratio of 15.0:1), an underestimation of 60.9%, $F(1, 1538) = 6203.29, p < .001$.

**Figure 14a – c.** Histograms of decremental matched ratios distributions comparing CRT and spotlight targets. CRT target distributions and corresponding physical ratios (indicated by an arrow) are represented in blue. Spotlight target distributions and corresponding physical ratios (indicated by an arrow) are represented in red. Note that the CRT and spotlight physical ratios in Fig. 14b are so close in value (3.24:1 and 3.20:1, respectively) that only one arrow is shown.

Next, decremental CRT targets were compared to decremental paper targets, both small and large (Figure 15a and 15b). As with the CRT and spotlight histograms, CRT data is associated with the color blue, and paper is associated
with the color red. The average matched ratio for decremental small CRT (non-articulated) targets was 3.5:1 (physical ratio was 3.24:1), an overestimation of 8.0%; whereas the average matched ratio for small paper targets was 3.16:1 (physical ratio was 2.78:1), an overestimation of 13.7%, \( F(1, 1598) = 47.48, p < .001 \). The average matched ratio for large CRT targets was 3.46:1 (physical ratio of 3.24:1), an overestimation of 6.8%; whereas the average matched ratio for large non-articulated paper targets was 2.79:1 (physical ratio of 2.85:1), an underestimation of 2.1%, \( F(1, 1598) = 80.49, p < .001 \).

\[ \text{Figure 15a – b. Histograms of decremental matched ratios distributions comparing CRT and paper targets. CRT target distributions and corresponding physical ratios (indicated by an arrow) are represented in blue. Paper target distributions and corresponding physical ratios (indicated by an arrow) are represented in red.} \]

To summarize, observers tended to match CRT decrements slightly more accurately compared to paper target matches, and especially compared to spotlight target matches. On average, observers’ CRT matches were inaccurate by approximately 6.8%, a relatively small amount. On the other hand, when matching spotlight targets observers were inaccurate by approximately 31.8%, and when matching paper targets observers were inaccurate by approximately
7.9%. Also, observers tended to slightly overestimate decremental CRT targets as well as both decremental spotlight and paper targets. There are two exceptions. In Figure 14c, observers actually underestimated large articulated spotlight targets by a large margin, 60.9%. Secondly, in Figure 15b, large non-articulate paper targets were slightly underestimated by 2.1%.
GENERAL DISCUSSION

Effects of stimuli type on luminance contrast estimates

There is ample evidence suggesting that the type of stimuli used is critically important when making luminance contrast estimates, whether it be reflectance or illumination edges (Gilchrist et al., 1983), CRTs (Arend & Spehar, 1993), or even materials such as wood or flannel (Beck, 1961). It is widely regarded that the absolute level of luminance is only one of many determinants of estimating lightness or perceived illumination. The current studies examined three kinds of stimuli. First, in pilot data and experiment one, observers viewed ambient light that was reflected off of a piece of paper. Second, those same observers also viewed a spotlight, emanating from a projector, which was reflected off a piece of paper. Third, in experiment two, observers viewed images originating from a direct light source, a CRT monitor. In experiment one, as well as pilot studies, average matched spotlight ratios were clearly inaccurate, ranging from overestimating the spotlight by 7.5% to underestimating the spotlight by 60.9%. Average matched paper ratios were likewise inaccurate, but much less so compared to spotlights. Across both pilot studies and experiment one, these ratios ranged from underestimating the paper target by 2.1% to underestimating the paper target by 24.2%. In experiment two, observers achieved near veridicality in many cases and ranged from overestimating the CRT target by just 0.1% to overestimating the CRT target by 8.0%.
Figure 16 shows this striking trend. Across experiment one and two, as well as pilot studies, observers inaccurately matched spotlight targets by a total of 29.7%; observers inaccurately matched paper targets by 12.6%; and observers inaccurately matched CRT targets by 4.8% (note that errors of over or underestimation are not reported here, only the degree of inaccuracy, thus absolute values were used in Figure 16). This pattern validates previous research that suggests that stimuli type has a strong influence over luminance contrast estimates. Why the luminance contrasts of some surfaces are easier to estimate than other surfaces is not yet known.

One notion is that as the original light goes through more and more known alterations, it becomes harder and harder to accurately judge. For instance,
imagine the setup in experiment one: light emanates from a projector, along with a small amount of ambient light from the room, and hits a black piece of paper. Some portion of that light is then reflected off the paper, say 10%, and reaches the retina of an observer, who perceives the light as a small spotlight on a larger piece of paper. In order to accurately judge the lightness or the illumination of this spotlight area, the observer’s visual system must somehow be able to judge two separate pieces of information (i.e., the level of illumination and the reflectance of the paper). Now imagine experiment two, where the observer gets to bypass any reflection of light, but rather obtains the light directly from the light source, without any alterations. Referencing Figure 1, this difference is akin to drawing a line straight from the sun to the observer’s eye, instead of drawing a line from the sun to the paper, and then to the eye. Simply, as light travels a more complex path, more errors may be made.

This is a rather simplistic view of the problem, but helps illustrate the important distinction between viewing surfaces in a simple controlled environment, such as in a dome (Gilchrist et al., 1999) or on a CRT monitor (Arend, 1993), and the highly complex images seen in everyday life.

**Effects of size on luminance contrast estimates**

Although not intuitive, previous research has demonstrated that stimulus size matters. As darker regions increase in size, observers perceive those regions as increasing in lightness (Gilchrist et al., 1999). The results from the current studies show that support for Gilchrist’s area rule was inconsistent across the two experiments. In particular, experiment one found support for the area rule
only for decremental conditions, but support for the area rule in experiment two was found only for incremental conditions. While this suggests that the area rule can be used as an explanation across multiple conditions, it raises more questions than it answers. Primarily, under which conditions does the area rule fail and why?

Regarding both incremental and decremental conditions, it is important to take into consideration a difference in stimuli between the current studies and that of Gilchrist et al. (1999). Gilchrist's stimuli did not employ the traditional target and background display, but rather used a large acrylic hemisphere that had a two-part pattern painted on its surface. In this instance, the observer can give no more attention to one area over another; the only difference between the two areas are their reflectance and size. Additionally, Gilchrist’s observers were instructed to match both regions, not just one. In the current studies, there are three differences that the observer can attend to: reflectance, size, and target/background distinction. Because of the target/background setup, as well as the fact that observers were instructed to attend to and make matches of the target, equal consideration (consciously and unconsciously) was not given to both areas. This may be a crucial caveat in the area rule: as a darker region increases in size, it will appear lighter, so long as all other properties are equal. Thus, for the area rule to apply, observers should not be told to attend to one area more than another, and all other properties that may unconsciously lead an observer to focus on one area should be removed, such as a display where one area constitutes a foreground and the other area constitutes a background.
Moreover, in both experiments one and two, there was a third and even a fourth area of the environment that could have canceled out area effects. In experiment one the stimulus was attached to a large sheet of black felt that was attached to the wall. This border of felt was rather large (17.10° X 37.21° of visual angle) compared to the actual stimulus display. Recall that Gilchrist had observers place their heads in a large dome where the entire visual field was made up of a dark and a lighter surface. In experiment one, observers were placed in a typical room that contained everyday objects such as a large filing cabinet, a desk, light fixtures, etc., all of which may have reduced or eliminated any effects of size. In experiment two, there was a black border that surrounded the display presented on the CRT monitor to match the black felt in experiment one, although this surrounding border was narrow (2.58° X 32.01° of visual angle) compared to the CRT stimulus display. In addition, all of the objects present in the room for experiment one were still in plain sight for experiment two. Keep in mind that this was done intentionally in order to investigate real world effects. The area rule may apply under strict experimental and highly simulated conditions (i.e., a hemispheric dome), but in everyday scenes effects of size do not seem to occur. Although it is possible that highly complex area effects were present in experiments one and two, any such effects were not intended to be investigated.

Taken as a whole, the results from both experiments suggest that effects of stimuli size in a complicated real world scene are either non-existent, or are small and easily drowned out by other more dominating effects.
Effects of articulation on luminance contrast estimates

In experiment two, articulated displays seemed to either impair observer’s estimates (as was the case for increments) or have no effect whatsoever (as was the case for decrements). Overall, these results do not support previous research that suggests that articulation can improve observer’s luminance contrast estimates (Gilchrist et al., 1999; Logvinenko, 2005; Maloney & Schirillo, 2002). There may be two explanations for this result.

First, Gilchrist et al. (1999) investigated articulation effects across several different types of displays, all of which had the articulated squares border each other (i.e., there were no gaps between the articulated squares). This is in contrast to experiment two, where articulated squares were separated so that the background was visible on all four sides of the square. This difference, which minimizes the possibility of taking local contrast ratios across the articulated patches, may be enough to explain why articulation did not have the predicted effect in the current study.

Second, in both Gilchrist et al. (1999) and Logvinenko (2005), articulated displays consisted of paper, while experiment two was comprised of images on a CRT monitor. As previously shown, luminance contrast estimates are quite accurate when matching CRT targets and, thus, any articulation effect would have to be very strong in order to influence observer’s CRT estimates. At best, a small articulation effect is drowned out by other influences that make observers accurate in their judgments, or at worst, articulation acts as a small distractor when making their judgments. Overall, these results suggest that the current
articulation display, consisting of small separated squares, does not act to improve CRT luminance contrast estimates. Further research is needed to investigate what kind of articulation displays, if any, would improve CRT performance.
CONCLUSION

This thesis clarifies three important effects regarding lightness perception. First, and most importantly, the combined results from experiment one and two strongly indicate that luminance contrast estimates vary as a function of stimuli type. Second, hypotheses concerning Gilchrist's area rule (Gilchrist et al., 1999) for complex displays were rejected overall. Third, the particular type of articulation used did not improve observers’ performances when viewing CRT displays.

Effects of stimuli type

Observers’ performance steadily approached veridicality as the displays being judged moved towards direct light, with no alterations. Thus, performance was best when viewing a CRT monitor. The light that reaches the retina in this case can be said to equal the original light, which can be designated by $L_0$ (where $L$ represents light, and 0 indicates the light has undergone zero alterations). On the other hand, judging the lightness of an illuminated piece of paper (i.e., a spotlight) requires light to be reconstructed at least once. Some percentage of the original light is absorbed by the paper and some percentage is reflected off of the paper and into the retina. This progression could then be represented by the symbol $L_1$, indicating that the original light, or luminance, has been altered once.
The theory put forth in this thesis is that as \( L_n \) increases, errors of lightness constancy should also increase. However, there is an important caveat: not all alterations of light have an equal effect. For instance, the paper and spotlight conditions in this thesis both underwent one alteration \((L_1)\), but paper conditions resulted in greater performance than did spotlight conditions. This is primarily because of errors associated with illumination discounting (Logvinenko, 2005). Notice that both conditions are essentially equal in terms of what they are composed of – both contain illumination falling upon a piece of paper – but illumination makes up a far greater percentage of the spotlight condition than it does for the paper condition, which is only minimally illuminated. Thus, the weighting of each alteration also needs to be taken into account, which can be represented by \( w \); so that both the paper and the spotlight conditions are indicated by \( L_1(w) \), where \( w \) is greater for the spotlight than it is for the paper.

In theory, in terms of lightness constancy, \( L_{0(0)} \) would be greater than \( L_{1(1)} \), which would be greater than \( L_{1(2)} \), which would be greater than \( L_{2(1)(1)} \), etc. This is a testable hypothesis. An experiment could be designed with identical conditions (identical in luminances, sizes, displays, etc.), using CRT monitors \((L_{0(0)})\), papers \((L_{1(1)})\), spotlights \((L_{1(2)})\), and translucent mirrors \((L_{2(1)(1)})\). Note that conditions with say three, four, or even more light alterations could technically be used, although three levels should be sufficient. This hypothetical experiment would help illuminate one reason why errors of lightness constancy differ across stimuli.
Effects of stimuli size

In both experiment one and two, hypotheses regarding stimuli sizes were only partially supported. It was concluded that size effects were not obtained because of differences in stimuli display (such as target/background distinctions), instructions, and complexity (i.e., the number of different surfaces visible). However, experiments that focused solely on size effects could greatly clarify when size effects are present or absent.

One could easily conduct an experiment in which stimuli ranged from simple to complex. Recall that Gilchrist et al. (1999) used simple stimuli composed of two surfaces within a hemispheric dome, whereas experiment one and two used a target/background display, with an additional 40 matching reflectance squares, in addition to other surfaces present in the immediate surroundings of the display. Stimuli displays that range from the simple to these more complex displays could be created in future experiments. It would also be important to use a combination of simple two-part surfaces with no clear foreground and background, as well as target/background displays, to define what effect, if any, the presence of a foreground and background has on size effects.

Effects of articulation

Articulation effects have historically been underrepresented in lightness perception research (Maloney & Schirillo, 2002). Unfortunately, the results from experiment two shed little light on the question of whether or not articulation improves luminance contrast estimates. The hypothesis that articulated images
would be more accurate than non-articulated images was not only rejected, the reverse was true; non-articulated images received more accurate estimates than did articulated images. Two potential explanations for this result were put forth. First, the arrangement of articulated squares (the squares were separated by small background gaps) differed from previous arrangements found in other studies (Gilchrist et al., 1999; Maloney & Schirillo, 2002). Second, because CRT performance was largely accurate even without articulation, the added stimuli may have served as a distractor. Future research should be able to determine under which conditions articulation has an effect by adjusting the configuration of the articulated squares, from those found in Gilchrist et al. (1999) to those in Logvinenko (2005) and the current study, as well as adjusting the type of stimuli used (i.e., paper, CRT, etc.).

Limitations

One potential limitation in this thesis is the fact that the physical ratios used in both experiments one and two were not equivalent. For example, incremental paper targets in experiment one had a physical ratio of 8.53:1 and decremental paper targets had a physical ratio of 2.78:1. While this disparity between physical ratios is sometimes irrelevant and other times advantageous, it can also be a limitation. For example, a difference found between average matched paper and spotlight ratios could be explained by the differences in physical ratios. That is, a 5% underestimation of a 3:1 ratio may not be equivalent to a 5% underestimation of a 10:1 ratio. However, it was helpful to have non-matching physical ratios across different conditions because it can
reveal other effects, such as the effect that the absolute luminance contrast ratios have on the direction of lightness constancy errors. Specifically, small physical ratios tend to be overestimated and large ratios tend to be underestimated, regardless of whether the ratio was a decremental or incremental display.

Implications

The most important implication of this thesis comes from understanding more regarding the effects of stimuli type on luminance contrast and lightness estimates. Future researchers should take note that whether their observers match CRT targets, paper targets, etc., the type of stimuli that they decide to use can have a strong impact on lightness constancy. This is particularly important in the current state of lightness perception studies, as those who use predominantly CRT stimuli (Arend & Spehar, 1993) can and do draw very different conclusions than those who predominantly use real surfaces and light (Gilchrist et al., 1999; Logvinenko, 2005). Although experiments using CRTs are convenient and more precise, greater consideration should be given to experimental conditions that employ real surfaces (i.e., paper) and light. Even in these modern times, the human visual system still works in a world which is mostly made up of real objects, as opposed to objects on a television or computer screen.
REFERENCES


APPENDIX A

(Room Pictures)

Experiment one

Experiment two
APPENDIX B

(Experiment one)

General instructions

“Here you can see that this region is brighter (or darker) than this adjacent region. That is because this part is illuminated {for spotlight conditions only} by the projector spotlight. Notice the edge here. It is created by the difference in illumination.”

Contrast matching instructions

“Your task is to choose a square from the left {or right} hand side of the upper array of squares, which has the same contrast to its background as the large rectangular spotlight/paper has to its background.”

Lightness matching instructions

“Your task is to choose a square from the left {or right} hand side of the upper array of squares, which has the same lightness as the large rectangular spotlight/paper. In other words, please choose a square which looks as if it is made of the same paper as the spotlight/paper.”
APPENDIX C
(Experiment two)

Contrast matching instructions

“Your task is to choose a square from the left (or right) hand side of the upper array of squares, which has the same contrast to its background as the large or small target has to its background. In other words, look at the edge here and compare the contrast of the target to its background, and choose the square that has the same contrast to its background. Lastly, on some trials, you will notice several additional squares underneath the bottom target; you should not match any of these squares to the target. Use only the top squares to make your matches.”

Lightness matching instructions

“Your task is to choose a square from the left (or right) hand side of the upper array of squares, which has the same lightness as the large or small target. In other words, please choose a square, which looks as if it is made of the same shade as this area. Lastly, on some trials, you will notice several additional squares underneath the bottom target; you should not match any of these squares to the target. Use only the top squares to make your matches.”
**Anchor**: a property of a scene that ties relative luminance values to specific values of perception, such as lightness values (Gilchrist et al., 1999).

**Articulation**: the presence of multiple shades of gray surfaces that can be used as a comparison in judging the lightness of a given surface. See the “Increment, Spotlight” and the “Decrement, Spotlight” picture in Figure 4 for an example of an articulated display.

**Brightness**: the subjective perception of the total amount of light of a given region.

**CRT**: cathode ray tube.

**Decrement**: a target whose lightness is darker than its background. See the two “Decrement” pictures in Figure 4 for an example.

**Illumination**: the total amount of light given off from a light source, such as a light bulb.

**Increment**: a target whose lightness is lighter than its background. See the two “Increment” pictures in Figure 4 for an example.

**Lightness**: the perceived reflectance of a surface.

**Lightness constancy**: the ability to have a surface appear to reflect the same proportion of light independent of the level of illumination falling upon it.

**Luminance**: the total amount of light that reaches the retina from a given source.
Luminance contrast: either a physical or perceptual term that refers to the ratio of the luminance of a target to the luminance of a background.

Mondrian stimulus: a patchwork of gray rectangles of varying luminances, named after the artist Piet Mondrian. See Figure 2 for an illustration.

Munsell chart: an array of patches increasing in reflectance from white to black used to make lightness matches of a target.

Perceived illumination: the subjective total amount of light that falls upon a surface or object.

Perceptual scission: the separation of the lightness of a surface from the illumination falling onto that surface.

Reflectance: the ratio of the amount of light that is reflected off of a surface to the total amount of light that falls upon the surface.

Simultaneous contrast: the occurrence of a surface seeming to change its appearance based on its surroundings. For example, a gray target with a black background will appear lighter than the same gray target with a white background (Blake & Sekuler, 2006).

Veridicality: the accurate perception of an object’s properties.
CIRRICULUM VITAE

Matt Riddle
2700 Reynolda Road, apt. 612
Winston-Salem, NC 27106
412-480-2330
riddm9@wfu.edu

Undergraduate Education:
Concentration: Psychology
Research Project: A Study of Reading Accuracy in Colorblindness
Community College of Alleghany County, Pittsburgh, PA, 2004.
Indiana University, Bloomington, IN, 2002-2004.

Graduate Education:
M.A. Psychology, Wake Forest University, Winston-Salem, NC, 2009 - Present

Relevant Coursework:


Mathematics: Research Design and Analysis in Psychology, Precalculus Mathematics, Calculus I, Probability and Statistics
Work Experience:

2004 – 2006: Internship at DDI (Development Dimensions International), an HR consulting firm

2006 – 2008: Project Associate for DDI Manhattan’s office.

Presentations:


Publications:


Honors:

Awarded Dean Scholarship to Indiana University (2002).

Other Skills:

Extensive knowledge of Microsoft Outlook, Word, PowerPoint, and Excel. Knowledge in SPSS.